LITERATURE SURVEY - Visualization and Prediction of Heart Diseases Using Data Science Framework

Heart diseases leading most causes of death globally according to World Health Organization cardiovascular or all heart related disease are responsible for 17.9 million death every year. An early detection and diagnosis of the disease is very important and maybe it's the key of cure. Heart is one the most vital organ in the human body. When we talk about heart diseases, we can have multiple conditions where heart is not working the way it should be like blockage in blood vessels. According to many researches that have been conducted through a period of time have found out that heart failure and heart disease has been the cruel cause of death in human beings. What aggravates this situation is that most of these diseases are being diagnosed at later stages at which it is very difficult to control. But if somehow, we can diagnose these diseases at its early stage, then we can surely cure the disease.

The main aim of this paper is to use various classification algorithms of data science framework to somehow detect the chances of having a heart disease. Also, the main aim of this research paper is to find out the most efficient classification algorithm that can help us to detect heart diseases at early stage. this algorithm can be used on heart records of the patient or by using it on classification reports. This research was conducted and tested upon various algorithms to test its accuracy like Logistic Regression, Random Forest, Vector Support and XG-Boost. After applying these algorithms of prediction model has been developed

Shouman, M., Turner, T. and Stocker, R. (2012) Using Data Mining Techniques in Heart Disease Diagnosis and Treatment

The availability of huge amounts of medical data leads to the need for powerful data analysis tools to extract useful knowledge. Researchers have long been concerned with applying statistical and data mining tools to improve data analysis on large data sets. Disease diagnosis is one of the applications where data mining tools are proving successful results.

Patel, J., Upadhyay, P. and Patel, D. (2016) Heart Disease Prediction Using Machine learning and Data Mining Technique

Heart disease is the main reason for death in the world over the last decade. Almost one person dies of Heart disease about every minute in the United States alone. Researchers have been using several data mining techniques to help health care professionals in the diagnosis of heart disease. However using data mining technique can reduce the number of test that are required. In order to reduce number of deaths from heart diseases there have to be a quick and efficient detection technique. Decision Tree is one of the effective data mining methods used. The goal of this study is to extract hidden patterns by applying data mining techniques, which are noteworthy to heart diseases and to predict the presence of heart disease in patients where this presence is valued from no presence to likely presence.

Hazra, mandal s, gupta, a. and mukherjee, a. (2017) heart disease diagnosis and prediction using machine learning and data mining techniques.

A popular saying goes that we are living in an "information age". Terabytes of data are produced every day. Data mining is the process which turns a collection of data into knowledge. The health care industry generates a huge amount of data daily. However, most of it is not effectively used. Efficient tools to extract knowledge from these databases for clinical detection of diseases or other purposes are not much prevalent. The aim of this paper is to summarize some of the current research on predicting heart diseases using data mining techniques, analyse the various combinations of mining algorithms used and conclude which

technique(s) are effective and efficient. Also, some future directions on prediction systems have been addressed.

Kumari, M. and Godara, S. (2011) Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction

The paper investigates the powerful of hybridizing two computational intelligence methods viz., Gray Wolf Optimization (GWO) and Artificial Neural Networks (ANN) for prediction of heart disease. Gray wolf optimization is a global search method while gradient-based back propagation method is a local search one.

The results demonstrate that the proposed model increases the convergence speed and the accuracy of prediction.

Masethe, H. and Masethe, M. (2014) Prediction of Heart Disease Using Classification Algorithms.

The heart disease accounts to be the leading cause of death worldwide. It is difficult for medical practitioners to predict the heart attack as it is a complex task that requires experience and knowledge. The health sector today contains hidden information that can be important in making decisions. Data mining algorithms such as J48, Naive Bayes, REPTREE, CART, and Bayes Net are applied in this research for predicting heart attacks. The research result shows prediction accuracy of 99%. Data mining enable the health sector to predict patterns in the dataset.