Ian Klasky

Capstone 2 Proposal

I'd like to predict diabetes in patients in order to decrease costs to all stakeholders in our current healthcare system. I'd specifically plan to only use behavioral features of patients, as these could be easier to gauge than by taking medical samples and measurements. If a non-diabetic patient starts exhibiting behaviors associated with diabetic patients, this model might indicate that behavioral changes might need to be made in order to avoid diabetes. This could also help healthcare providers better anticipate the number of diabetes patients they'll have, shrinking the variance of costs associated with treating the disease.

The data I'll be utilizing comes from the 2013-2014 National Health and Nutritional Examination Survey, provided by the CDC. Over 10,000 people in the U.S. were surveyed and each asked around 1,000 questions. One of the questions asked if the patient has ever had diabetes. The response to this question will be my target variable. The data can be found here: https://wwwn.cdc.gov/nchs/nhanes/Search/DataPage.aspx?Component=Questionnaire&CycleBeginYear=2013