Visualizing And Predicting Heart Diseases With An Interactive Dashboard

Team ID : PNT2022TMID34621

Team Size : 5

Team Leader : BIBISHEK G S STEEPHEN

Team member : BIBIN S

Team member : AJAY A J

Team member : GINO M

Team member: AJAY E

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 analyzing data that excludes inferences and statistical modeling. 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 analyzing 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.

Heart Diseases

Heart disease is perceived as the deadliest disease in the human life across the world. In particular, in this type of disease the heart is not capable in pushing the required quantity of blood to the remaining organs of the human body in order to accomplish the regular functionalities . Some of the symptoms of heart disease include physical body weakness, improper breathing, swollen feet, etc. The techniques are essential to identify the complicated heart diseases which results in high risk in turn affect the human life . Presently, diagnosis and treatment process are highly challenging due to inadequacy of physicians and diagnostic apparatus that affect the treatment of heart patients . Early diagnosis of heart disease is significant to minimize the heart related issues and to protect it from serious risks . The invasive techniques are implemented to diagnose heart diseases based on medical history, symptom analysis report by experts, and physical laboratory report. Moreover, it causes delay and imprecise diagnosis due to human

intervention. It is time consuming, computationally intensive and expensive at the time of assessment . Heart disease can be predicted based on various symptoms such as age, gender, pulse rate etc. Data analysis in healthcare assists in predicting diseases, improving diagnosis, analyzing symptoms, providing appropriate medicines, improving the quality of care, minimizing cost, extending the life span and reduces the death rate of heart patients**.**

Dataset

The dataset to define the proposed algorithm is the Cleveland heart disease raw dataset with 76 features of 303 patients. During the pre-processing method, some samples are removed to eradicate error due to inconsistency of data. The prediction of heart disease is made with 209 samples with seven independent features like age, chest pain type, blood pressure, blood glucose level, ECG in rest, heart rate and four types of chest pain and the habitual of physical exercise. Age is considered as the main risk factor for heart diseases as coronary fatty streaks develops in the adolescence stage. Male are at higher risk of coronary diseases than females, hence the data set considered here is for only male. Angina is the discomfort caused when the muscles of heart is not supplied with sufficient oxygen rich blood. High blood pressure is one of the major causes of heart disease as it damages arteries. Blood pressure combined with diabetes can increase the risk even more. Heart

rate with high blood pressure increases the risk of heart diseases. Heart beat rate is directly proportional to the risk of coronary disease. The symptom of heart disease includes feeling gripping and tight usually on the chest but spread to shoulders up to the stomach. The types of angina are atypical angina, typical angina, asymptomatic and non-anginal pain.

Prediction of Coronary Heart Disease Using Risk Factor Categories

The objective of this study was to examine the association of Joint National Committee (JNC-V) blood pressure and National Cholesterol Education Program (NCEP) cholesterol categories with coronary heart disease (CHD) risk, to incorporate them into coronary prediction algorithms, and to compare the discrimination properties of this approach with other noncategorical prediction functions.

Methods and Results—This work was designed as a prospective, single-center study in the setting of a community-based cohort. The patients were 2489 men and 2856 women 30 to 74 years old at baseline with 12 years of follow-up. During the 12 years of follow-up, a total of 383 men and 227 women developed CHD, which was significantly associated with categories of blood pressure, total cholesterol, LDL

cholesterol, and HDL cholesterol (all $P<.001$). Sex-specific prediction equations were formulated to predict CHD risk according to age, diabetes, smoking, JNC-V blood pressure categories, and NCEP total cholesterol and LDL cholesterol categories. After adjustment for other factors, ≈28% of CHD events in men and 29% in women were attributable to blood pressure levels that exceeded high normal (≥130/85).

Heart Disease Prediction using Machine Learning

One of the leading causes of deaths around the globe is heart disease. Heart is an organ that is responsible for the supply of blood to each part of the body. Coronary artery disease (CAD) and chronic heart failure (CHF) often lead to heart attack. Traditional medical procedures (angiography) for the diagnosis of heart disease have higher cost as well as serious health concerns. Therefore, researchers have developed various automated diagnostic systems based on machine learning (ML) and data mining techniques. ML-based automated diagnostic systems provide an affordable, efficient, and reliable solutions for heart disease detection. Various ML, data mining methods, and data modalities have been utilized in the past. Many previous review papers have presented systematic reviews based on one type of data modality. This study,

therefore, targets systematic review of automated diagnosis for heart disease prediction based on different types of modalities, i.e., clinical feature-based data modality, images, and ECG. Moreover, this paper critically evaluates the previous methods and presents the limitations in these methods. Finally, the article provides some future research directions in the domain of automated heart disease detection based on machine learning and multiple of data modalities.

. Major Types of Heart Diseases

1. Coronary Artery Disease (CAD)

CAD is a heart disease which commonly occurs as result of the build of fatty deposits (plaque) inside the arteries responsible for supplying blood to the heart muscles. The obstruction in the arteries reduces blood flow to heart muscles which results in the impairment of the heart functions. This phenomenon is known as myocardial ischemia. The partial or complete blockage of arteries results in inevitable damage done to the heart also known as a heart attack. The human heart has four chambers that are divided into upper receiving chamber (right and left atria) and lower pumping chambers (right and left ventricle (LV)). Echocardiography helps in detecting CAD by examining or monitoring the heart for the evolution of CAD and wall

motion abnormalities that begin to arise . CAD can be diagnosed through LV measurement and wall motion scoring. Therefore, monitoring of LV is essential to avoid protracted damages that will affect size, shape, and function of the LV. Echocardiography is an imaging method that captures different cardiac views, structure, and their movement from ultrasound videos. Heart functional and morphological assessment is done to diagnose the cardiac disease through echocardiography . Furthermore, echocardiography is also utilized for quantitative analysis of the LV ejection fraction and cardiac output .

2. Congestive Heart Failure (CHF)

Congestive heart failure also known as chronic heart failure is a condition whereby the heart fails to pump a sufficient amount of blood to the body to meet oxygen demand . CHF is a chronic disease that affects the heart muscles. There are various risk factors behind CHF but the most common risk factors consist of high blood pressure, old age, obesity, and diabetes. Congestive heart failure is more common in men as compared to women. The term heart failure does not refer to the complete cease of the heart, but it actually diminishes the normal functionality of the heart as compared to a healthy person . To test the systolic heart failure, a typical clinical test ejection fraction (EF) is done.For a normal person, the value of ejection fraction is more than 55%, while for diastolic heart failure, the threshold value of ejection

fraction is below 55%. In diastolic heart failure, the heart contracts normally but rigid and inflexible while it is relaxing and being filled with blood. Due to the stiffness of the heart, it is unable to be properly filled with blood to push back into the lungs which causes or leads to heart failure. The ejection fraction in diastolic heart failure is normal or hike.

3. Abnormal Heart Rhythms

Abnormal heart rhythms, also known as arrhythmias, are a condition whereby the heart beats too slow/too fast or irregularly due to a problem in the heart electrical system. The electrical system provides the heart with a clue of when to beat and supply blood to each part of the body. Palpitations, tiredness, losing consciousness, dizziness, and breathlessness are the most common symptoms of an abnormal heart rhythm. The symptoms of heart failure are arduous to notice; therefore, it is also known as the silent killer. Doctors recommend various medical tests for the diagnosis of heart failure, such as echocardiogram, where blood flow through the heart is monitored with the help of ultrasound waves. Electrocardiogram (ECG) is another way to diagnose heart problems related to the heart's rhythm. Holter monitoring is a portable device used to record continuous ECG data of the patient. Cardio computerized tomography (CT) scans provide the facility of an X-ray cross-sectional view of the patient's heart, to detect heart failure. Cardiac

magnetic resonance imaging (MRI) helps to generate an image of the heart and tissues of the heart through the use of powerful magnets and radio waves.

Heart Disease Prediction using Artificial Intelligence

Artificial Intelligence techniques have been widely used in clinical decision support systems for prediction and diagnosis of various diseases with good accuracy. These classifying techniques are very effective in designing clinical support systems due to their ability to get hidden patterns and relationships in medical data provided by medical professionals. One of the most important applications of such systems is in the diagnosis of heart diseases because it is one of the leading causes of deaths all over the world. Almost all systems that predict heart diseases using clinical dataset having parameters and inputs from complex tests conducted in labs. None of the systems predicts heart diseases supporting risk factors like age, case history, diabetes, hypertension, high cholesterol, tobacco smoking, alcohol intake, obesity or physical inactivity, etc. Heart disease patients have many of those visible risk factors in common which may be used

very effectively for diagnosis. A system based on such risk factors would not only help medical professionals but it would give patients a warning about the probable presence of heart disease even before the patient visits a hospital or goes for costly medical check-ups. Hence this paper presents a technique for prediction of heart disease using major risk factors with help of different Classifying Algorithms. This technique involves four major classification algorithms such as K Neighbors, Support Vector, Decision Tree, Random Forest algorithms.

The existing system modules generates comprehensive report by implementing the strong prediction algorithm. In this project the input details are obtained from the patient and the doctor. Then from the doctor inputs, using ai algorithms heart disease is analysed.There are studies showing that reducing these risk factors for heart disease can actually help in preventing heart diseases. There are many studies and researches on the prevention of heart disease risk. Now, the obtained result is compared with the result of existing models with in the same domain and found to be improved. The main aims of the existing system to compare and check the before patient whose having disease outputs and new patient disease and determine future possibilities of the heart disease to a particular patient By

Implementing the above-mentioned model we will get the goal of developing a system with increased rate of accuracy of estimating the new patient getting heart attack percentage. The data of heart disease patients collected from the UCI laboratory is used to discover patterns with K Neighbours Classifier, Support Vector Classifier, decision Tree Classifier, Random Forest Classifier. The results are compared for performance and accuracy with these AI algorithms. The model which is proposed for Heart Disease Prediction System is invented for using different algorithms of AI and approach. But by using all the existing systems theaccuracy is very less.

The data for 50 people was collected from surveys done by the American Heart Association. Most of the heart disease patients had many similarities in the risk factors. The TABLE I below shows the identified important risk factors and the corresponding values and their encoded values in brackets, which were used as input to the system.Data analysis has been carried out in order to transform data into useful form, for this the values were encoded mostly between a range [-1, 1]. Data analysis also removed the inconsistency and anomalies in the data. This was needed. Data analysis was needed for correct data preprocessing. The removal of missing and incorrect inputs will

help the neural network to generalize well.The proposed application is developed using python and is capable of identifying if a patient has heart disease or not.