Correlation Between Heart Activity, Certain Biomedical Traits and Heart Disease

Introduction

Heart disease is the leading cause of death in both men and women. 647,000 people die of heart disease every year. In america about 805,000 people have a heart attack every year and 605,000 of these are a first heart attack. And about 1 in 5 heart attacks are silent so the damage has been done without the person being aware of it.

In this study we will look into correlations between heart trends, and or other biomedical traits that can help us predict the presence of heart disease, also we would look into the trends for gender and age in relation to heart disease. If there is a correlation between some heart activities and, or other biomedical traits this study can be used to help with heart disease prevention. As a high risk patient including men, the elderly, individuals with diabetes and persons with family members who have suffered from heart disease, early detection and prevention is important as heart disease is a silent threat and shows no symptoms before occurrence.
Prevention is paramount because in the event of a heart attack if the any of the heart muscles are damaged they cannot regrow, also if any of the valves in the heart become stiff and calcified there is no way to restore the flexibility and it must be replaced or repaired.

Benefits of the study

Currently cardiovascular disease in the costliest disease in america with the price tag of $555 million in year 2016. On Average an employee with cardiovascular disease costs their employer $1,100 more per year due to loss in productivity. If the study can help with prediction of heart disease, preventive measures can be taken to reduce cost.
The data used will be provided by kaggle and UCI machine learning repository.

Method of Solving

- Looking through the data set to ensure that is clean
- Changing the categorical values to 0 and 1 to better analyse the data
- Running 4 machine learning algorithms on the data to see which factors seem to play a more important role in the prediction of the disease.
- Finding the most accurate of the 4 methods and comparing the most important factors from the machine learning to the target values (either a 0 or 1, where 1 is having heart disease and 0 is not having heart disease)
- The comparisons will include comparing one factor to the target value, comparing multiple factors to the target value and comparing of some significant factors to themselves to check for high correlation.

Deliverables

- Machine learning code
- Diagrams
- Pdf

References:
https://www.cdc.gov/heartdisease/facts.htm