Member: Nichollette Acosta(800829090)

My project will apply data mining methods to predict secondary structures of proteins. Protein prediction is important for drug design and the design of novel enzymes. Data mining proves to be useful by helping to predict parameters related to the secondary protein structure. These parameters are the alpha helix, beta sheets and hairpin turn.

My data source will be from NCBI database. <http://www.ncbi.nl.nih.gov/structure/>. Data mining methods that may be involved are clustering, supervised and unsupervised analysis, classification, or similarity analysis. Upon completion of the project I expect to have an intermediate understanding of data mining in protein prediction, and have more inquiries to look into in biostatistics.