Ajit Kolekar

DSC 530 - Final Project

Heart Failure Prediction

Date - 11/22/2020

**Problem Statement:** Determine factors that could contribute to the death by heart failure using historical data and to predict how the death by heart failure can be avoided by taking necessary steps.

**Research Questions:**

* What impact does age have on the risk of death due to heart failure?
* Does diabetes contribute to death due to heart failure?
* How does ejection fraction relate to the heart’s health?
* Does high blood pressure contribute to death due to heart failure?
* How does serum creatinine relate to the heart’s health?

After doing further analysis with the data, I decided to use following five variables for my exploratory data analysis: age, diabetes, ejection\_fraction, high\_blood\_presure, and serum\_creatinine.

Based on my general understanding, I believed that old people, people with diabetes, and people with high blood pressure would have more chances of death due to heart failure. I also picked two variables that could impact heart health in medical terms. My exploratory data analysis conducted through logistic regression model indicated that diabetes and high blood pressure did not impact the death due to heart failure. Age and Ejection Fraction had a minimal impact and Serum Creatinine had significant impact on the death by heart failure.

One of the limitations of my analysis was the lack of data available. The dataset included only 299 rows, but this was the most appropriate dataset for my analysis. It would be great if more data is collected and made available for research, so that better analysis can be conducted.

Another limitation of my analysis was that the available data was of patients who have already had a heart incident. The analysis would help doctors to elongate the life of such patients by looking at the factors. Due to this limitation, the outcomes of my analysis make sense.

Several people who are at high risk of getting heart attack do not know that there are ways that can be used to prevent the heart attack. Collecting data of such people that have not had a heart incident yet, and conducting analysis on the data will not only benefit the people at risk, but also the overall community.