**Capstone Proposal**

**Problem Statement:**

Using Machine Learning to help predict Heart Attack using some risk indicators.

**Context:**

According to the American Heart Association, Cardiovascular disease (CVD), listed as the underlying cause of death, accounted for 874,613 deaths in the United States in 2019. It also claims more lives each year in the United States then all forms of cancer and Chronic Lower Respiratory Disease combined. CVD accounted for 12% of total US health expenditures in 2017 and 2018 and accounted for approximately 19.05 million global deaths in 2020.

**Criteria for Success:**

Predict heart attack using some risk indicators

**Constrains within Solution Space:**

* The dataset contains 303 patients with 14 variables
* All the data store as Int or float datatype
* There is no missing data

**Stakeholders:**

* Kaggle: <https://www.kaggle.com/code/bucktuddrussell/heart-attack-prediction/data> provided dataset
* patient with or without heart problems

**Key Data Sources:**

Kaggle: <https://www.kaggle.com/code/bucktuddrussell/heart-attack-prediction/data>

List of Variables:

|  |  |
| --- | --- |
| Age | Age of the patient |
| Sex | Sex of the patient |
| exng | Exercise induced angina (1= yes, 0=no) |
| Caa | Number of major vessels (0-3) |
| Cp | Chest pain type chest pain type (0= typical angina, 1=atypical angina, 2=non-anginal pain, 3=asymptomatic) |
| trtbps | Resting systolic blood pressure (in mm Hg) |
| Chol | Cholestoral in mg/dl fetched via BMI sensor |
| fbs | Fasting blood surgar >120 mg/dl, 1=true, 0= false |
| rest\_ecg | Resting electrocardiographic results, 0=normal, 1= having ST-T wave abnormality (T wave inversions and/or ET elevation or depression of > 0.05 mV) |
| thalachh | Maximum heart rate achieved |
| oldpeak | Previous peak |
| Slp | Slope |
| thal | Thallium stree test result (0-3) |
| Output | 0 = less chance of heart attack, 1=more chance of heart attack |