**Summary and findings of the Heart disease dataset**

Heart disease is one of the biggest causes for morbidity and mortality among the population of the world. Prediction of cardiovascular disease is regarded as one of the most important subjects in the section of clinical data analysis. The amount of data in the healthcare industry is huge. Data mining turns the large collection of raw healthcare data into information that can help to make informed decision and prediction. it is difficult to identify heart disease because of several contributory risk factors such as diabetes, high blood pressure, high cholesterol, abnormal pulse rate and many other factors. Due to such constraints scientists have turned towards modern approaches like Data Mining and Machine Learning for predicting the disease.

 In this data science project, I will be applying Machine Learning approaches for classifying whether a person is suffering from a heart disease or not, using one of the most used datasets, which is the Cleveland Heart Disease dataset from the UCI Repository.

First, data is divided into two parts using component splitting. In this experiment, data is split based on a ratio of 80:20 for the training set and the prediction set. The training set data is used in the logistic regression component for model training, while the prediction set data is used in the prediction component.

The two inputs of the prediction component are the model and the prediction set. The prediction result shows the predicted data, actual data, and the probability of different results in each group.

In summary this This dataset is old and small by today's standards. However, it did allow to create a simple model and then use various machine learning tools and techniques to look inside. At the beginning I thought that factors such as cholesterol and age would be major factors in the model. This dataset didn't show that. Instead, the number of major factors and aspects of re sting ECG  (displays resting electrocardiographic results) dominated. I actually feel like I've learnt a thing or two about heart disease!

I suspect this sort of approach will become increasingly important as machine learning has a greater and greater role in health care.