**A**

**Semester Project-III**

**Report**

**On**

# “Heart Attack Prediction using Machine Learning”

# By

1. Ram Girish Paliwal.
2. Lalit Gulab Sangore.
3. Kanchan Prakash Mahajan.



# Department of Computer Science & Engineering (Data Science)

**The Shirpur Education Society’s**

**R. C. Patel Institute of Technology, Shirpur - 425405.**

**[2022-23]**

**A**

**Semester Project-III**

**Report**

**On**

**“Heart Attack Prediction using Machine Learning”**

In partial fulfilment of requirements for the degree of

Bachelor of Technology

In

Computer Science & Engineering (Data Science)

**Submitted By**

1. Ram Girish Paliwal.
2. Lalit Gulab Sangore.
3. Kanchan Prakash Mahajan.

**Under the Guidance of**

## Prof. Manisha Patil.



**The Shirpur Education Society’s**

**R. C. Patel Institute of Technology, Shirpur - 425405.**

**Department of Computer Science & Engineering (Data Science) [2022-23]**



**The Shirpur Education Society’s**

**R. C. Patel Institute of Technology**

**Shirpur, Dist. Dhule (M.S.)**

**Department of Computer Science & Engineering (Data Science)**

# CERTIFICATE

This is to certify that the Semester Project-II entitled “**Heart Attack Prediction Using Machine Learning”** has been carried out by team:

1. **Ram Girish Paliwal**
2. **Lalit Gulab Sangore**
3. **Kanchan Prakash Mahajan**

under the guidance of **Prof.** **Manisha Patil** in partial fulfilment of the requirement for the degree of Bachelor of Technology in Computer Science & Engineering (Data Science) (Semester- V) of Dr. Babasaheb Ambedkar Technological University, Lonere during the academic year 2022-23.

**Date:**

**Place: Shirpur**

**Guide Semester Project-III Coordinator**

**Prof. Manisha Patil Prof. Priti Sanjekar**

**H.O.D. Director**

**Prof. Dr. R. B. Wagh Prof. Dr. J. B. Patil**

## ACKNOWLEDGEMENT

We are grateful and sincerely appreciate the effort of our respected project in-charge **Prof. Priti Sanjekar** who acted as a fulcrum for us and supported us during the ups and downs of our project. We are profoundly indebted to our project guide **Prof. Manisha Patil.** for innumerable acts of timely advice, encouragement and we sincerely express our gratitude to him.

We extend our sincere thanks to **Prof. Dr. R. B. Wagh**, Head of Department, Computer Engineering, for providing us the right ambience for carrying out this work.

Many people have contributed to the success of this project work. Although a single sentence hardly suffices, we would like to thank Almighty God for blessing us with his grace. We would like to express our sincere appreciation to our Principal, **Dr. J. B. Patil**, who spend his precious time and gave some ideas and suggestion towards this project. We express our immense pleasure and thankfulness to all the teachers and staff of the **Department of Computer Science and Engineering (Data Science)** for their cooperation and support.

**Project Team:**

**Ram Girish Paliwal**

**Lalit Gulab Sangore**

**Kanchan Prakash Mahajan**

## PAGE INDEX

|  |  |  |  |
| --- | --- | --- | --- |
| **Chapter** |  | **Topic** | **Page No.** |
|  |  | ABSTRACT | 1 |
| 1. |  | Introduction | 2 |
| 2. |  | Need | 3 |
| 3. |  | Software And Hardware Requirements | 4 |
|  | 3.1 | Software Requirements |  |
|  | 3.2 | Hardware Requirements |  |
| 4. |  | Literature Survey | 5 |
| 5. |  | Implementation Details | 11 |
|  | 5.1 | Algorithms |  |
|  | 5.2 | Implementation steps |  |
|  | 5.3 | Snapshots of Outputs |  |
| 6. |  | CONCLUSION | 20 |
|  |  | BIBLIOGRAPHY |  |

## ABSTRACT

Low heart rate causes a risk of death, heart disease, and cardiovascular diseases. Therefore, monitoring the heart rate is critical because of the heart’s function to discover its irregularity to detect the health problems early. Rapid technological advancement (e.g., artificial intelligence and stream processing technologies) allows healthcare sectors to consolidate and analyze massive health-based data to discover risks by making more accurate predictions. Therefore, this work proposes a real-time prediction system for heart rate, which helps the medical care providers and patients avoid heart rate risk in due time. The proposed model uses machine learning basics and is programmed in python and uses libraries such as pandas, warnings, Train Test Split and also uses other libraries such as SMOTE, logical regression and many others. It declares its result from a looking at the lifestyle of the users and other aspects in their life such as smoking and other damaging habits and also checks whether the individual is diagnosed with other diseases. After checking such data and following its algorithm it arrives at a conclusion, indicating whether the individual will have a heart attack in the next 10 years or not.

**1**

**CHAPTER – 1**

## Introduction

As per the World Health Organization (WHO) , heart disease is one of the main death reasons in the world, which reported around 17.7 million deaths worldwide every year. In particular, the heart disease (HD) is caused by any condition affecting the heart when the heart is unable to do its normal function. Consequently, technology plays a vital role to early detect the high heart rate to avoid the risk of heart disease progression. Early diagnosis of HD is essential because HD treatment is most effective during the early stages of the disease. To date, many research works have been done using statistical comparative analysis, machine learning, and historical data to estimate the risk factors of diseases. The proposed models, including ours still have improvements it can make, for example our model uses an index and depending on that it may not always be accurate but we still can get very close as the model follows a particular algorithm derived from previous hospital data.

2

**CHAPTER – 2**

**NEED**

**NEED OF HEART ATTACK PREDICTION**

In today’s world due to unhealthy lifestyle and practices like smoking & alcohol have raised the risk of heart attacks worldwide. This risk can be prevented if the person follows certain routine and medications.

Physicians have been trying to predict heart attacks for as long as there have been heart attacks. Traditionally, they have relied on standard assessments of cholesterol, blood pressure, lifestyle factors and health conditions such as diabetes to predict whether a patient is likely to suffer a heart attack.

so the need of heart attack prediction arises. The objective of this study is to effectively predict if the patient suffers from heart disease. The health professional enters the input values from the patient’s health report. The data is fed into model which predicts the probability of having heart disease.

3

**CHAPTER – 3**

### Software and Hardware Requirements

#### 3.1 Software Requirements

|  |  |  |
| --- | --- | --- |
| **Sr.**  **No.** | **Name of Resource** | **Specifications** |
| **1.** | Operating System | Windows 11 |
| **2.** | OpenCV | Python |
| **3.** | Pytorch | Python |

#### 3.2 Hardware Requirements

|  |  |  |
| --- | --- | --- |
| **Sr.**  **No.** | **Name of Resource** | **Specifications** |
| **1.** | Computer system | Intel (R) core (TM) i5-4170  CPU@ 3.20GHz(2 in 1 display) |
| **2.** | Primary Memory | 8GB |
| **3.** | Secondary Memory | 512 GB SSD |
| **4.** | GPU | NVIDIA GTX 1650 TI - QMAX |

4

### CHAPTER – 4

### Literature Survey

There is number of works has been done related to attack prediction systems using different machine learning algorithms in medical Centres.

Senthil Kumar Mohan et al,[1] proposed Effective Heart Attack Prediction Using Hybrid Machine Learning Techniques in which strategy that objective is to finding critical includes by applying Machine Learning bringing about improving the exactness in the expectation of cardiovascular malady. The expectation model is created with various blends of highlights and a few known arrangement strategies. We produce an improved exhibition level with a precision level of 88.7% through the prediction model for heart attack with hybrid random forest with a linear model (HRFLM) they likewise educated about Diverse data mining approaches and expectation techniques, Such as, KNN, LR, SVM, NN, and Vote have been fairly famous of late to distinguish and predict heart attack.

Sonam Nikhar et al [2] has built up the paper titled as Prediction of Heart Attack Using Machine Learning Algorithms by This exploration plans to give a point by point portrayal of NaÃ¯ve Bayes and decision tree classifier that are applied in our examination especially in the prediction of Heart Attack. Some analysis has been led to think about the execution of prescient data mining strategy on the equivalent dataset, and the result uncovers that Decision Tree beats over Bayesian classification system.

Aditi Gavhane, Gouthami Kokkula, Isha Pandya, Prof. Kailas Devadkar (PhD), [3] Prediction of Heart Attack Using Machine Learning, In this paper proposed system they used the neural network algorithm multi-layer

5

perceptron (MLP) to train and test the dataset. In this algorithm there will be multiple layers like one for input, second for output and one or more layers are hidden layers between these two input and output layers. Each node in input layer is connected to output nodes through these hidden layers. This connection is assigned with some weights. There is another identity input called bias which is with weight b, which added to node to balance the perceptron. The connection between the nodes can be feedforwarded or feedback based on the requirement.

Abhay Kishore et al,[4] developed Heart Attack Prediction Using Deep Learning in which This paper proposes a heart attack prediction system using Deep learning procedures, explicitly Recurrent Neural System to predict the probable prospects of heart related infections of the patient. Recurrent Neural Network is a very ground-breaking characterization calculation that utilizes Deep Learning approach in Artificial Neural Network. The paper talks about in detail the significant modules of the framework alongside the related hypothesis. The proposed model deep learning and data mining to give the precise outcomes least blunders. This paper gives a bearing and point of reference for the advancement of another type of heart attack prediction platform. Prediction stage.

Lakshmana Rao et al,[5] Machine Learning Techniques for Heart Attack Prediction in which the contributing elements for heart attack are more (circulatory strain, diabetes, current smoker, high cholesterol, etc..). So, it is difficult to distinguish heart attack. Different systems in data mining and neural systems have been utilized to discover the seriousness of heart attack among people. The idea of CHD ailment is bewildering, in addition, in this manner, the attack must be dealt with warily. Not doing early identification, may impact the heart or cause sudden passing. The perspective of therapeutic science furthermore, data burrowing is used for finding various

6

sorts of metabolic machine learning a procedure that causes the framework to gain from past information tests, models without being expressly customized. Machine learning makes rationale dependent on chronicled information.

Mr. Santhana Krishnan.J and Dr. Geetha.S, [6] Prediction of heart attack using machine learning algorithm This Paper predicts heart attack for Male Patient using Classification Techniques. The detailed information about Coronary Heart attacks such as its Facts, Common Types, and Risk Factors has been explained in this paper. The Data Minin tool used is WEKA (Waikato Environment for Knowledge Analysis), a good Data Mining Tool for Bioinformatics Fields. The all three available Interface in WEKA is used here; Naive Bayes, Artificial Neural Networks and Decision Tree are Main Data Mining Techniques and through this techniques heart attack is predicted in this System. The main Methodology used for prediction is Decision Trees like CART, C4.5, CHAID, J48, ID3 Algorithms, and Naive Bayes Techniques.

Avinash Golande et al,[7] proposed Heart Attack Prediction Using Effective Machine Learning Techniques in which Specialists utilize a few data mining strategies that are available to support the authorities or doctors distinguish the heart attack. Usually utilized methodology utilized are decision tree, k- closest and NaÃ¯ve Bayes. Other unique characterization-based strategies utilized are packing calculation, Part thickness, consecutive negligible streamlining and neural systems, straight Kernel self- arranging guide and SVM (Bolster Vector Machine). The following area obviously gives subtleties of systems that were utilized in the examination.

V.V. Ramalingam et Al,[8] proposed Heart attack prediction using machine learning techniques in which Machine Learning algorithms and techniques have been applied to various medical datasets to automate the analysis of

7

large and complex data. Many researchers, in recent times, have been using several machine learning techniques to help the health care industry and the professionals in the diagnosis of heart related attacks. This paper presents a survey of various models based on such algorithms and techniques and analyse their performance. Models based on supervised learning algorithms such as Support Vector Machines (SVM), K- Nearest Neighbour (KNN), NaÃ¯ve Bayes, Decision Trees (DT), Random Forest (RF) and ensemble models are found very popular among the researchers and systems have been applied to different clinical datasets to robotize the investigation of huge and complex information. Numerous scientists, as of late, have been utilizing a few Machine Learning algorithms and techniques have been applied to various medical datasets to automate the analysis of large and complex data. Many researchers, in recent times, have been using several machine learning techniques to help the health care industry and the professionals in the diagnosis of heart related attacks. This paper presents a survey of various models based on such algorithms and techniques and analyze their performance. Models based on supervised learning algorithms such as Support Vector Machines (SVM), K- Nearest Neighbour (KNN), NaÃ¯ve Bayes, Decision Trees (DT), Random Forest (RF) and ensemble models are found very popular among the researchers. strategies to enable the wellbeing to mind industry and the experts in the analysis of heart related sicknesses. This paper presents a review of different models dependent on such calculations and methods and analyze their exhibition. Models in light of directed learning calculations, for example, Support Vector Machines (SVM), K- Nearest Neighbour (KNN), NaÃ¯ve Bayes, Decision Trees (DT), Random Forest (RF) and group models are discovered extremely well known among the scientists

8

.

9

|  |  |  |  |
| --- | --- | --- | --- |
|  | | | |
| YEAR | AUTHOR | PURPOSE | TECHNIQUES USED | | ACCURACY |
| 2016 | Sonam | Prediction | 1) NaÃ¯ve Bayes Classifier | | Decision tree |
|  | Nikhar[2] | of Heart |  | | has better |
|  |  | Attack |  | | accuracy as |
|  |  | Using | 2)Decision tree | | compared to |
|  |  | Machine |  | | naÃ¯ve Bayes |
|  |  | Learning |  | | classifier. |
|  |  | Algorithms |  | |  |
| 2018 | Aditi | Prediction | 1) Multi-layer | | The MLP |
|  | Gavhane[3] | of Heart | perceptron | | gives best |
|  |  | Attack | algorithm | | Accuracy |
|  |  | Using |  | |  |
|  |  | Machine |  | |  |
|  |  | Learning |  | |  |
| 2018 | V.V. | Heart | 1) NaÃ¯ve Bayes | | SVM has |
|  | Ramalingam[8] | attack |  | | more |
|  |  | prediction | 2) Support | | accuracy than |
|  |  | using | Vector Machine | | other |
|  |  | machine |  | | techniques. |
|  |  | learning | 3) K Nearest | |  |
|  |  | techniqus | Neighbour | |  |
|  |  |  | 4) Decision Tree | |  |
|  |  |  |  | |  |
|  |  |  | 5) Random Forest | |  |
|  |  |  |  | |  |
|  |  |  | 6) Ensemble | |  |
|  |  |  | Model | |  |
| 2019 | Abhay | Heart | 1) RNN | | RNN |
|  | Kishore1[4] | Attack |  | | accuracy 92% |
|  |  | Prediction |  | |  |
|  |  | Using |  | |  |
|  |  | Deep |  | |  |
|  |  | Learning |  | |  |

### CHAPTER – 5

### Implementation Details

#### 5.1 Algorithm

Step 1: Start

Step 2: Creating Data

Step 3: Labelling Images

Step 4: Training a CNN on the capture dataset

Step 5: Predict the gesture

Step 6: Stop

11

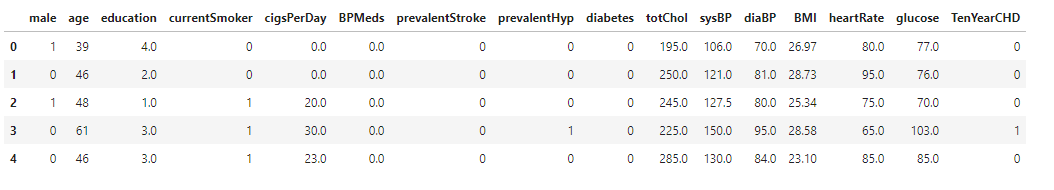
**5.2 Implementation Details**

To implement this model, we have used ML with python and many other advanced libraries of python like pandas (to handle data frame and importing data), sklearn (to implement statistical modeling and ML models), warnings (to handle warnings which we get while program is running).

Now we will be seeing the step-by-step implementation of our ML model.

Step 1: Importing Libraries and Data.

We will import libraries like pandas and warnings to import our dataset and to handle warnings in our program.  
Then we will import our dataset.  
  
 

Our imported dataset will look like as shown below:

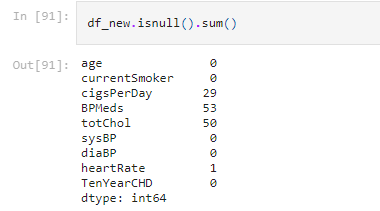
12

Step 2: Selecting necessary columns.

We will select only required columns for prediction. Like 'age','currentSmoker','cigsPerDay','BPMeds','totChol','sysBP','diaBP','heartRate','TenYearCHD'.

Step 3: Exploratory Data Analysis.  
  
We will do Exploratory Data Analysis (EDA) on our dataset. EDA is applied to investigate the data and summarize the key insights.  
In this we will be checking if our data has missing values or not. If it has missing values then we are going to handle them using mean or median or by simply dropping them from our dataset. First, we will check the skewness of columns which have missing value if skewness is between 1 to –1 then we will use mean to handle the missing values, If the skewness is more than 1 or less than –1 then we will use median to handle missing values.Then we will check shape of our data frame and all information of our data frame like data types contained by columns memory used by our data frame etc.

Checking for missing values:



13

Handling missing values:



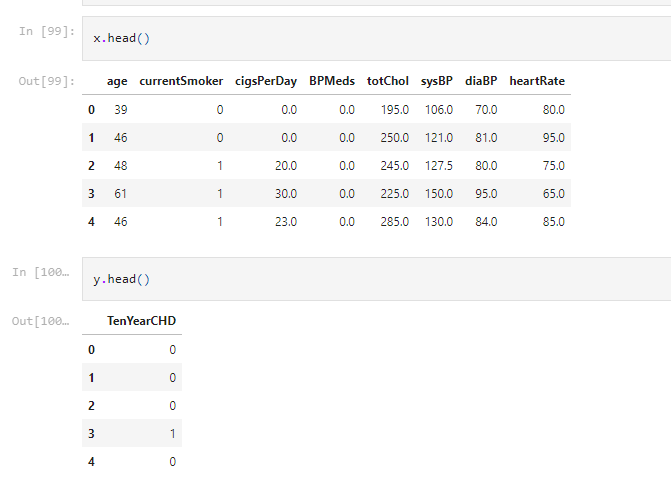
Step 4: Identifying target variable.

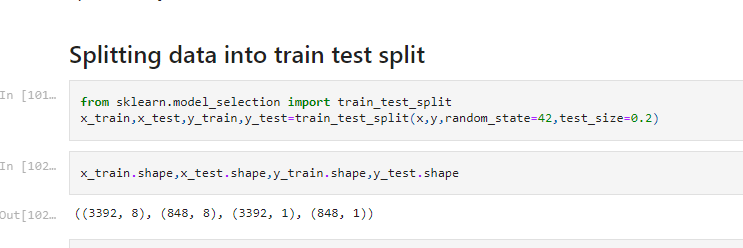
Here we will separate the target columns and other columns.  
In this data frame our target column is ‘TenYearCHD’   
We will store this column is ‘y’ and we store other columns in ‘x’ by dropping ‘TenYearCHD’ column.



14

After doing the above step we will get the following output:

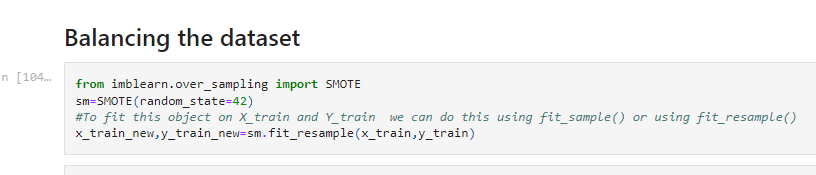


Step 5: Splitting Data into train test split.  
Now we will split our data into training data and test data.  
We will do so by using model\_selection method using sklearn library.   
We will use 80% of data for training and 20% for testing the model.

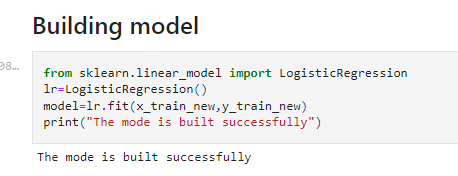
15

Step 6: Balancing the data  
Check whether our data is balanced or imbalanced for prediction by using value\_count method.  
If data is imbalanced, we will balance it by using SMOTE (Synthetic Minority Oversampling Technique). It is a statistical technique for increasing the number of cases in your dataset in a balanced way.  
We will import SMOTE from imblearn library.  
Then we will fit our previous training data into it to balance it.

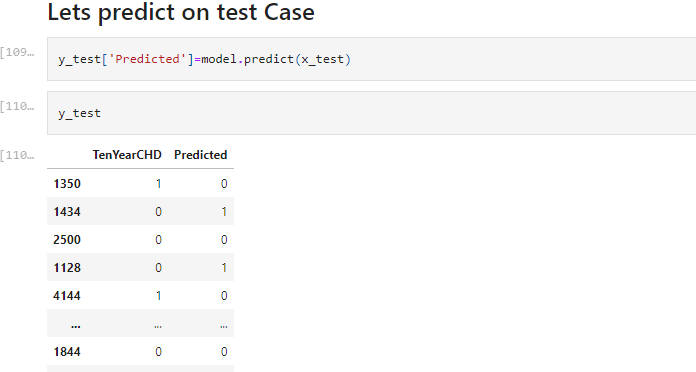
Then, we will check again whether our data got balanced or not.



**Step 7: Building Model.**

After balancing the data. We will build or model. In this ML model we will use Logistic Regression algorithm To use Logistic Regression algorithm we will use linear\_model class from sklearn.  
After building model we will fit out training data into it.

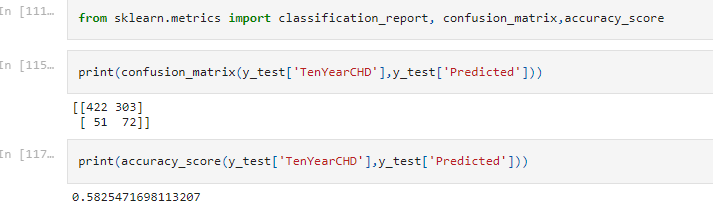
16

**Step 8: Predicting on test case.**  
We will use our test data to predict the values.  


**Step 9: Model Evaluation.**

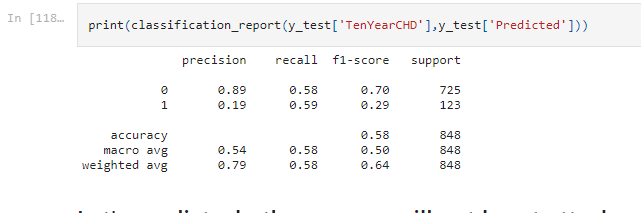
Model evaluation is the process of using different evaluation metrics to understand a machine learning model's performance, as well as its strengths and weaknesses.  
We will use confusion\_matrix , classification\_report , accuracy\_score to check the performance of our model.  
We will import this by using metrics module from sklearn.

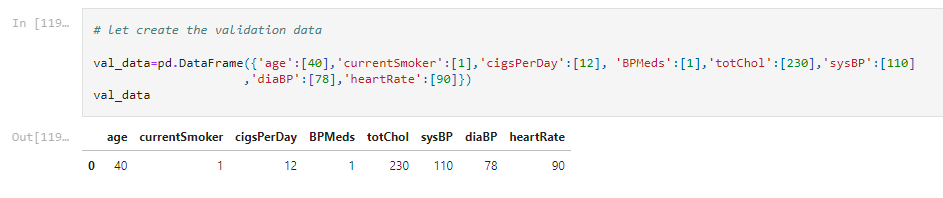
Confusion matrix and Accuracy score:



17

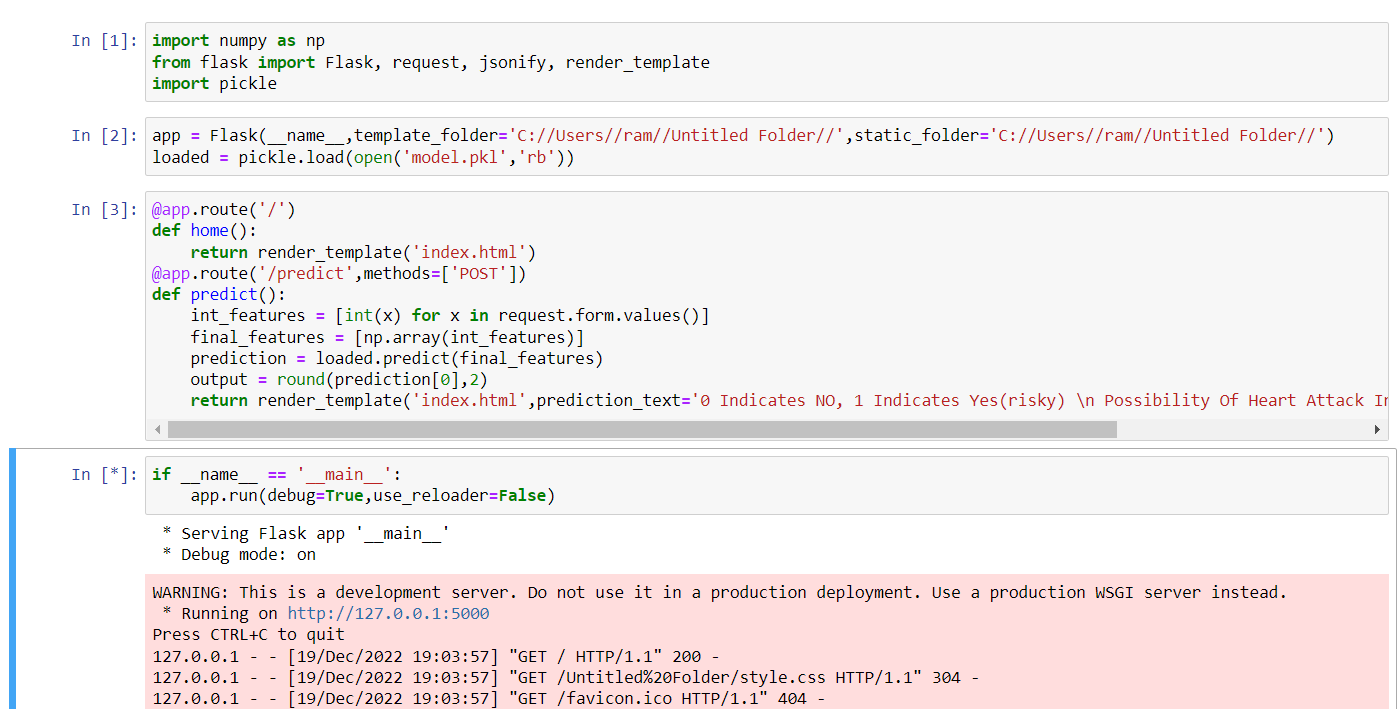
Classification report:



**Step 10: Predicting for Validation data.**  
  
In this step we will created validation data (Validation data provides the first test against unseen data, allowing data scientists to evaluate how well the model makes predictions based on the new data).

In this way we have built our model successfully which may predict whether the person will have a heart attack in the next ten years or not.

18

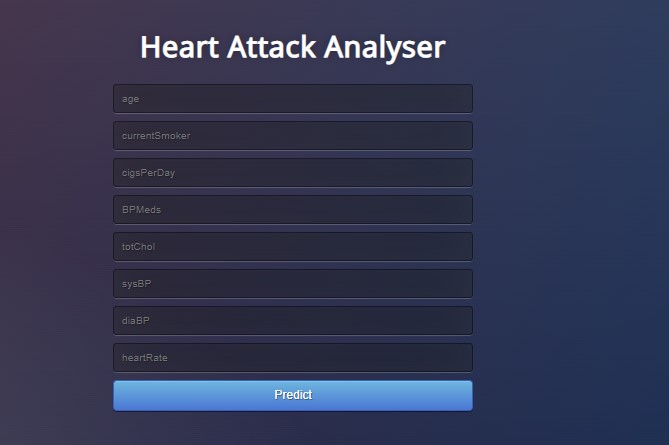


This is a flask program in which we give index.HTML to the render\_template function. Now when we implement this it will create url that will directly redirect to the index.HTML. we rename this file as app.py now when we run this app.py file that is a flask file it will throw an exception that is SystemExit 1 so we use reloader as false that will handle this exception of SystemExit.

Now after executing this app.py file that is a flask it will directly to the index.HTML file.

19

**5.3 Snapshots of output**

****

****

20

**CHAPTER – 6**

## Conclusion

In this project we have used a Machine learning algorithm to predict the Heart Attack .With the help of Machine learning we can easily handle huge amounts of clinical data for best prediction . Men seem to be more susceptible to heart attack than women. Every 1 in 4 men are likely to have heart disease whereas in case of women every 1 out of 5 women is likely to have heart disease . Increase in Age, number of cigarettes and Blood Pressure also show increasing odds of having heart attack. Interestingly total cholesterol shows no significant change in the odds of CHD. This could be due to the presence of good cholesterol in the total cholesterol reading. The early prognosis of cardiovascular diseases would aid in making better decisions on lifestyle changes in high risk patients and in turn reduce any future heart problems.

In the end Logistic regression is providing the more accurate prediction of the heart attack.So we can say that with help of several kinds of Machine Algorithm we can easily predict many kinds of related information.

21

### BIBLIOGRAPHY

[1] Soni J, Ansari U, Sharma D & Soni S (2011). Predictive data mining for medical diagnosis: an overview of heart attack 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 data mining classification techniques. International Journal of Computer Applications, 47(10), 44-8.

[3] Dataset reference from Kaggle (<https://kaggle.com>)

[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] 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.

22