Visualizing and Predicting Heart Diseases with an Interactive Dash Board

TEAM MEMBERS:
1 Bathri D [190801021]
2 Charavanan V [190801025]
3 Ezhil Maran V [190801042]
4 Suneeta R [190801174]

LITERATURE SURVEY:

1. Effective Heart Disease Prediction using Hybrid Machine Learning Algorithms
Link: https://ieeexplore.ieee.org/abstract/document/8740989/

Aims at finding significant features by applying machine learning techniques resulting in improving the accuracy in the prediction of cardiovascular disease. The prediction model is introduced with different combinations of features and several known classification techniques. The prediction models are developed using 13 features and the accuracy is calculated for modelling techniques. The model compares the accuracy, classification error, precision, F-measure, sensitivity and specificity. An enhanced performance level with an accuracy level of 88.7% is achieved through the prediction model for heart disease with the hybrid random forest with a linear model (HRFLM).

2. A novel approach for heart disease prediction using strength scores with significant predictors
Link: https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-021-01527-5

Cardiovascular disease is the leading cause of death in many countries. Physicians often diagnose cardiovascular disease based on current clinical tests and previous experience of diagnosing patients with similar symptoms. Patients who suffer from heart disease require quick diagnosis, early treatment and constant observations. To address their needs, many data mining approaches have been used in the past in diagnosing and predicting heart diseases. Previous research was also focused on identifying the significant contributing features to heart disease prediction, however, less importance was given to identifying the strength of these features. This paper is motivated by the gap in the literature, thus proposes an algorithm that measures the strength of the significant features that contribute to heart disease prediction. The study is aimed at predicting heart disease based on the scores of significant features using Weighted Associative Rule Mining. A set of important feature scores and rules were identified in diagnosing heart disease and cardiologists were consulted to Confirm the validity of these rules. The experiments performed on the UCI open dataset, widely used for heart disease research yielded the highest confidence score of 98% in predicting heart disease.

3. Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction
Link:https://www.academia.edu/download/79534142/5a18f6653b56138cd5196d20e2f39de189e3.pdf

Intends to provide a survey of current techniques of knowledge discovery in databases using data mining techniques that are in use in today's medical research particularly in Heart Disease Prediction. The outcome of predictive data mining technique on the same dataset reveals that Decision Tree outperforms and sometime Bayesian classification is having similar accuracy as that of decision tree but other predictive methods like KNN, Neural Networks, Classification based on clustering are not performing well. The second conclusion is that the accuracy of the Decision Tree and Bayesian Classification further improves after applying a genetic algorithm to reduce the actual data size to get the optimal subset of attributes sufficient for heart disease prediction.

4. Enhanced Heart Disease Analysis and Prediction System [EHDAPS] Using Data Mining
Link: https://www.semanticscholar.org/paper/Improved-Study-of-Heart-Disease-Prediction-System-Dangare-Apte/f4de0213b4a5777ff39d5a94cd574713799ca221

Data mining is the process of analysing data from different perspectives and summarizing it into useful information. Data mining techniques are used for variety of applications. Data mining techniques have been very effective in designing clinical support systems because of their ability to discover hidden patterns and relationships in clinical data. One of the most important applications of such systems is in diagnosis of heart disease. The main objective of Enhanced Heart Disease Analysis and Prediction System (EHDAPS) is predicting the heart disease using historical heart database. To develop this system, medical terms such as sex, blood pressure, and cholesterol like seventeen input attributes are used. In this paper association among various attributes which are the causative factors of heart diseases are analysed. The patient's records are observed before prediction and the factors are grouped as per its severity levels. In this system the level of causative factors is categorized using K-Means clustering technique and it distinguishes the risky and non-risky factors. Frequent risk factors are mined from the clinical heart database using Apriori algorithm. The risk factors are taken for this study to predict the risk level and find the co-ordination among the factors that helps the medical people to predict the disease with minimum tests and treatments.