

Literature Survey

Date	19 September 2022
Team ID	PNT2022TMID53274
Team Members	Sidharth B, Pandiyarasan D, Alagappan S, Kaviraj S
Project Name	Visualizing and Predicting Heart Diseases with an Interactive Dash Board

Abstract:

This document describes various methods of data mining, big data and machine learning models for predicting the heart disease. Data mining and machine learning plays an important role in building an important model for medical system to predict heart disease or cardiovascular disease. Medical experts can help the patients by detecting the cardiovascular disease before occurring. Now-a-days heart disease is one of the most significant causes of fatality. The prediction of heart disease is a critical challenge in the clinical area. But time to time, several techniques are discovered to predict the heart disease in data mining. In this survey paper, many techniques were described for predicting the heart disease.

Key Words: Data mining, prediction, cardiovascular disease, heart disease, machine learning

Introduction:

A study in 2016 found that human beings are collectively generated data more than ten exabytes, or 5×10^{18} bytes from various sources (Lyman and Varian 2003). 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 [15].

“Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques” proposed by Senthilkumar Mohan, Chandrasegar Thirumalai et al.

(2019) was efficient technique using hybrid machine learning methodology. The hybrid approach is combination of random forest and linear method. The dataset and subsets of attributes were collected for prediction. The subset of some attributes were chosen from the pre-processed knowledge(data) set of cardiovascular disease .After prep-processing , the hybrid techniques were applied and diagnosis the cardiovascular disease [4].

M.Satish, et al. (2015) used different Data Mining techniques like Rule based, Decision Tree, Navie Bayes, and Artificial Neural Network. An efficient approach called pruningclassification association rule (PCAR) was used to generate association rules from cardiovascular disease warehouse for prediction of Heart Disease. Heart attack data warehouse was used for pre-processing for mining. All the above discussed data mining technique were described [6].

Mamatha Alex P and Shaicy P Shaji (2019) designed “Prediction and Diagnosis of Heart Disease Patients using Data Mining Technique”. This paper uses techniques of Artificial Neural Network, KNN, Random Forest and Support Vector Machine. Comparing with the above mentioned classification techniques in data mining to predict the higher accuracy for diagnosing the heart disease is Artificial Neural Network[9].

Literature survey:

Predicting the analysis of heart disease symptoms using medicinal data mining methods

Medicinal data mining methods are used to analyze the medical data information resources. Medical data mining content mining and structure methods are used to analyze the medical data contents. The effort to develop knowledge and experience of frequent specialists and clinical selection data of patients collected in databases to facilitate the diagnosis process is considered a valuable option. Diagnosis of heart disease is a significant and tedious task in medicine. The term Heart disease encompasses the various diseases that affect the heart. The exposure of heart disease from various factors or symptom is an issue which is not complimentary from false presumptions often accompanied by unpredictable effects.

Aakash Chauhan et al. (2018) presented “Heart Disease Prediction using Evolutionary Rule Learning”. This study eliminates the manual task that additionally helps in extracting the information (data) directly from the electronic records. To generate strong association rules, we have applied frequent pattern growth association mining on patient’s dataset. This will facilitate (help) in decreasing the amount of services and shown that overwhelming majority of the rules helps within the best prediction of coronary sickness [2].

Ashir Javeed, Shijie Zhou et al. (2017) designed “An Intelligent Learning System based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection”. This paper uses random search algorithm (RSA) for factor selection and random forest model for diagnosing the cardiovascular disease. This model is principally optimized for using grid search algorithmic program.

Two forms of experiments are used for cardiovascular disease prediction. In the first form, only random forest model is developed and within the second experiment the proposed Random Search Algorithm based random forest model is developed. This methodology is efficient and less complex than conventional random forest model. Comparing to conventional random forest it produces 3.3% higher accuracy. The proposed learning system can help the physicians to improve the quality of heart failure detection [3].

Conclusion:

In this paper, a literature survey of review delivers the concept of various techniques has been studied for diagnosing the cardiovascular disease. Use of big data, machine learning along with data mining can provide promising results to bring the most effective accuracy in analysing the prediction model. The main aim of this paper diagnosing the cardiovascular disease or the heart disease and using different methods and many approaches to get prediction.

References:

- [1] V. Manikantan & S.Latha, "Predicting the Analysis of Heart Disease Symptoms Using Medicinal Data Mining Methods", International Journal on Advanced Computer Theory and Engineering, Volume-2, Issue-2, pp.5-10, 2013
- [2] Aakash Chauhan , Aditya Jain , Purushottam Sharma , Vikas Deep, "Heart Disease Prediction using Evolutionary Rule Learning", "International Conference on "Computational Intelligence and Communication Technology" (CICT 2018).
- [3] Ashir Javeed, Shijie Zhou, Liao Yongjian, Iqbal Qasim, Adeeb Noor, Redhwan Nour4, Samad Wali And Abdul Basit , "An Intelligent Learning System based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection" , IEEE Access 2017
- [4] Senthilkumar Mohan, Chandrasegar Thirumalai, and Gautam Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques", IEEE Access 2019.