Visualizing and Predicting Heart Diseases with an Interactive Dash Board

TEAM LEADER: MUKUNTHAN B

TEAM MEMBER: MATHAN C

TEAM MEMBER: SRI VIJAY B

TEAM MEMBER: SASIKUMARAN T

LITERATURE SURVEY

1. Heart Disease Prediction using Exploratory Data Analysis

Healthcare industries generate enormous amount of data, so called big data that accommodates hidden knowledge or pattern for decision making. The huge volume of data is used to make decision which is more accurate than intuition. Exploratory Data Analysis (EDA) detects mistakes, finds appropriate data, checks assumptions and determines the correlation among the explanatory variables. In the context, EDA is considered as analysing data that excludes inferences and statistical modelling. Analytics is an essential technique for any profession as it forecast the future and hidden pattern. Data analytics is considered as a cost effective technology in the recent past and it plays an essential role in healthcare which includes new research findings, emergency situations and outbreaks of disease. The use of analytics in healthcare improves care by facilitating preventive care and EDA is a vital step while analysing data. In this paper, the risk factors that causes heart disease is considered and predicted using K-means algorithm and the analysis is carried out using a publicly available data for heart disease. The dataset holds 209 records with 8 attributes such as age, chest pain type, blood pressure, blood glucose level, ECG in rest, heart rate and four types of chest pain. To predict the heart disease, K-means clustering algorithm is used along with data analytics and visualization tool. The paper discusses the pre-processing

methods, classifier performances and evaluation metrics. In the result section, the visualized data shows that the prediction is accurate.

2. Analysis of Various Heart Disease Prediction Techniques

Today, heart diseases have become one of the leading causes of deaths in nationwide. The best prevention for this disease is to have an early system that can predict the early symptoms which can save more life. Recently research in data mining had gained a lot of attention and had been used in different kind of applications including in medical. The use of data mining techniques can help researchers in predicting the probability of getting heart diseases among susceptible patients. Among prior studies, several researchers articulated their efforts for finding a best possible technique for heart disease prediction model. This study aims to draw a comparison among different algorithms used to predict heart diseases. The results of this paper will helps towards developing an understanding of the recent methodologies used for heart disease prediction models. This paper presents analysis results of significant data mining techniques that can be used in developing highly accurate and efficient prediction model which will help doctors in reducing the number of deaths cause by heart disease. The prediction analysis is the technique of data analytics which can predict further possibilities based in the current information. This research work, is based on heart disease prediction in data mining. The dataset prediction data set has 13 number of attributes for the heart disease prediction. In the previous research work, the SVM classifier is applied for the heart disease prediction. Due to large number of attributes in the dataset, SVM classifier is not able to classify all the attributes due to which accuracy is low for the prediction. In this paper, various techniques for the heart disease prediction are reviewed in terms of certain parameters.

3. Prediction of heart disease at early stage using data analytics:

The various technologies of data Analytics (DA) models for forecast of heart disease are discussed. Data analytics plays an important role in building an intelligent model for medical systems to detect heart disease (HD) using data sets of the patients, which involves risk factor associated with heart disease. Medical practitioners can help the patients by predicting the heart disease before occurring. The large data available from medical diagnosis is analyzed by using data analytics tools and useful information known as knowledge is extracted. Analytics is a method of exploring massive sets of data to take out patterns which are hidden and previously unknown relationships and knowledge detection to help the better understanding of medical data to prevent heart disease. There are many DA techniques available namely Classification techniques involving Naïve Bayes (NB), Decision tree (DT), Neural network (NN), Genetic algorithm (GA), Artificial intelligence (AI) and Clustering algorithms like K-NN, and Support vector machine (SVM). Several studies have been carried out for developing prediction model using individual technique and also by combining two or more techniques. This paper provides a quick and easy review and understanding of available prediction models using data analytics from 2004 to 2016. The comparison shows the accuracy level of each model given by different researchers.