A Project/Dissertation Report

on

Heart Disease Detection Using Classification Algorithms

Submitted in partial fulfillment of the

requirement for the award of the degree of

B.Tech (Computer Science & Engineering)

GALGOTIAS UNIVERSITY

(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

Under The Supervision of
Name of Supervisor : Ms. Kiran Singh

Submitted By

NAME	ADMISSION NUMBER
DHRUV AGNIHOTRI	**20SCSE1010027**
ASHISH TiWARI	**20SCSE1010517**
AMAN MISHRA	**20SCSE1010371**

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA
INDIA
MAY, 2022

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA

CANDIDATE’S DECLARATION

We hereby certify that the work which is being presented in the thesis/project/dissertation, entitled **“HEART DISEASE DETECTION USING CLASSIFICATION ALGORITHM”** in partial fulfillment of the requirements for the award of the **B.Tech**-submitted in the School of Computing Science and Engineering of Galgotias University, Greater Noida, is an original work carried out during the period of month, Year to Month and Year, under the supervision of Name… Designation, Department of Computer Science and Engineering/Computer Application and Information and Science, of School of Computing Science and Engineering , Galgotias University, Greater Noida

The matter presented in the thesis/project/dissertation has not been submitted by me/us for the award of any other degree of this or any other places.

	DHRUV AGNIHOTRI	20SCSE1010027
	ASHISH TIWARI	20SCSE1010517
`	AMAN MISHRA	20SCSE1010371

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Supervisor Name

Designation

CERTIFICATE

The Final Thesis/Project/ Dissertation Viva-Voce examination of DHRUV AGNIHOTRI 20SCSE1010027 , ASHISH TIWARI 20SCSE1010517 , AMAN MISHRA 20SCSE1010371 has been held on __________________ and his/her work is recommended for the award of B.Tech in Computer Science and Engineering .

Signature of Examiner(s) **Signature of Supervisor(s)**

Signature of Project Coordinator **Signature of Dean**

Date: November, 2013

Place: Greater Noida

Abstract

The correct prediction of heart disease can prevent life threats, and incorrect prediction can prove to be fatal at the same time. Every day the cases of coronary heart illnesses are growing at a rapid rate and it's very important and regarding to predict one of these diseases beforehand. This diagnosis is a difficult undertaking i.e. it have to be accomplished exactly and successfully. The studies report in particular specializes in which affected person is much more likely to have a coronary heart disorder primarily based on numerous medical attributes. The main objective of the project is to get a better accuracy to detect the heart-disease using algorithms in which the target output counts that a person having heart disease or not.

In this report different machine learning algorithms are applied to compare the results and analysis of the UCI Machine Learning Heart Disease dataset. The dataset consists of 14 main attributes used for performing the analysis. Various promising results are achieved and are validated using accuracy and confusion matrix. The dataset consists of some irrelevant features which are handled and data are also normalized for getting better results. And this study can be combined with some multimedia technology like mobile devices is the additional concern to be discussed.

Using different types of parameters in the dataset we can predict the cardiac-disease. The study of existing systems has resulted in the discovery of the best data classification algorithm among various other algorithms. The algorithms used for comparison were **Logistic Regression, Decision Tree, MLP, KNN and Random Forest** in which **Random Forest** proved to be comparatively accurate.

This report depicts the use of a Machine Learning Model using classification algorithm. This is beneficial in field of medical science as it could reduce the load of medical practitioner and treat the patient more efficiently. Also, cardiovascular disease can be discovered at an early stage which reduce the heart attacks among patients. Treatment can be done according to the severity of the detected disease.

The term Heart disease covers all diseases and disorders related to heart and blood vessels. Nowadays there is an increase in heart disease. The diagnosis and treatment for this sums up to a huge amount. In solution to this problem many research have been made to reduce the expense and to get quick result. This model would be able to develop a more accurate system for detection of heart disease using classification algorithms.

Table of Contents

CHAPTER-1
Introduction

Heart disease describes a range of conditions that affect your heart. Today, cardiovascular diseases are the leading cause of death worldwide with 17.9 million deaths annually, as per the World Health Organization reports. Various unhealthy activities are the reason for the increase in the risk of heart disease like high cholesterol, obesity, increase in triglycerides levels, hypertension, etc. There are certain signs which the American Heart Association lists like the persons having sleep issues, a certain increase and decrease in heart rate (irregular heartbeat), swollen legs, and in some cases weight gain occurring quite fast; it can be 1-2 kg daily .All these symptoms resemble different diseases also like it occurs in the aging persons, so it becomes a difficult task to get a correct diagnosis, which results in fatality in near future.

But as time is passing, a lot of research data and patients records of hospitals are available. There are many open sources for accessing the patient's records and researches can be conducted so that various computer technologies could be used for doing the correct diagnosis of the patients and detect this disease to stop it from becoming fatal. Nowadays it is well known that machine learning and artificial intelligence are playing a huge role in the medical industry. We can use different machine learning models to diagnose the disease and classify or predict the results. A complete genomic data analysis can easily be done using machine learning models. Models can be trained for knowledge pandemic predictions and also medical records can be transformed and analysed more deeply for better predictions.

All the diseases or conditions that affect the heart are generally termed as heart diseases. Now days due to advancement in technology most of the hospitals store their patient's information and medical issues in an information system. Huge amounts of data which taken in the form of numbers, text, and images are stored and managed by these systems. But all this data is most of the time never used to support clinical decisions. This gives rise to an important question that how can this data be turned into useful content so that it can help the medical practitioners to take important decisions in the medical field or support these decisions in an effective way. So this is the main question which needs to be cratered so that all this important data can be used instead of just consuming memory.

Python is most powerful programming language having numerous libraries which is used in this project with machine learning model. Machine learning is a subset model of artificial intelligence network in which uses complex algorithms and deep learning neural networks. Cardio vascular disease is a widespread disease in all over a region. This type of disease may cause due to smoking, high blood pressure, diabetes, overweight, hyper tension, cholesterol etc. that has to be accumulated because of the fatty foods or unlimited intake of foods or non-moving to anywhere.

This disease may occur by various heart problems such as coronary-artery disease, cardio-vascular, stroke, heart failure and much more. Chest pain (cp), resting blood pressure, cholesterol, resting electrocardiographic results, fasting blood sugar(fbs), maximum heart achieved, exercise induced angina, ST depression induced by exercise relative to rest, slope of the peak exercise ST segment, number of major vessels colored by fluoroscopy etc., are the major reasons for causing heart problems but we have a attributes of individual person like height, weight, systolic blood pressure, diastolic blood pressure, cholesterol, glucose, smoke, alcohol, active (physically active person).

CHAPTER-2

Literature Survey

[1]In this paper aurthor used different Data Mining techniques like Rule based, Decision Tree, Navie Bayes, and Artifical Neural Network. An efficient approach called pruning classification 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].
In this paper authors used two algorithms: hill climbing and decision tree. Before applying the classification algorithms, the data is pre-processed. The data set used is Cleveland data set. The Knowledge Extraction is done based on Evolutionary Learning (KEEL), an open-source data mining tool which is used to fill the missing values in the data set. Hill climbing algorithm is then used to find the best subset of rules. The parameters and their values used are - Confidence: minimum confidence value is 0.25, MinItemsets: the minimum number of item-sets per leaf is two, Threshold: a value of 10 is used to find the best subset of rules for the hill-climbing algorithm by the authors. A decision tree is constructed in the top-down approach for each level and a node is selected by a test for the actual node chosen using a hill-climbing algorithm. The decision tree knows how to generate the basic rules. First generated rules are referred to as original rules and from these rules Pruned Rules are generated. From these rules without duplicates and classified rules are obtained. And finally a small number of rules called Class wise Rule distribution is generated. The accuracy of the system is about 86.7%.

[2]In this paper author has implemented hybrid machine learning for heart disease prediction. The data set used is Cleveland data set. The first step is data pre-processing step. In this the tuples are removed from the data set which have missing the values. Attributes age and sex from data set are also not used as the authors think that it's personal information and has no impact on predication. The remaining 11 attributes are considered important as they contain vital clinical records. They have proposed own Hybrid Random Forest Linear Method (HRFLM) which is combination of Random Forest (RF) and Linear method (LM). In HRFLM algorithm the authors have used four algorithms. First algorithm deals with partitioning the input dataset. It is based on decision tree which is executed for each sample of the dataset. After identifying the feature space, the dataset is split into the leaf nodes. Output of first algorithm is Partition of data set. After that in second algorithm they apply rules to the data set and output here is the classification of data with those

rules. In third algorithm features are extracted using Less Error Classifier. This algorithm deals with finding the minimum and maximum error rate from the classifier. Output of this algorithm is the features with classified attributes. In forth algorithm they apply Classifier which is hybrid method based on the error rate on the Extracted Features. Finally they have compared the results obtained after applying HRFLM with other classification algorithms such as decision tree and support vector machine. In result as RF and LM are giving better results than other, both the algorithms are put together and new unique algorithm HRFLM is created. The accuracy of HRFLM initially increased with number splits and then has become constant at a particular level. The accuracy obtained is 88.7% which higher than the SVM and decision tree. The authors suggest further improvement in accuracy by using combination of various machine learning algorithms and also by concentrating on developing novel feature selection techniques which would help in extracting significant features.

[3]In this paper authors deal with various supervised machine learning algorithms such as Random Forest, Support Vector Machine, Logistic Regression, Linear Regression, Decision Tree with 3 fold, 5 fold and 10 fold cross-validation techniques. They have used Cleveland data set having 303 tuples, with some tuples having missing attributes. In the preprocessing of data they just removed the missing value tuple from the data set which are six in number and then from the remaining 297 tuples, they divided the data as training 70% and testing 30%. First algorithm applied is Linear Regression. In this, they have defined the dependency of one attribute over others which can be linearly separated from each other. Basically the classification takes place with the help of the group of attributes used for binary classification. They have obtained best results in 10 fold which is 83.82%. Logistic regression classification is done using a sigmoid function. This algorithm applied for heart disease prediction shows maximum accuracy with 3 and 5 fold cross-validation and it is 83.83%. Support Vector Machine is the classification algorithm in supervised machine learning. In this the classification is done by hyperplane. The maximum accuracy achieved by SVM in 3 fold cross-validation is 83.17%. For Decision Tree in this paper, the authors have used different number splits and different number of leaf nodes to find the maximum accuracy. With 37 number splits and 6 leaf nodes maximum accuracy is achieved which is 79.12%. When used with cross-validation, accuracy achieved by the decision tree 79.54% with 5 fold. Random forest algorithm used on nonlinear data set gives better results as compared to the decision tree. Random forest is the group of decision tree created by the different root nodes. From this group of decision tree, voting can be done first and then classification can be done from the one getting maximum votes. Authors have used different number splits, different number of tree per observation and different

number of folds for cross-validation. For random forest, 85.81% accuracy is achieved by 20 Number of splits, 75 Number of trees and 10 number of folds.

[4]In this paper authors deal with machine learning algorithms such as decision tree and Naive Bayes algorithm for prediction of heart disease. In first algorithm the decision tree is built using certain conditions which gives True or False decisions. Other algorithms like SVM, KNN are results based on vertical or horizontal split conditions depends on dependent variables. But decision tree for a tree like structure having root node, leaves and branches base on the decision made in each of tree Decision tree also help in the understating the importance of the attributes in the dataset. They have also used Cleveland data set. Dataset splits in 70% training and 30% testing. This algorithm gives a 91% accuracy. The second algorithm is Naive Bayes. It is used for classification. It can handle complicated, nonlinear, dependent data and hence is found suitable for heart disease dataset as this dataset is also complicated, dependent and nonlinear in nature. This algorithm gives an 87% accuracy.

CHAPTER-3
Functionality/Working of Project

(i.) Data collection

Overall process of predicting heart disease carries following procedure:

- We have collected data from dataset provider –Kaggle.com. The dataset which is published as in the title of Heart Disease dataset .The dataset collected consists of **1024** records of patients, data carries **14** features.
- Dataset is the information or a tool essential to do any kind of research or a project.

(ii.) Data Analysis

- This , process involves studying the attributes and features of the dataset and analyzing their correlation with each other.
- Further, in this process we find out the different measures of statistics and also describe the dataset and analyze it on the basis of the different first few and last few tuples of the whole dataset.
- Now the attributes which are used in this project are described as follows and for what they are used or resemble:

  (i)**Age**—age of patient in years.

  (ii)**sex**—(1 = male; 0 = female).

  (iii)**Cp**—chest pain type.

  (iv)**Trestbps**—resting blood pressure (in mm Hg on admission to the hospital). The normal range is 120/80 (if you have a normal blood pressure reading, it is fine, but if it is a little higher than it should be, you should try to lower it. Make healthy changes to your lifestyle).

  (v)**Chol**—serum cholesterol shows the amount of triglycerides present. Triglycerides are another lipid that can be measured in the blood. It should be less than 170 mg/dL (may differ in different Labs).

  (vi)**Fbs**—fasting blood sugar larger than 120 mg/dl (1 true). Less than 100 mg/dL (5.6 mmol/L) is normal, and 100 to 125 mg/dL (5.6 to 6.9 mmol/L) is considered prediabetes.

  (vii)**Restecg**—resting electrocardiographic results.

(viii)**Thalach**—maximum heart rate achieved. The maximum heart rate is 220 minus your age.

(ix)**Exang**—exercise-induced angina (1 yes). Angina is a type of chest pain caused by reduced blood flow to the heart. Angina is a symptom of coronary artery disease.

(x)**Oldpeak**—ST depression induced by exercise relative to rest.

(xi)**Slope**—the slope of the peak exercise ST segment.

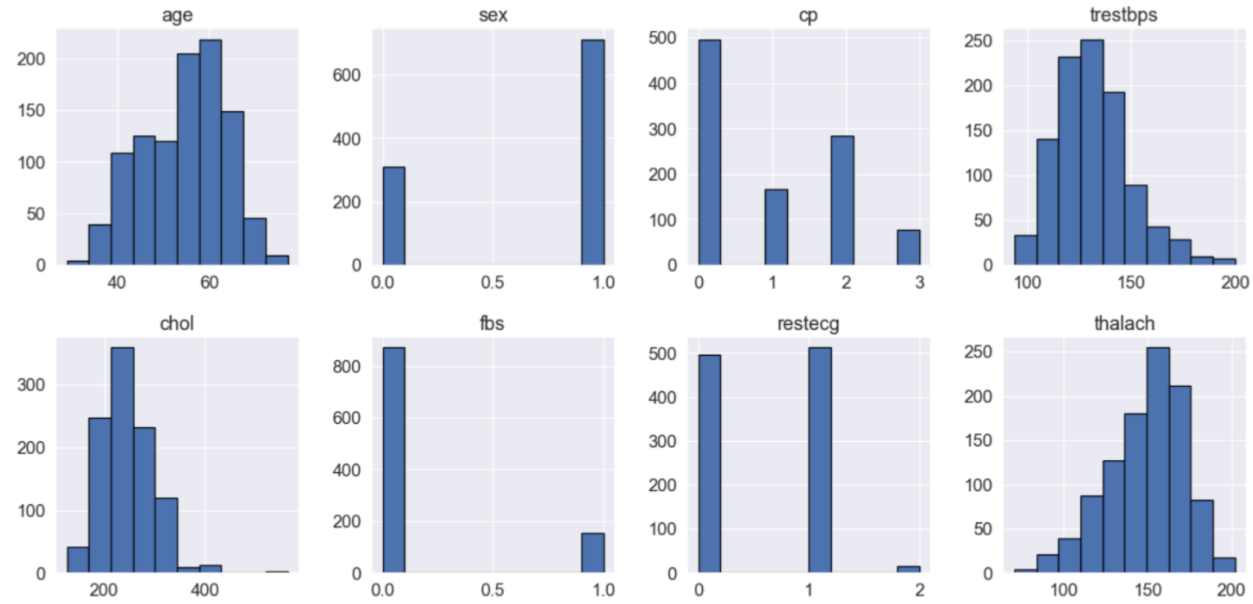
(xii)**Ca**—number of major vessels (0–3) colored by fluoroscopy.

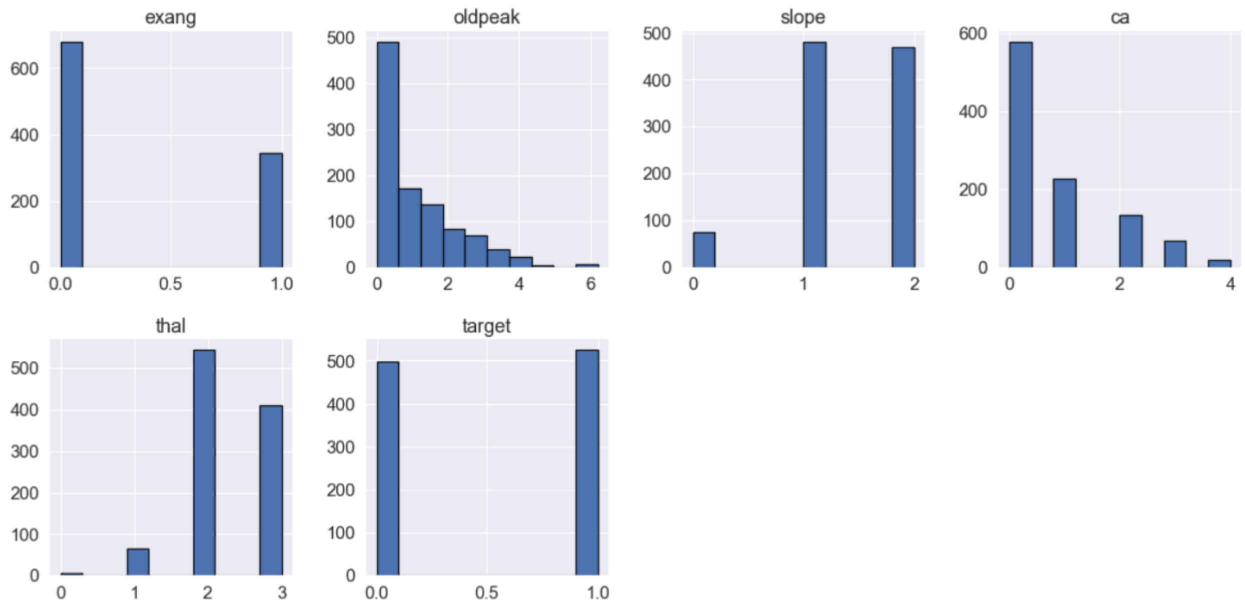
(xiii)**Thal**—no explanation provided, but probably thalassemia (3 normal; 6 fixed defects; 7 reversible defects).

(xiv)**Target (T)**—no disease = 0 and disease = 1, (angiographic disease status).

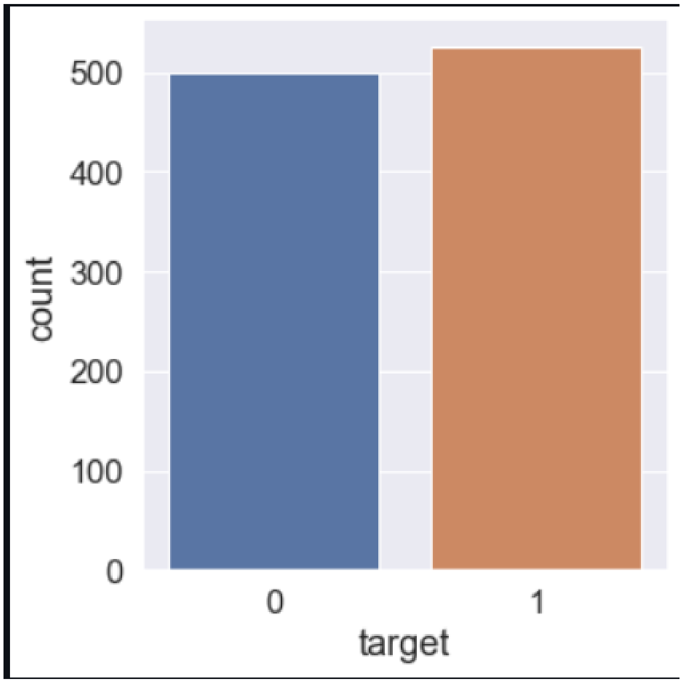
(iii.) **Data Visualization**

- Plotting different attributes.

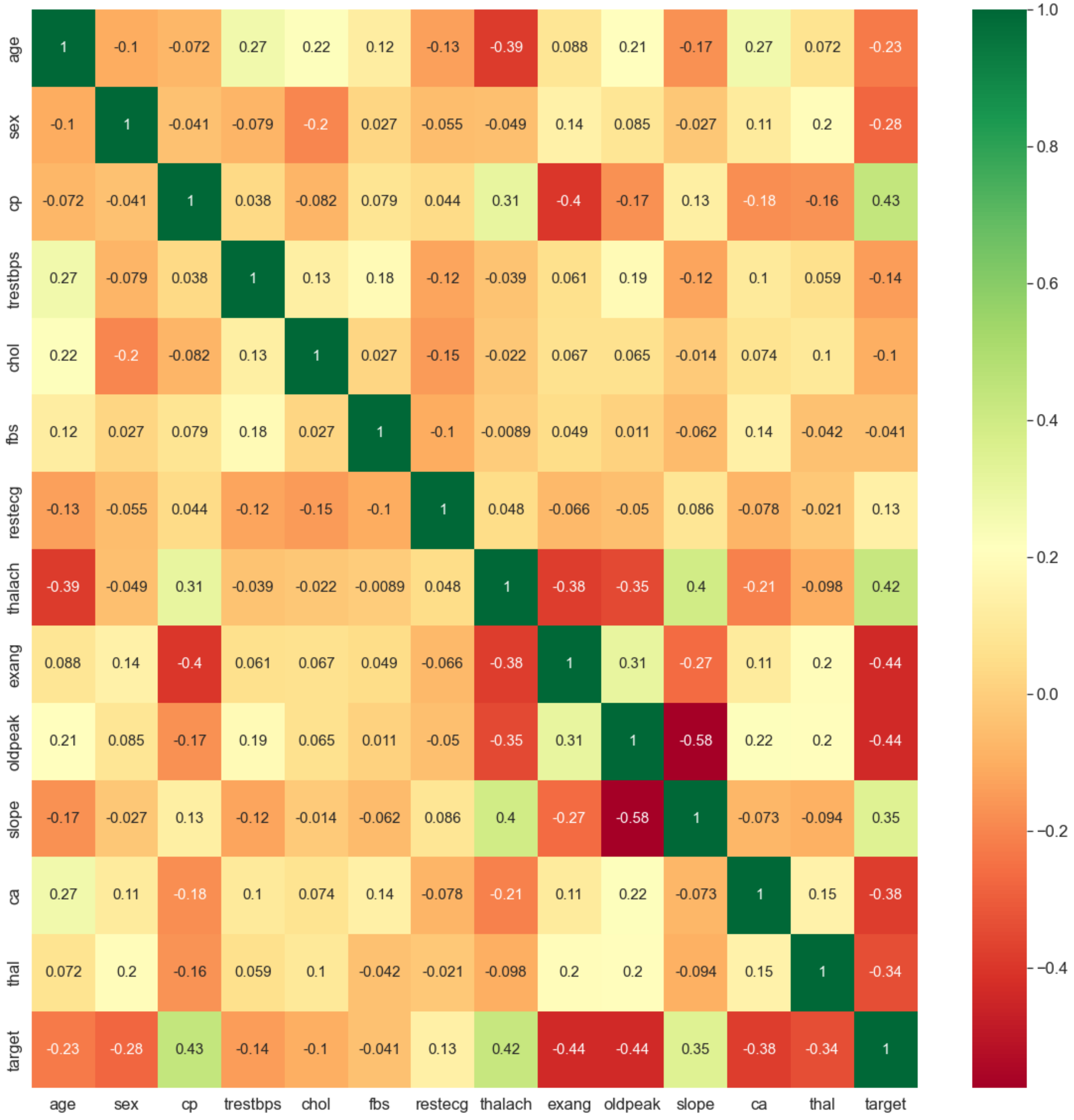




- ➢ Plotting target vs count.



- Plotting correlation matrix.



	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
age	1	-0.1	-0.072	0.27	0.22	0.12	-0.13	-0.39	0.088	0.21	-0.17	0.27	0.072	-0.23
sex	-0.1	1	-0.041	-0.079	-0.2	0.027	-0.055	-0.049	0.14	0.085	-0.027	0.11	0.2	-0.28
cp	-0.072	-0.041	1	0.038	-0.082	0.079	0.044	0.31	-0.4	-0.17	0.13	-0.18	-0.16	0.43
trestbps	0.27	-0.079	0.038	1	0.13	0.18	-0.12	-0.039	0.061	0.19	-0.12	0.1	0.059	-0.14
chol	0.22	-0.2	-0.082	0.13	1	0.027	-0.15	-0.022	0.067	0.065	-0.014	0.074	0.1	-0.1
fbs	0.12	0.027	0.079	0.18	0.027	1	-0.1	-0.0089	0.049	0.011	-0.062	0.14	-0.042	-0.041
restecg	-0.13	-0.055	0.044	-0.12	-0.15	-0.1	1	0.048	-0.066	-0.05	0.086	-0.078	-0.021	0.13
thalach	-0.39	-0.049	0.31	-0.039	-0.022	-0.0089	0.048	1	-0.38	-0.35	0.4	-0.21	-0.098	0.42
exang	0.088	0.14	-0.4	0.061	0.067	0.049	-0.066	-0.38	1	0.31	-0.27	0.11	0.2	-0.44
oldpeak	0.21	0.085	-0.17	0.19	0.065	0.011	-0.05	-0.35	0.31	1	-0.58	0.22	0.2	-0.44
slope	-0.17	-0.027	0.13	-0.12	-0.014	-0.062	0.086	0.4	-0.27	-0.58	1	-0.073	-0.094	0.35
ca	0.27	0.11	-0.18	0.1	0.074	0.14	-0.078	-0.21	0.11	0.22	-0.073	1	0.15	-0.38
thal	0.072	0.2	-0.16	0.059	0.1	-0.042	-0.021	-0.098	0.2	0.2	-0.094	0.15	1	-0.34
target	-0.23	-0.28	0.43	-0.14	-0.1	-0.041	0.13	0.42	-0.44	-0.44	0.35	-0.38	-0.34	1

Similarly, all the attributes and their correlations are plotted.

(iv.) **Data Preprocessing**

- Segregation of target data and feature data as training and test data.
- Scaling the values in the data to be values between 0 and 1 in which and scale all the values before training the Machine Learning models.

(v.) **Applying Algorithms**

- Comparing 5-machine learning algorithms such as Logistic Regression, Decision tree, Random forest classifier and K- nearest neighbor & MLP as well ,to get the better accuracy to which highest parameter may cause disease.
- For each algorithm, there is a pseudo code helpful to develop any kind of programming language. In python, there is a simple way to establish any kind of algorithm in which simple and short code easier to predict accuracy.

Machine Learning Algorithms:

- The algorithms used in this project is highly helpful to predict the accurate result to detect heart disease in which factors that cause a disease can be detected. The following algorithms have built in this project.
  - **K-Nearest Neighbor algorithm:** KNN is a supervised classifier that carry-outs a observations from within a test set to predict classification labels. KNN is one of the classification technique used whenever there is a classification. It has a few assumptions includes dataset has little noise, labeled and it should contains relevant features. By applying KNN in large datasets takes long time to process. The accuracy gained with this algorithm is **79 %**.
  - **Random Forest Classifier:** Random forest classifier is a powerful tool in the machine learning library. With this classifier, we will be able to get higher accuracy and training time should be less. Initially, we have to build a model and by splitting variables into training and test set. After splitting the data, train the dependent variables and predict the response. By using the random forest classifier, the accuracy predicted result is of **100 %**.

- **Decision tree classifier:** In this algorithm, preprocessing made initially by splitting data into training and test data .Feature scaling can be done because of normalizing the values before prediction. Import a decision tree classifier to fit the training sets of dependent and independent variables in which Gini-index criterion is used to predict the accuracy or response for the test set. The accuracy gained with this algorithm is **90 %**.

- **Logistic Regression:** This algorithm is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables. Logistic regression predicts the output of a categorical dependent variable. The accuracy gained with this algorithm is **92 %**.

- **MLP:** Multi-layer Perceptron (MLP) is a supervised learning algorithm that learns a function f (·) : R m → R o by training on a dataset, where is the number of dimensions for input and is the number of dimensions for output. The accuracy gained with this algorithm is **75 %**.

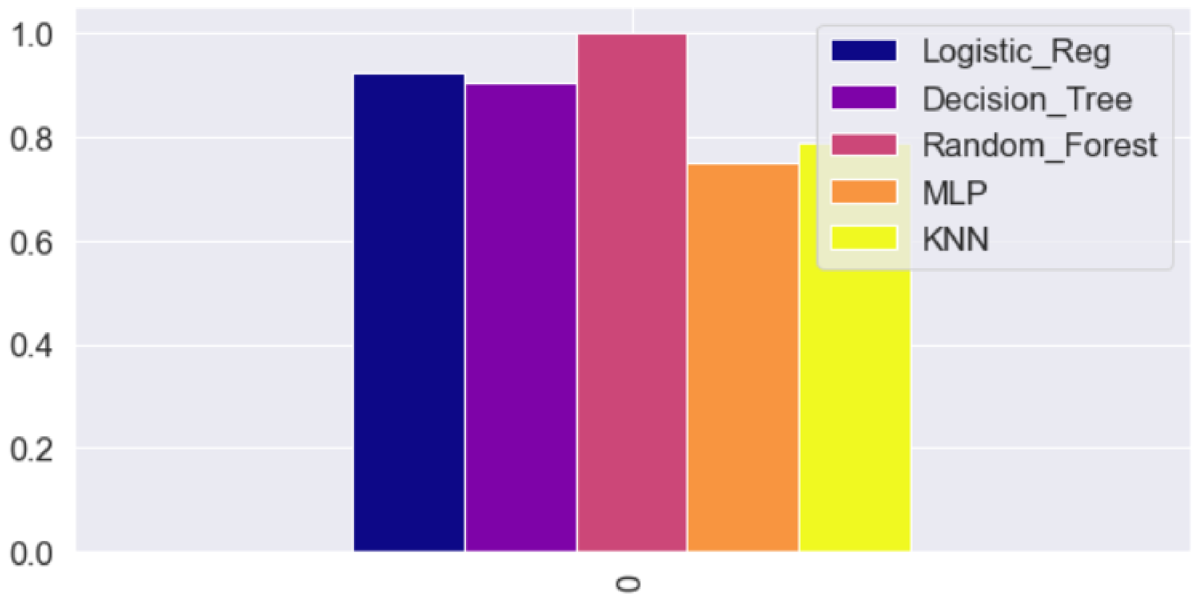
❖ **Our main goal is to predict the accuracy for future problems that the disease may cause and which algorithm gives more accuracy that can be made for the target output counts that a person having Heart disease or not.**

The aim of prediction methodology is to design a model using **Random Forest algorithm** which will produce output or results based on the **14** attributes provided by the medical practitioner.

The imported dataset can be processed and correlated to each other and visualize the correlation for each attribute with another attribute to each other.

Now , after all these visualizations and the processing of the dataset and applying the above discussed machine learning algorithms and finding the accuracy of predictions and finally deciding the best algorithm using the confusion matrix, & use the same to build upon the system.

Comparative Plotting of all the Algorithms used:



Tools & Technology used:

- ❖ Python libraries are the pre-requisites for making prediction in which **SKLEARN** is basically used in machine learning predictions. From **SKLEARN**, we will be able to preprocess the data by splitting the attributes and labels, test and train data, and also scale the values in the data to be values between 0 and 1 by importing the library **STANDARDSCALAR**.

- ❖ Also **SEABORN** is another library used in our prediction to correlate each and every attributes together. At last the confusion matrix decides accuracy perfectly by importing **CONFUSION MATRIX**.

- ❖ And in the visualization part the library used is matplotlib and the other libraries used for numerical computations & analysis of dataset are **NUMPY** and the **PANDAS**.

❖ Model description through Flow Diagram:

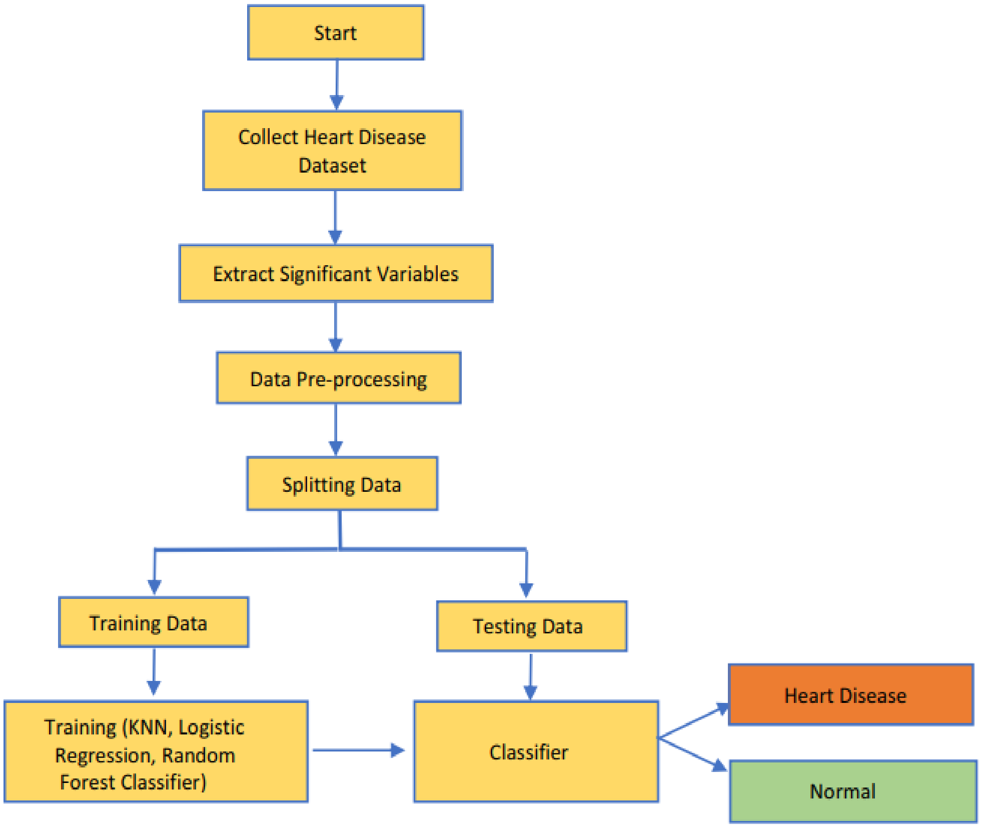

Start
Collect Heart Disease Dataset
Extract Significant Variables
Data Pre-processing
Splitting Data
Training Data
Testing Data
Training (KNN, Logistic Regression, Random Forest Classifier)
Classifier
Heart Disease
Normal

Figure 1. Proposed Model

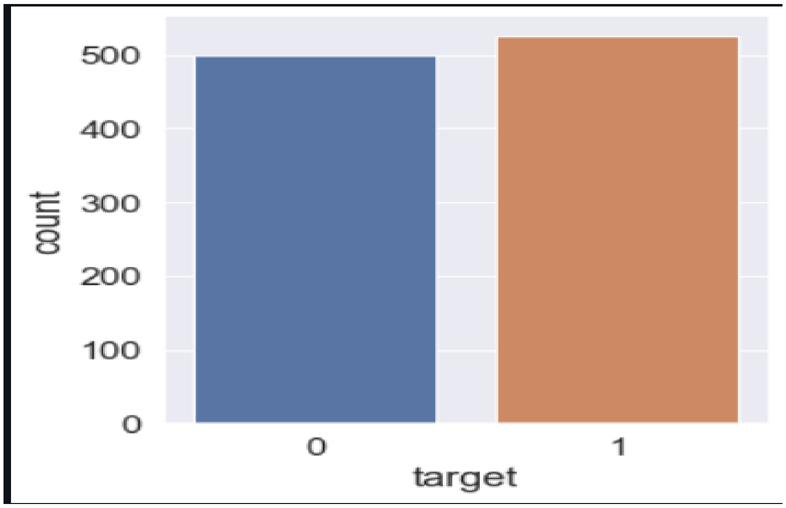
CHAPTER-4
Conclusion and Future Scope

Conclusion

A Heart disease detection model has been developed using five ML classification modelling techniques. This project predicts people with cardiovascular disease by extracting the patient medical history that leads to a fatal heart disease from a dataset that includes patients' medical history such as chest pain, sugar level, blood pressure, etc. This Heart Disease detection system assists a patient based on his/her clinical information of them been diagnosed with a previous heart disease**. The algorithms used in building the given model are Logistic regression, Random Forest Classifier ,KNN, MLP and Decision Tree.** The accuracy of our model using Random Forest classifier is 100 %. Use of more training data ensures the higher chances of the model to accurately predict whether the given person has a heart disease or not. By using these, computer aided techniques we can predict the patient fast and better and the cost can be reduced very much. There are a number of medical databases that we can work on as these Machine learning techniques are better and they can predict better than a human being which helps the patient as well as the doctors. Therefore, in conclusion this project helps us predict the patients who are diagnosed with heart diseases by cleaning the dataset ,on our model which is better than the previous models having an accuracy of 100 % in Random Forest Classifier. Also, it is concluded that accuracy of Random Forest Classifier is highest between the five algorithms that we have used i.e. 100 %. Figure below shows of people that are listed in the dataset are suffering from Heart Disease and how many are not suffering.

Future Scope

With the rising number of deaths due to heart disease, it is becoming increasingly important to build a system that can effectively and accurately forecast heart disease. The project's goal was to find the most efficient machine learning algorithm for detecting heart diseases. Using a typical dataset from kaggle, this study analyses the accuracy score of Decision Tree, Logistic Regression, Random Forest, MLP, and KNN algorithms for predicting heart disease. According to the findings of this study, the Random Forest algorithm is the most efficient algorithm for predicting heart disease, with a score of 100 percent accuracy. In the future, the study might be improved by creating a web application based on the Random Forest method and employing a larger dataset than the one used in this analysis, which would help to deliver better results and aid health professionals in successfully and efficiently forecasting heart disease.



References:

[1] Soni J, Ansari U, Sharma D & Soni S (2011). Predictive data mining for medical diagnosis: an overview of heart disease prediction. International Journal of Computer Applications, 17(8), 43-8

[2] Dangare C S & Apte S S (2012). Improved study of heart disease prediction system using Hybrid Machine techniques. International Journal of Computer Applications, 47(10), 44-8.

[3] Ordonez C (2006). Association rule discovery with the train and test approach for heart disease prediction. IEEE Transactions on Information Technology in Biomedicine, 10(2), 334-43.

[4] Shinde R, Arjun S, Patil P & Waghmare J (2015). An intelligent heart disease prediction system using k-means clustering and Naïve Bayes algorithm. International Journal of Computer Science and Information Technologies, 6(1), 637-9.

[5] Bashir S, Qamar U & Javed M Y (2014, November). An ensemble-based decision support framework for intelligent heart disease diagnosis. In International Conference on Information Society (i-Society 2014) (pp. 259-64). IEEE. ICCRDA 2020 IOP Conf. Series: Materials Science and Engineering 1022 (2021) 012072 IOP Publishing doi:10.1088/1757-899X/1022/1/012072 9 **[6**] Jee S H, Jang Y, Oh D J, Oh B H, Lee S H, Park S W & Yun Y D (2014). A coronary heart disease prediction model: the Korean Heart Study. BMJ open, 4(5), e005025.

[7] Ganna A, Magnusson P K, Pedersen N L, de Faire U, Reilly M, Ärnlöv J & Ingelsson E (2013). Multilocus genetic risk scores for coronary heart disease prediction. Arteriosclerosis, thrombosis, and vascular biology, 33(9), 2267-72.

[8] Jabbar M A, Deekshatulu B L & Chandra P (2013, March). Heart disease prediction using lazy associative classification. In 2013 International Mutli-Conference on Automation, Computing,Communication, Control and Compressed Sensing (iMac4s) (pp. 40- 6). IEEE.

[9] Dangare Chaitrali S and Sulabha S Apte. "Improved study of heart disease prediction system using data mining classification techniques." International Journal of Computer Applications 47.10 (2012): 44-8.

[10] Soni Jyoti. "Predictive data mining for medical diagnosis: An overview of heart disease prediction." International Journal of Computer Applications 17.8 (2011): 43-8.

[11] Chen A H, Huang S Y, Hong P S, Cheng C H & Lin E J (2011, September). HDPS: Heart disease prediction system. In 2011 Computing in Cardiology (pp. 557-60). IEEE.

[12] Parthiban, Latha and R Subramanian. "Intelligent heart disease prediction system using CANFIS and genetic algorithm." International Journal of Biological, Biomedical and Medical Sciences 3.3 (2008).

[13] Wolgast G, Ehrenborg C, Israelsson A, Helander J, Johansson E & Manefjord H (2016). Wireless body area network for heart attack detection [Education Corner]. IEEE antennas and propagation magazine, 58(5), 84-92.
[14] Patel S & Chauhan Y (2014). Heart attack detection and medical attention using motion sensing device -kinect. International Journal of Scientific and Research Publications, 4(1), 1-4. **[15]** Zhang Y, Fogoros R, Thompson J, Kenknight B H, Pederson M J, Patangay A & Mazar S T (2011). U.S. Patent No. 8,014,863. Washington, DC: U.S. Patent and Trademark Office. **Etc.**