

# **1. Visualization and Prediction of Heart Diseases Using Data Science Framework**

## **Author :**

Vaibhav Gupta, Vaibhav Aggarwal, Shagun Gupta, Neeti Sharma, Kiran Sharma, Neetu Sharma.

## **Abstract :**

Heart is one the most vital organ in the human body. When we talk about heart diseases, we can have multiple conditions where heart is not working the way it should be like blockage in blood vessels. According to many researches that have been conducted through a period of time have found out that heart failure and heart disease has been the cruel cause of death in human beings. What aggravates this situation is that most of these diseases are being diagnosed at later stages at which it is very difficult to control. But if somehow, we can diagnose these diseases at its early stage, then we can surely cure the disease. The main aim of this paper is to use various classification algorithms of data science framework to somehow detect the chances of having a heart disease. Also, the main aim of this research paper is to find out the most efficient classification algorithm that can help us to detect heart diseases at early stage. This algorithm can be used on heart records of the patient or by using it on classification reports. This research was conducted and tested upon various algorithms to test its accuracy like Logistic Regression, Random Forest, Vector Support and XG-Boost. After applying these algorithms of prediction model has been developed.

## **2.Heart Disease Prediction using Exploratory Data Analysis**

**Author :**

R.Indrakumari, T.Poongodi, Soumya Ranjan Jena

**Abstract :**

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 analysing data that excludes inferences and statistical modelling. 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 analysing 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. To predict the heart disease, K-means clustering algorithm is used along with data analytics and visualization tool. The paper discusses the pre-processing methods, classifier performances and

evaluation metrics. In the result section, the visualized data shows that the prediction is accurate.

### **3.Heart Disease Prediction Using Tableau**

#### **Author :**

R. Indrakumari, Priyanka Shukla, Akanksha Sehgal

#### **Abstract :**

Heart disease is defined as any ailment that harms the heart. The terms “heart disease” and “cardiovascular disease” are frequently used interchangeably. Cardiovascular disease often refers to problems affecting heart and leading to blockages, such as atherosclerosis, whereas heart disease is an umbrella term that encompasses various cardiac disorders. Heart disease can impact not just the arteries but also the heart muscle, valves, rhythm, and other vital parts of a healthy heart. This chapter exemplifies the application of data science, notably Tableau, in the healthcare business using the Cleveland Heart Disease dataset.

# **4. Predicting the Analysis of Heart Disease Symptoms Using Medicinal Data Mining Methods**

**Author :**

V. Manikantan, S. Latha

**Abstract :**

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. Association rule mining procedures are used to extract item set relations. Item set regularities are used in the rule mining process. The data classification is based on MAFIA algorithms which result in accuracy, the data is evaluated using entropy based cross validations and partition techniques and the results are compared. Here using the C4.5 algorithm as the training algorithm to show rank of heart attack with the decision tree. Finally, the heart disease database is clustered using the K-means clustering algorithm, which will remove the data applicable to heart attack from the database. The results showed

that the medicinal prescription and designed prediction system is capable of prophesying the heart attack successfully.

## **5.Heart Disease Prediction Using Data Mining preprocessing and Hierarchical Clustering**

**Author :**

Dr.A.V.Senthil Kumar

**Abstract :**

The diagnosis of diseases is a crucial and difficult job in medicine. The recognition of heartdisease from diverse features or signs is a major issue which is not free from false presumptions often accompanied by unpredictable effects. The healthcare industry gathers enormous amounts of heart disease data that unfortunately, are not mined to determine concealed information for effective diagnosing. Due to this rapid growth is the main motivation for researchers to mine useful information from these medical databases. As the volume of stored data increases, data mining techniques play an important role in finding patterns and extracting knowledge to provide better patient care and effective diagnostic capabilities. Heart disease prediction suffers from the problem of missing data, statistical tests will lose power, results may

be based, or analysis may not be feasible at all. There are several ways to handle the problem, for example through imputation. To overcome this problem initially, the data set containing 13 medical attributes were obtained from the Cleveland heart disease database missing attributes data is replaced with the help of imputation method. With imputation, missing values are replaced with estimated values according to an imputation method or model. In this paper, preprocessed dataset from EM is given as input to clustering method for heart disease prediction. In this paper, an efficient approach non negative matrix factorization with hierarchical clustering methods (NMF-HC) is proposed for the intelligent heart disease prediction. The dataset is clustered with the aid of NMF-HC clustering algorithm. The NMF-HC is trained using the preprocessed data sets. The proposed NMF-HC works as promising tool for prediction of heart disease.

## **6.Heart Disease Prediction Using Classification Techniques with Feature Selection Method**

**Author :**

Uma K, M. Hanumathappa

**Abstract :**

Heart disease is now turned out most deadly disease throughout the world. Due to misdiagnosis of heart disease more people losing their lives. Hence, there is a need of automate the system for correct diagnose the heart disease based on the historical data. To aid early and correct

diagnose of heart disease, many data mining techniques are used to predicting the disease. The high volume of medical data offered data mining techniques to discover the hidden pattern. Classification technique is one among the data mining techniques predict the heart disease. This paper presents, classification techniques applied for prediction of heart disease in two scenario such as dataset with all 13 attributes and 6 attributes selected by attribute selection method. For achieving the results, the selected classification techniques are Support vector machine, Neural Network (Multilayer perception), Bagging, Classification via regression and Simple logistic. And correlates the accuracy and time taken to build the prediction model for all used classification techniques in two different scenario.

## **7.Prediction of Heart Disease using Machine Learning Algorithms:A Survey**

**Author :**

Himanshu Sharma, M.A. Rizvi

**Abstract :**

According to recent survey by WHO organisation 17.5 million people dead each year. It will increase to 75 million in the year 2030[1].Medical professionals working in the field of heart disease have their own limitation, they can predict chance of heart attack up to 67% accuracy[2], with the current epidemic scenario doctors need a support system for more accurate prediction of heart disease. Machine learning algorithm and deep learning opens new door opportunities for precise predication of heart attack. Paper provideslot information about state of art methods

in Machine learning and deep learning. An analytical comparison has been provided to help new researches' working in this field.

## **8.Epidemiology and risk profile of heart failure**

**Author :**

A.L. Bui, T.B. Horwich, G.C. Fonarow

### **Abstract :**

Heart failure (HF) is a major public health issue, with a prevalence of over 5.8 million in the USA, and over 23 million worldwide, and rising. The lifetime risk of developing HF is one in five. Although promising evidence shows that the age-adjusted incidence of HF may have plateaued, HF still carries substantial morbidity and mortality, with 5-year mortality that rival those of many cancers. HF represents a considerable burden to the health-care system, responsible for costs of more than \$39 billion annually in the USA alone, and high rates of hospitalizations, readmissions, and outpatient visits. HF is not a single entity, but a clinical syndrome that may have different characteristics depending on age, sex, race or ethnicity, left ventricular ejection fraction (LVEF) status, and HF etiology. Furthermore, pathophysiological differences are observed among patients diagnosed with HF and reduced LVEF compared with HF and preserved LVEF, which are beginning to be better appreciated in epidemiological studies. A number of risk factors, such as ischemic heart disease, hypertension, smoking, obesity, and diabetes, among others, have been identified that both predict the



incidence of HF as well as its severity. In this Review, we discuss key features of the epidemiology and risk profile of HF.

## **9. Predictive data mining for medical diagnosis:**

### **An overview of heart disease prediction**

#### **Author :**

Soni J., Ansari U., Sharma D., Soni S.

#### **Abstract :**

The successful application of data mining in highly visible fields like e-business, marketing and retail has led to its application in other industries and sectors. Among these sectors just discovering is healthcare. The healthcare environment is still „information rich“ but „knowledge poor“. There is a wealth of data available within the healthcare systems. However, there is a lack of effective analysis tools to discover hidden relationships and trends in data. This research paper 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. Number of experiment has been conducted to compare the performance of predictive data mining technique on the same dataset and the outcome reveals that Decision Tree outperforms and some time Bayesian classification is having similar accuracy as 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 genetic algorithm to reduce the actual data size

to get the optimal subset of attribute sufficient for heart disease prediction.

## **10.Prediction of heart disease using classification algorithms in Proceedings of the world congress on engineering and computer science**

**Author :**

Masethe, H. D., & Masethe, M. A.

**Abstract :**

The heart disease accounts to be the leading cause of death worldwide. It is difficult for medical practitioners to predict the heart attack as it is a complex task that requires experience and knowledge. The health sector today contains hidden information that can be important in making decisions. Data mining algorithms such as J48, Naive Bayes, REPTREE, CART, and Bayes Net are applied in this research for predicting heart attacks. The research result shows prediction accuracy of 99%. Data mining enable the health sector to predict patterns in the dataset.