Group member: George Wilfert, Nick Tran

Topic: We will focus mainly on modules 1 using Machine Learning for Bioinformatics with the goal of gaining a better understanding of the mechanisms of Alzheimer's disease.

Our focus and goal:

Our focus will be on Alzheimer's disease. We will be using public biomedical data which will be heavily focused on the dataset from UCI. We have also found other sources of data from NCBI, GEO, and Metabolomics Workbench that we can utilize for different sources to reference. Within these cases, there will be use of tools to study patterns that might be related to the cause of Alzheimer's. We will apply machine learning techniques, like the classification methods we learned in class. Specifically we want to look more into biological mechanisms that drive this disease such as key focus areas that have to do with Amyloid beta and Tau protein dysfunction. There are also a lot of studies pointing to a genetic component to the disease that would give us the ability to dive deeper into the second module of the class, understanding the gene sequencing from a RNA and DNA standpoint.

Technical Skills dive into:

First, Nick will plan to take all the features from the genetic data from a dataset Next, he will use these features to train a classifier using Python. distinguish sequences that may be associated with Alzheimer's disease from those that are not. In parallel, run BLAST searches to compare our unknown sequences to known Alzheimer's-related genes, which will help validate the machine learning model and improve our understanding of the genetic similarities and differences. Geroge will generate a predictive model and a detailed analysis that could contribute to understanding the molecular mechanisms behind Alzheimer's disease. There is also a focus on looking at the phenotypes and being able to determine statistically significant relationships between control and Alzheimer's patients in a similar fashion to the gene expression in space flight examples we did in class.