Visualizing and Predicting of Heart Diseases with an Interactive Dash Board

Proposed Solution

Our idea is to predict the heart disease and its category using the available datasets for heart disease. It consists age, chest pain type, blood pressure, blood glucose level, heart rate and ECG. All these values will be trained by machine learning techniques and visualized with the help of visualization tool Tableau.

- ❖ K–means Clustering Algorithm

K-means clustering is an unsupervised class of machine learning algorithm. Usually, unsupervised algorithm project the desired output without referring any value. In this algorithm, the data are clustered in such a way that it has highest intraclass similarity and minimal interclass similarity. This algorithm uses the sum of squares distance from the centroid within the cluster. The algorithm divides the data into k clusters with a centroid. This algorithm interactively finds the center that reduces the distance among individual points in a cluster and the cluster center.

- ❖ Tableau

Tableau is one of the business intelligence software used to analyze data and visualize the insights in the form of graph and charts. User can develop and share an interactive dashboard which shows the hidden pattern, trends, density and variation of data. Dashboards are created with the data set after applying k-means algorithm. It provides visual appealing clusters in order to predict the occurrence of heart disease from the given dataset.

Novelty

- ➢ Data collection
- ➢ Data pre-processing and
- ➢ The classification of data.

The data are collected from a standard dataset that contains 303 records. The 15 parameters, such as age, sex, chest pain type (CP), and cholesterol (chol), with some domain values associated with them, considered to predict the probability of heart disease

Feasibility of Idea

- The Datasets which we are using contains the sufficient data without any misinterpretation .
- By using Machine Learning technique, we can predict the outcome of that particular data effectively and more precisely.
- After prediction of diseases the outcome is visualized using Tableau to view the output in an elegant way.

Business Model

Heart disease and stroke, the principal components of cardiovascular disease (CVD), are the first and third leading causes of death in the United States. In 2002, employers representing 88 companies in the United States paid an average of 18,618 dollars per employee for health and productivity-related costs. A sizable portion of these costs are related to CVD.Employers can yield a 3 dollar to 6 dollar return on investment for each dollar invested over a 2 to 5 year period and improve employee cardiovascular health by investing in comprehensive worksite health-promotion programs, and by choosing health plans that provide adequate coverage and support for essential preventive services. The most effective interventions in worksites are those that provide sustained individual follow-up risk factor education and counseling and other interventions within the context of a comprehensive health-promotion program:

- screening, health risk assessments, and referrals
- environmental supports for behavior change (e.g., access to healthy food choices)
- financial and other incentives
- corporate policies that support healthy lifestyles
- electronic medical records.

Comprehensive worksite health-promotion programs, health plans that cover preventive benefits, and effective healthcare systems will have the greatest impact on heart disease and stroke and are likely to reduce employers' health and productivity-related costs.

Scalability of Solution

- The main objective of this research is to develop a heart prediction system, the system can discover and extract hidden knowledge associated with diseases from heart dataset.
- This system aims to exploit on interactive dashboard on medical dataset to assist in the prediction of heart disease.
- Reduce the cost of medical tests.
- To help avoid human biases.
- To find parameter values in prediction like accuracy, elapsed time and energy consumption.