Sudhanshu Memane, 2021, 9:6
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

Heart Disease Prediction Using Naive Bayes Classifier

Sudhanshu Memane , Akash Patel , Anjal Patel , Omkar Dive , Prof. Uzma Sheikh
Department of Computer Engineering,
Trinity College of Engineering and Research,
Pune, Maharashtra, India

Abstract- In healthcare world for providing appropriate results and making effective decisions on data, some advanced data mining techniques are used. In this study, a Heart Disease Prediction System (HDPS) is developed using Naive Bayes algorithm for predicting the risk level of heart disease. The system uses 14 medical parameters such as age, sex, blood pressure, cholesterol, and obesity for prediction. The HDPS predicts the likelihood of patients getting heart disease. E.g., Relationships between medical factors related to heart disease and patterns, to be established. The obtained results have illustrated that the designed diagnostic system can effectively predict the risk level of heart diseases.

Keywords- heart disease, HDPS predicts. Etc.

I. INTRODUCTION

The goal is to predict the health of a patient from collective data, so as to be able to detect configurations at risk for the patient, and therefore, in cases requiring emergency medical assistance. Heart Disease is even highlighted as a silent killer which leads to the death of the person without obvious symptoms. The early diagnosis of heart disease plays a vital role in making decisions on lifestyle changes in high-risk patients and in turn reduce the complications. This project aims to predict future heart disease by analyzing data of patients which classifies whether they have heart disease or not using machine-learning algorithms.

The main feature will be the machine learning, in which we will be using algorithm such as Naïve Bayes Algorithm, which will predict accurate disease and also, will find which algorithm gives a faster and efficient result by comparatively comparing. performed using training data set consists of 3000 instances with 13 different attributes which has mentioned earlier. The data set is divided into two parts that is 70% of the data are used for training and 30% used for testing.

II.LITERATURE SURVEY

1. "Pola Raju, Durga Prasad, & Tech Scholar" (2017) This paper proposed Prediction of Heart Disease using Multiple Regression Model and it proves that Multiple Linear Regression is appropriate for predicting heart disease chance. The work is

2. "Deepika & Seema" (2017)
This study focuses on techniques that can predict chronic disease by mining the data containing in historical health records using Naïve Bayes, Decision tree, Support Vector Machine (SVM) and Artificial Neural Network (ANN). A comparative study is performed on classifiers to measure the better performance on an accurate rate. From this experiment, SVM gives highest accuracy rate, whereas for diabetes Naïve Bayes gives the highest accuracy.

3. "Beyene & Kamath" (2018)
This study recommended different algorithms like Naive Bayes, Classification Tree, KNN, Logistic Regression, SVM and ANN. The Logistic Regression gives better accuracy compared to other algorithms suggested Heart Disease Prediction System using Data Mining Techniques. WEKA software used for

International Journal of Science, Engineering and Technology

An Open Access Journal

automatic diagnosis of disease and to give qualities of services in healthcare centres. The paper used various algorithms like SVM, Naïve Bayes, Association rule, KNN, ANN, and Decision Tree. The paper recommended SVM is effective and provides more accuracy as compared with other data mining algorithms.

4. "A & Naik" (2016)

This study recommended to develop the prediction system which will diagnosis the heart disease from patient's medical data set. 13 risk factors of input attributes have considered to build the system. After analysis of the data from the dataset, He used k-means and naïve Bayes to predict heart disease. To extract knowledge from database, data mining techniques such as clustering, classification methods can be used. 13 attributes with total of 300 records were used from the Cleveland Heart Database. This model is to predict whether the patient have heart disease or not based on the values of 13 attributes

5. "Sultana, Haider, & Uddin" (2017)

This study proposed an analysis of cardiovascular disease. This paper proposed data min in techniques to predict the disease. It is intended to provide the survey of current techniques to extrac information from dataset and it will useful for health care practitioners. The performance can be obtained based on the time taken to build the decision tree for the system. The primary objective is to predict the disease with a smaller number of attributes.

6. "Sai & Reddy" (2017)

This study proposed heart disease prediction using ANN algorithm in data mining. Due to increasing expenses of heart disease diagnosis disease, there was a need to develop new system which can predict heart disease. Prediction model is used to predict the condition of the patient after evaluation on the basis of various parameters like heart beat rate, blood pressure, cholesterol etc.

III.EXISTING SYSTEM

In The Existing System The Data Set Is Typically Small, For Patients And Diseases With Specific Conditions. These Systems Are Mostly Designed For The More Casual Symptoms Such As Mild Pain, Etc. The Pre-Selected Characteristics May Sometimes Not Satisfy The Changes In The Disease And Its Influencing Factors Which Could Lead To Inaccuracy In Results. As We Live In Continuously Evolving World, The Symptoms Of Diseases Also Evolve Over A Course Of Time. Also, Most Of The Current Systems Make The Users Wait For Long Periods By Making Them Answer Lengthy Question naires.

V.PROPOSED SYSTEM

In this system we are implementing effective heart attack prediction system using naïve bayes algorithm. we can give the input as in csv file or manual entry to the system. after taking input the algorithms apply on that input that is naïve bayes. after accessing data set the operation is performed and effective heart attack level is produced. the proposed system will add some more parameters significant to heart attack with their weight, age and the priority levels are by consulting expertise doctors and the medical experts. the heart attack prediction system designed to help the identify different risk levels of heart attack like normal, low or high and also giving the prescription details with

IV.RELATED TO THE PREDICTED RESULT

Naïve Bayes Classifier

Naïve Bayes classifier is based on Bayes theorem. This classifier uses conditional independence in which attribute value is independent of the values of other attributes. The Bayes theorem is as follows: Let X= {x1, x2, ..., Xn} be a set of n attributes. In Bayesian, X is considered as evidence and H be some hypothesis means, the data of X belongs to specific class C. We have to determine P (H|X), the probability that the hypothesis H holds given evidence i.e., data sample X. According to Bayes theorem the P (H|X) is expressed as: P(H|X) = P(X|H) P(H) / P(X)

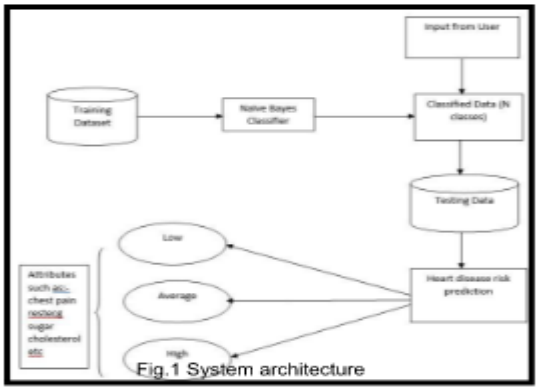

Fig.1 System architecture

Fig.1, using Bayesian classifiers, the system will discover the concealed knowledge associated with diseases from historical records of the patients having heart disease. Bayesian classifiers predict the class membership probabilities, in a way that the probability of a given sample belongs to a particular class statistically. Bayesian classifier is based on Bayes' theorem. We can use Bayes theorem to determine the probability that a proposed diagnosis is correct, given the observation. A simple probabilistic, the naive Bayes classifier is used for classification based on which is based on Bayes' theorem.

According to naïve Bayesian classifier the occurrence or an occurrence of a particular feature of a class is considered as independent in the presence or absence of any other feature. When the dimension of the inputs is high and more efficient result is expected, the chief Naïve Bayes Classifier technique is applicable. The Naïve Bayes model identifies the physical characteristics and features of patients suffering from heart disease. For each input, it gives the possibility of attribute of the expectable state. Naïve Bayes is a statistical classifier which assumes no dependency between attributes. This classifier algorithm uses conditional independence, means it assumes that an attribute value of a given class is independent of the values of other attributes. The advantage of using Naïve Bayes is that one can work with the Naïve Bayes model without. using any Bayesian methods.

V.CONCLUSION

This paper gives research of multiple researches done in this field. Our Proposed System aims at bridging gap between Doctors and Patients which will help both classes of users in achieving their goals. The proposed system scalable, reliable and an expandable system. The proposed working model can also help in reducing treatment costs by providing Initial diagnostics in time.

The model can also serve the purpose of training tool for medical students and will be a soft diagnostic tool available for physician and cardiologist. As we have developed a generalized system, in future we can use this system for the analysis of different data sets. The performance of the health's diagnosis can be improved significantly by handling numerous class labels in the prediction process, and it can be another positive direction of research.

REFERENCES

[1] Meenakshi, K., Maragatham, G., Agarwal, N., &Amp; Ghosh, I. (2018). A Data Mining Technique For Analyzing And Predicting The Success Of Movie.

[2]. Desai, S. D., Giraddi, S., Narayankar, P. Pudakalakatti, N. R., "Back-Propagation Neural Network Versus Logistic Regression In Heart Disease Classification" (2017)

[3]. A. L. Bui, T. B. Horwich, And G. C. Funaro, Epidemiology And Risk Prole Of Heart Failure (2011)

[4]. M. Durairaj And N. Ramasamy, "A Comparison Of The Perceptive Approaches For Preprocessing The Data Set For Predicting Fertility Success Rate" (2016)

[5]. Nandhini, S., Debnath, M., Sharma, A. And Pushkar 2018. Heart Disease Prediction Using Machine Learning.

[6]. Garate Escamila, A. K., Hassani, A. H. And Andres, E. 2020. Classification Models For Heart Disease Prediction Using Feature Selection And PCA. Informatics In Medicine Unlocked 19:1-11.

[7]. Kavitha, S., Baskaran, K. R. And Sathyavathi, S. 2018. Heart Disease With Risk Prediction Using Machine Learning Algorithms. International Journal Of Recent Technology And Engineering 7(48):314-317.

[8]. Pola Raju, K., Durga Prasad, D., &Amp; Tech Scholar, M. (2017). Prediction Of Heart Disease Using Multiple Linear Regression Model.

Author's Details

1. Student, Computer Department, Trinity College Of Engineering And Research, Maharashtra, India, Sudhanshumemane@Gmail.Com
2. Student, Computer Department, Trinity College Of Engineering And Research, Maharashtra, India, Asp4020@Gmail.Com
3. Student, Computer Department, Trinity College Of Engineering And Research, Maharashtra, India, Anjal.Patel9@Gmail.Com
4. Student, Computer Department, Trinity College Of Engineering And Research, Maharashtra, India, Omkardive77@Gmail.Com
5. Professor, Computer Department, Trinity College Of Engineering And Research, Maharashtra, India Uzmasheikh.Tcoer@Kjei.Edu.In