PROJECT OBJECTIVES

Team id	PNT2022TMID46792
Project name	Visualizing and Predicting Heart Diseases with an Interactive Dash Board

Main objective

Exploratory Data Analysis (EDA) is a method to analyze data using advanced techniques to expose hidden structure, enhances the insight into a given dataset, identifies the anomalies and builds parsimonious models to test the underlying assumptions. Exploratory Data Analysis (EDA) is classified into Graphical or non-graphical and Univariate or multivariate Univariate data consider one data column at a time while multivariate method considers more than two variables while analyzing.

The diagnostic methods of diseases are of two types namely, Invasive and Non-invasive . Invasive diagnostic method includes incise procedures in which instruments are used to cut the skin, mucus membrane and connective tissues. In contrast, non-invasive methods are used to diagnose diseases without opening the skin.

Some of the machine learning algorithms based on non-invasive methods are Support Vector Machine(SVM),K- means clustering, K-Nearest Neighbour (KNN), Artificial Neural Network (ANN), Naive Bayes, Logistic Regression and rough set .

Predicting and diagnosing heart disease is the biggest challenge in the medical industry and it is based of actors like physical examination, symptoms and signs of the patient [1-3]. Factors which influence heart diseases are cholesterol level of the body, smoking habit, and obesity, family history of diseases, blood pressure and working environment. Machine learning algorithms play a vital and accurate role in predicting heart disease . The advancement of technologies allows machine language to pair with big data tools to handle unstructured and exponentially growing data . In the paper, K means clustering method is proposed in big data environment and the visualization is made with the tableau dashboard.