

# **VISUALIZING AND PREDICTING HEART DISEASES WITH AN INTERACTIVE DASHBOARD**

**TEAM ID: PNT2022TMID20513**

## **TEAM MEMBERS**

- 1.SHENBAGARUBAN
- 2.JEGAN
- 3.SUBASHINI
- 4.KARPAGAMAHESWARI

## **1.INTRODUCTION**

### **1.1 PROJECT OVERVIEW**

The leading cause of death in the developed world is heart disease. Therefore, there needs to be work done to help prevent the risks of having a heart attack or stroke. Analytics is an essential technique for any profession because it predicts the future and uncovers hidden patterns. In recent years, data analytics has been regarded as a cost-effective technology, and it plays an important role in healthcare, including new research findings, emergency situations, and disease outbreaks. The use of analytics in healthcare improves care by facilitating preventive care, and EDA is a critical step when analysing data. The risk factors that cause heart disease are considered and predicted using the K-means algorithm, and the analysis is carried out using publicly available heart disease data. The K-means clustering algorithm, in conjunction with data analytics and visualization tools, is used to predict heart disease. Pre-processing methods, classifier performance, and evaluation metrics are all covered. The visualized data in the result section shows that the prediction is correct

### **1.2 PURPOSE**

The objective of this project is to check whether the patient is likely to be diagnosed with any cardiovascular heart diseases based on their medical attributes such as gender, age, chest pain, fasting sugar level, etc. Typically, the patient's physical examination, signs, and symptoms of heart disease serve as the foundation for the diagnosis. Due to their lifestyle choices and the state of the environment today, individuals are susceptible to many diseases. To prevent the severity of these disorders, early detection and prediction of their occurrence are crucial. Predictive analytics in healthcare can raise the standard of care, gather more clinical data for individualized treatment, and correctly identify each patient's medical condition. For the purpose of making wise decisions, healthcare businesses gather enormous amounts of data that may contain some hidden information. Some sophisticated data mining techniques are utilized to deliver accurate results and make data-driven judgments. In this study, we established a project for estimating the degree of heart disease risk using a neural network. For prediction, the algorithm makes use of 12 medical variables, including age, sex, blood pressure, cholesterol, and obesity.

## 2.LITERATURE SURVEY

<b>TITLE</b>	<b>Author</b>	<b>Year released</b>	<b>Concept</b>
PREDICTING HEART DISEASE	M. Preethi1, Dr. J. Selvakumar	2008	This paper describes various methods of data mining, bigdataand machine learning models for predicting the heart disease.Dataminingand machine learning plays an important role in building an important model for medical systemt opredict heart diseaseor cardiovascular disease.Medical experts can helpthe patients by detecting the cardiovascular disease before occurring.Now-a-daysheart disease is one of the most significant causes of fatality. The prediction of heart disease is challenge in the clinical area. Buttime totime, several techniques are discovered to predict the heart disease in data mining.

Design and Development of Real-Time Heart Disease Prediction System for Elderly People Using Machine Learning	Guttappa Sajjan	2019	Everyday blood is pumped to all parts of the body by the heart. It beats 100,000 times and pumps around 19,000 litres of blood through our body (Lett, Ali and Whooley, 2008). The blood delivers oxygen and nutrients to our tissues and carries away waste. Various types of heart diseases are caused due to the abnormalities in normal blood flow. The heart-related diseases are which are commonly known as cardiovascular disease (CVD). World Health Organization reports that CVD contributes to 31% of the overall deaths compared to other diseases. Percentage of deaths due to CVD is found to be increasing with age. The percentage of deaths due to CVD is 11.8% for people between 20-39 years, 38.55% for those in 40-59 years of age and 73.3% for those between 60-79 years (Nag and Ghosh, 2016). Therefore, it would be very useful to have a system to detect cardiac abnormalities at an early stage and necessary tools to predict heart disease can save many lives.
Heart Disease Prediction System Using Machine Learning	Ranjit Shrestha <sup>1</sup> and Jyotir Moy Chatterjee <sup>2</sup>	2019	The major killer cause of human death is Heart Disease (HD). Many people die due to this disease. Lots of researchers have been discovering new technologies to prognosticate the disease early before it's too late for helping healthcare as well as people. These processes are still under research phase. Machine Learning (ML) is faster-emerging technology of Artificial Intelligence. For predicting HD, a lot of research scholars contribute their effort in Web-based Heart Disease Prediction System (HDPS) by applying DT and NB ML algorithms. We are using the UCI repository HD dataset to train a model by comparing DT and NB algorithm for HDPS Web application. The main aim of this project is to build an efficient prediction model and deploy for prediction of disease. An HDPS Model is built by using NB algorithm that provides 88.163% accuracy among others.
	1. Animesh Hazra, 2. Subrata		A popular saying goes that we are living in an "information age". Terabytes of data are produced every day. Data mining is the process which turns

Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques:	Kumar Mandal,3. Amit Gupta,4.Arko mita Mukherjee,5. Asmita Mukherjee	2010	a collection of data into knowledge. The health care industry generates a huge amount of data daily. However, most of it is not effectively used. Efficient tools to extract knowledge from these databases for clinical detection of diseases or other purposes are not much prevalent.The aim of this paper is to summarize some of the current research on predicting heart diseases using data mining techniques, analyse the various combinations of mining algorithms used and conclude .
BIG DATA ANALYTICS IN HEART DISEASES PREDICTION	Ahmed Ismail, Samir Abdlerazek, I.M. El-Henawy	2005	The healthcare data can be employed to deveda health prediction system that can improve in heart disease prevention. Big data on health care, including patient records, clinical notes, diagnosis, parents and family past ailments, hospitals, and scan results can aid in the phase of disease identification and prediction. The emerging machine learning method offers an important framework for forecasting cardiac diseases. An advanced Support Vector Machine (SVM) classifier was used by the program to conduct parameter tuning to improve classification accuracy and performance. The proposed work aims to develop a real-time prediction system for health issues based on big medical data processing on the cloud. In the proposed scalable system, the medical parameters are sent to Apache Spark to extract the attributes from the data and to apply the proposed machine learning algorithm aiming to predict the healthcare risks and send them as alerts and recommendations to the users and the healthcare providers as well.

## 2.2 REFERENCES

[1]" Predicting the Risk of Heart Failure With EHR Sequential Data Modelling," Bo Jin, Chao Che.

[2]"Heart Attack Prediction and Visualization of Contributing Factors Using

Machine Learning” by Megha Banerjee, Reetodeep Hazra, Suvranil Saha, Megha Bhushan, Subhankar Bhattacharjee.

[3]“A novel approach for heart disease prediction using strength scores with significant Predictors” by Armin Yazdani, Kasturi Dewi Varathan, Yin Kia Chiam, Asad Waqar Malik, and Wan Azman Wan Ahmad.

[4]“Heart Disease Risk Prediction Using Machine Learning Classifiers with Attribute Evaluators” by Karna Vishnu Vardhana Reddy, Irraivan Elamvazuthi, Azrina Abd Aziz, Sivajothi Paramasivam, Hui Na Chua, and S. Pranavanand

### **2.3 Problem statement Definition**

Heart acts a major role in the corporeal organisms. The diseases of the heart want more perfection and exactness for diagnosis and analyses. Heart disease is a dangerous disease. This disease occurs due to various problems such as overpressure, blood sugar, high blood pressure, Cholesterol, etc. in the human body By using Python and machine learning, this paper is analyzed and predicted heart disease. We can predict this disease by using various attributes in the data set. We have collected a data set consisting of 13 elements and 383 individual values to analyze the patient's performance. The main aim of the paper is to get better accuracy to detect heart disease using the ML algorithm,python .

### 3.IDEATION & PROPOSED SOLUTION

#### 3.1 EMPATHY MAP CANVAS

##### SAYS

- 1.Is age become a factor for heart disease
- 2.Variation in breathing rate leads to heart disease
- 3.Weight gain may be due to rise in cholesterol level ?
4. i need a stress free satisfying life to avoid heart diseases

##### THINKS

1. i cant do my regular activities
2. is my improper food habit becomes a issue
3. Am i leading a stressful life
4. my smoking and alcoholic behaviour leads to this situation

##### DOES

1. Tracking a healthy food cycle
2. Avoiding habits like smoking and alcohol
3. Taking BMP test at regular intervals
4. Do regular exercise

##### FEELS

1. Easy to identify and eliminate unhealthy habits
2. Improving a healthy life
3. Controlling unhealthy habits

## 3.2 IDEATION AND BRAIN STORMING

### PROBLEM STATEMENT:

How to increase the accuracy and predict the risk of heart disease using a dashboard?

### BRAINSTORM:

#### PERSON 1

1. Heart disease (HD) is one of the most common diseases nowadays
2. Due to number of contributing factors, such as high blood pressure, diabetes, cholesterol fluctuation, exhaustion and many others
3. Many analytics tools available
4. Number of Machine Learning (ML) algorithms, such as, Naïve Bayes, Stochastic Gradient Descents (SGD)
5. Due to the complex nature of the HD, suggested tests, which has to be prioritized [3], and proposed techniques have to be selected carefully
6. Results showed were consistent for all used classifiers for achievable prediction accuracy with a detection rate of 82%

#### PERSON 2

1. The prediction of cardiac disease helps practitioners make more accurate decisions regarding patients' health
2. .
3. Therefore, the use of machine learning (ML) is a solution to reduce and understand the symptoms related to heart disease.
4. Can use large volume of dataset to increase accuracy
5. must run numerous test on old patients data
6. .
7. can seek the help of doctors in the case study
8. A hybrid intelligent system can be used as a tech stack

### PERSON 3

1. Heart is the most important organ of a human body. It circulates oxygen and other vital nutrients through blood to different parts of the body and helps in the metabolic activities.
2. Thus, even minor problems in heart can affect the whole organism.
3. Researchers are diverting a lot of data analysis work for assisting the doctors to predict the heart problem.
4. So, an analysis of the data related to different health problems and its functioning can help in predicting with a certain probability for the wellness of this organ.
5. we can use dataset all over india to compare with the input dataset.
6. The future scope of this system aims at giving more sophisticated prediction models, risk calculation tools and feature extraction tools for other clinical risks.
7. Using medical profiles such as age, sex, blood pressure and blood sugar it can predict the likelihood of patients getting a heart disease.

### PERSON 4

1. Machine Learning can play an essential role in predicting presence/absence of Locomotor disorders, Heart diseases and more
2. The comparison of the usage frequency and accuracy of different supervised learning algorithms
3. 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.
4. 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.
5. To predict the heart disease, K-means clustering algorithm is used along with data analytics and visualization tool
6. The future scope of this system aims at giving more sophisticated prediction models, risk calculation tools and feature extraction tools for other clinical risks



## GROUP IDEAS

### TECH STACK

1. We can use a wide range of tech stacks available
2. Machine Learning and Artificial intelligence

### REPORT AND ANALYSIS

1. Report and analysis Dashboard must give complete analysis of the given data
2. The analysed data should also be printable

## 3.3 PROPOSED SOLUTION

S.NO	PARAMETER	DESCRIPTION
1.	Problem Statement	The leading cause of death in the developed world is heart disease. As a result, work must be done to reduce the risks of having a heart attack or stroke. It is infeasible for a common man to frequently undergo tests for ECG and so on. Hence, it requires a replacement that is both convenient and dependable.
2.	Idea / Solution description	The proposed solution proposes an interactive dashboard for visualising and forecasting heart disorders, in which the user may observe his/her medical report analysis as well as the projected end result. IBM Cognos will be used to create the dashboard. Machine learning Algorithms will be used to forecast cardiac disease.
3.	Novelty / Uniqueness	Makes recommendations to the user based on that person's medical analysis. Also provide preventive measures when required.
4.	Social Impact/ Customer Satisfaction	It helps with disease prediction at an early stage and frequently alerts the user to their current health status. Both the user and the doctor can benefit from the system's improved decision-making regarding cardiac disease.
5.	Business Model (Revenue	Can be deployed by Hospitals or NFOs, so that it

	Model)	makes the analysis in a fast manner.
6.	Scalability of the Solution	The solution can work effectively on long and small datasets. It can also be changed to predict various other diseases depending on the dataset.

## ALGORITHMS USED:

### 1. KNN Algorithm:

One of the simplest machine learning algorithms, based on the supervised learning method, is K Nearest Neighbour.

The K-NN algorithm makes the assumption that the new case and the existing cases are comparable, and it places the new instance in the category that is most like the existing categories.

A new data point is classified using the K-NN algorithm based on similarity after all the existing data has been stored. This means that utilising the K-NN method, fresh data can be quickly and accurately sorted into a suitable category.

Although the K-NN approach is most frequently employed for classification problems, it can also be

Utilised for regression.

Since K-NN is a non-parametric technique, it makes no assumptions about the underlying data.

### 2. Random Forest Classifier:

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model. As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

### **3.4 Problem Solution fit:**

#### **3.4.1. Customer Segment(s)**

All adults. Especially people who are elder than 40 years and those who are on the verge of getting heart disease due to various factors such as age, obesity, diabetes, stress, etc.

#### **3.4.2. Jobs-To-Be-Done / Problems J&P**

To predict and identify the heart disease patient. It is a very useful strategy that was used to control how the model can be utilized to increase the accuracy of the prediction of Heart Attack in each.

8

#### **3.4.3. Triggers TR**

The generation currently living now leads an extremely unhealthy lifestyle. People worry about the sharp rise in mortality from heat-related illnesses. They, therefore, desire to adopt a better lifestyle.

#### **3.4.4. Emotions: Before / After EM**

People frequently worry that their health will decline. They suffer unneeded tension and emotional breakdowns as a result of this. Our prediction system would enable them to keep track of their health independently and assist them in overcoming their erroneous concerns.

#### **3.4.5. Available Solutions**

EDA: Exploratory data analysis is the key step for getting meaningful results.

Pros: Improve understanding of variables by extracting averages, mean, minimum, and maximum values, etc. Discover errors, outliers, and missing values in the data. Identify patterns by visualizing data in graphs such as box plots, scatter plots, and histograms.

Cons: Exploratory research comes with disadvantages that include offering inconclusive results, lack of standardized analysis, a small sample population, and outdated information that can adversely affect the authenticity of the information.

## **4.REQUIREMENT ANALYSIS**

## 4.1 Functional Requirement

FR No.	Functional Requirement (Epic)	Sub Requirement (Story /Sub-Task)
FR-1	User Registration	Registration through Email
FR-2	User Confirmation	Confirmation via Email
FR-3	Visualizing Data	Visualize the trends on heart disease through Dashboard created using IBM Cognos Analytics
FR-4	Generation Report	Report can be viewedby the users

## 4.2 Non-Functional Requirement

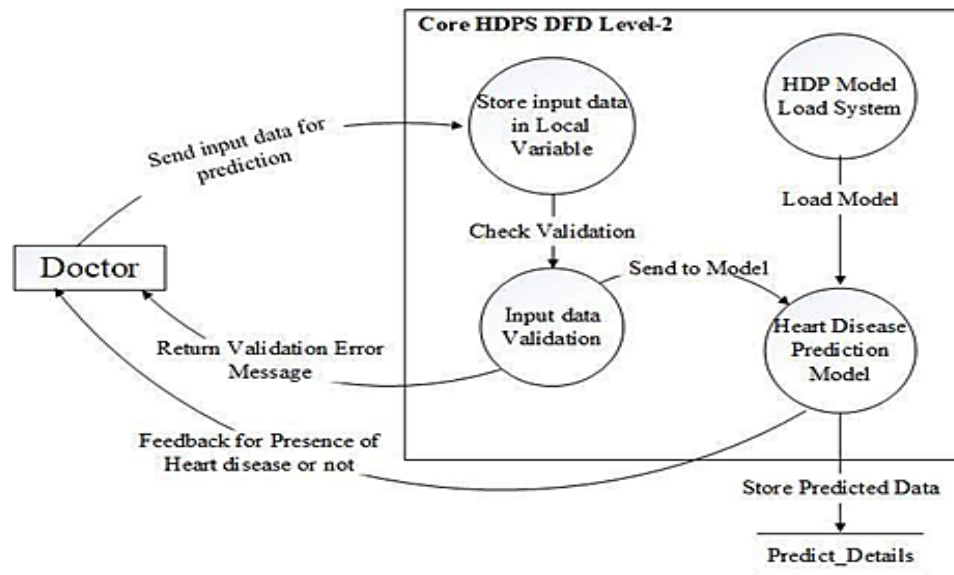
NFR-1	Usability	Provide a simple UI. Actions can be easily performed by a few clicks. Features will be understandable.
NFR-2	Security	2 step authorization (for register) Have a backup dataset
NFR-3	Reliability	Error must be low.(Improve accuracy) Must work without glitches
NFR-4	Performance	It is affected by the implementing algorithm. Depending on the error metrics we have to choose an algorithm with high response time.
NFR-5	Availability	Must be available for the user 24 x 7 without interruptions
NFR-6	Scalability	Should withstand a high number of users and large datasets.

## 5. PROJECT DESIGN

### 5.1 DATA FLOW DIAGRAM - 1



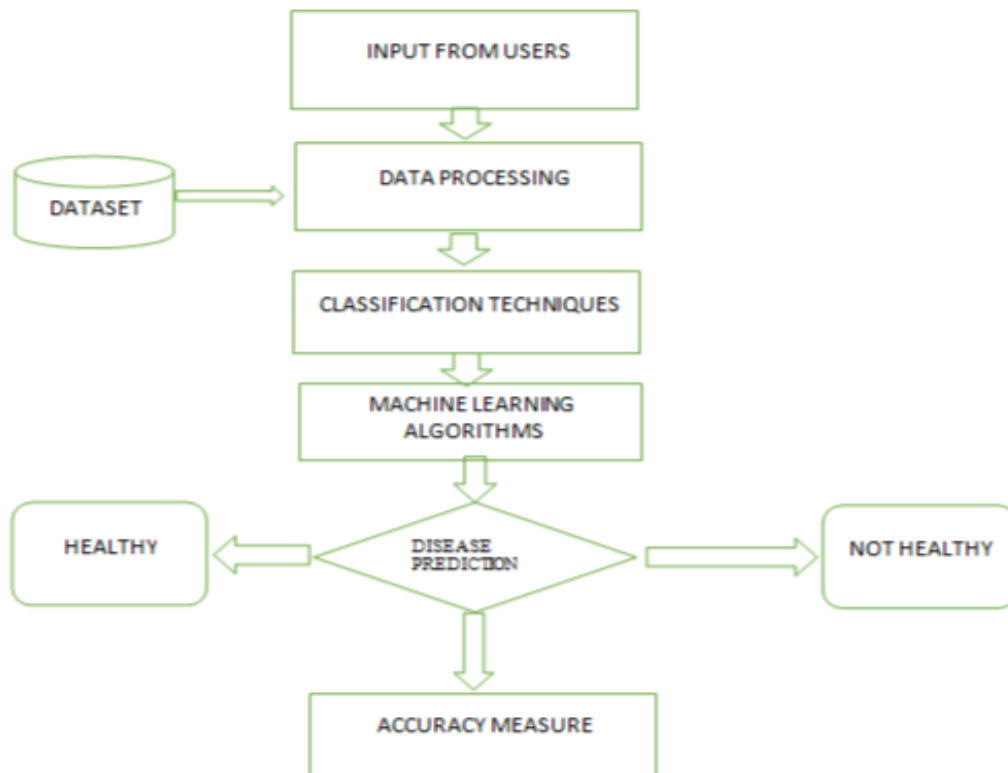
## 5.2 DATA FLOW DIAGRAM - 2



## 5.2 Solution Architecture:

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
- Define features, development phases, and solution requirements.
- Provide specifications according to which the solution is defined, managed, and delivered.



### 5.3 User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Web user)	System Requirement	USN-1	I. Hardware Requirement i. Laptop or PC <ul style="list-style-type: none"> <li>• I5 processor system or higher</li> <li>• 4GB RAM or higher</li> <li>• 128GB ROM or higher Android Phone (12.0 and above)</li> </ul>	These are all the specification available in your PC.	High	Sprint-2
		USN-2	II. Software Requirement ii. Laptop or PC	Install your application. This system can be used	Medium	Sprint-1



			<ul style="list-style-type: none"> <li>Windows 10 or higher</li> <li>Android studio</li> </ul>	to predict the presence of heart disease.		
	Registration	USN-3	As a user, I can register for the application by entering my email, password, and confirming my password	I can access my account/dashboard	High	Sprint-1
		USN-4	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
		USN-5	As a user, I can register for the application through email	I can register & access the dashboard with Gmail login	Medium	Sprint-1
	Login	USN-6	As a user, I can log into the application by entering email & password	I can register & access the dashboard with Gmail login	High	Sprint-1
	Dashboard	USN-7	As a user, I can view visualizations and a detailed report after entering my details.	I can view the correct results if the details entered are accurate..	High	Sprint-2
		USN-8	Profile-view & update your profile	I can see the profile.	High	Sprint-1
		USN-9	Change Password-user can change the password	I can able to change the password	High	Sprint-1
		USN-10	Home-Analyze your Heart	I can detect the health condition from where ever I want	High	Sprint-1
		USN-11	The user will have to fill in the below 13 fields for the system to predict a disease - Age in year - Gender - Chest pain Type - Fasting Blood Sugar - Resting Electrographic results(Restecg) - Exercise Induced Angina(Exang) - The slope of the peak exercise ST segment - CA - Number of major vessels colored by fluoroscopy - Thal - Trest Blood pressure - Serum Cholesterol - Maximum heart rate achieved(Thalach)- ST depression induced by exercise (Old peak)	These are the categories available in that application.	High	Sprint-2
		USN-12	View Doctors-view doctor detail by searching by names or filter by speciality	Using this application, people can know that the speciality doctor.	Medium	Sprint-1

## 6.PROJECT PLANNING & SCHEDULING :

### 6.1 Product Backlog, Sprint Schedule,and Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Datasets	USN-1	As an analyst, I will develop code for data preparation and data description.	5	High	Shenbagaruban Jegan Subashini karpagamaheswari
Sprint-2	Cleaning, exploring data and creating model	USN-2	As an Analyst I will develop code for data exploration.	5	High	Shenbagaruban Jegan Subashini karpagamaheswari
Sprint-3	Data visualization	USN-3	As an Analyst I can develop code for data visualization.	5	High	Shenbagaruban Jegan Subashini karpagamaheswari
Sprint-4	Data Prediction	USN-4	As a Data analyst, I will create code for different types of models in explored data	5	High	Shenbagaruban Jegan Subashini karpagamaheswari

### 6.2 Sprint Delivery Plan:

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	10	5 Days	24 Oct 2022	29 Oct 2022	10	29 Oct 2022
Sprint-2	10	5 Days	31 Oct 2022	05 Nov 2022	10	05 Nov 2022
Sprint-3	10	5 Days	07 Nov 2022	12 Nov 2022	10	12 Nov 2022
Sprint-4	10	5 Days	14 Nov 2022	19 Nov 2022	10	19 Nov 2022

#### Velocity:

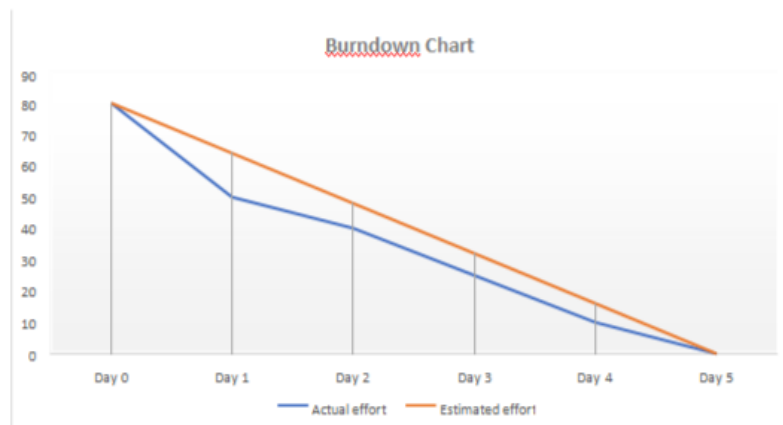
Imagine we have a 05-day sprint duration, and the velocity of the team is 10 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

$$AV = \text{Sprint Duration} / \text{Velocity} = 10 / 5 = 2$$

## Burndown Chart:

A burn down chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time.

**Goal: 60 hours in 5 days**



## 7. RESULTS

### 7.1. Performance Metrics

Metrics are measurements and parameters obtained throughout the quality assurance procedure. They may make reference to several test types. As you might have guessed, performance testing data gives you the ability to evaluate the efficiency of performance testing. Alternatively said, these measurements demonstrate how well software reacts to user scenarios and manages user flow in real time.

The following two categories of data are appropriate:

1. Measurements are data that are kept track of when testing, such as how long it takes to react to a request.
2. Metrics, which include various types of percentages, average indicators, and other metrics, are computations performed with the aid of certain formulas.

### Accuracy Percentages:

The number of wrong predictions on the test set as a whole divided by all of the test set predictions yields the error rate. Since accuracy and error rate are complementary quantities, we can always compute one from the other.

Accuracy = 1 – Error Rate

Error rate=1-Accuracy

#### Heart Disease Test

##### Heart Disease Test Form

Age	Sex		
<input type="text"/>	-- Select an Option --		
Chest Pain Type	Resting Blood Pressure in mm Hg	Serum Cholesterol in mg/dl	Fasting Blood Sugar > 120 mg/dl
-- Select an Option --	<input type="text"/>	<input type="text"/>	-- Select an Option --
Resting ECG Results	Maximum Heart Rate	Exercise Induced Angina	ST Depression Induced
-- Select an Option --	<input type="text"/>	-- Select an Option --	<input type="text"/>
Slope of the Peak Exercise ST Segment	Number of Vessels Colored by Flourosopy	Thalassemia	
-- Select an Option --	-- Select an Option --	-- Select an Option --	
<input type="button" value="Result"/>			

### Response Time:

The response time is not too long as our project as we have used real time data analysis. So, once the user enters his/her data in the Heart Disease prediction phase then the data will immediately be displayed so the response time is very less.

## 7.CODING AND SOLUTIONING

```
import numpy as np
import pickle
import sklearn
from flask import Flask, request, render_template
model = pickle.load(open('models.pkl', 'rb'))

app = Flask(__name__)
@app.route('/')
def home():
    return render_template('home.html')
@app.route('/signin')
def signin():
    return render_template('signin.html')

@app.route('/signup')
def signup():
    return render_template('signup.html')

@app.route('/predict', methods = ['POST'])
def predict():

    features = [float(i) for i in request.form.values()]
    #Convert features to array
    array_features = [np.array(features)]
    #Predict features
```

```
#prediction = model.predict(array_features)
#output = prediction
#if output == 1:
    #return render_template('Heart_Disease_Classifier.html', result = 'The
patient is not likely to have heart disease!')
#else:
    #return render_template('Heart_Disease_Classifier.html', result = 'The
patient is likely to have heart disease!')

#if __name__ == '__main__':
    #debug(True)
import numpy as np
import pickle
from flask import Flask, render_template, request, redirect, url_for, flash
import sqlite3
model = pickle.load(open('models.pkl', 'rb'))
app = Flask(__name__)
app.secret_key = "7847541"

def get_db():
    conn = sqlite3.connect('user_details.db')
    conn.row_factory = sqlite3.Row
    return conn

@app.route('/')
def index():
    return render_template('index.html', title='Home')

@app.route('/about')
def about():
    return render_template('about.html', title='About')

@app.route('/signin', methods=('GET', 'POST'))
def signin():
```

```

error = None
if request.method == 'POST':
    name = request.form['name']
    password = request.form['password']
    db = get_db()
    user = db.execute(
        'SELECT name FROM user_details WHERE password = ?', (password,
    )
    ).fetchone()

    if user is None:
        error = 'Incorrect Username/Password.'

    if error is None:
        return render_template('index.html', title="Home", succ="login
successfull!")
    flash(error)
    db.close()

    return render_template('signin.html', title='Sign In', error=error)

@app.route('/signup', methods=('POST', 'GET'))
def signup():
    if request.method == 'POST':
        name = request.form['name']
        email = request.form['email']
        password = request.form['password']
        db = get_db()
        curr = db.cursor()
        curr.execute(
            'INSERT INTO user_details (name, email, password) VALUES (?,
?, ? );', (name, email, password )
        )

```

```
        db.commit()
        curr.close()
        db.close()

        return render_template('index.html', title="Home",
succ="Registration Successfull!")
        return render_template('signup.html', title='Sign Up')

@app.route('/Heart_Disease_Classifier')
def Heart_Disease_Classifier():
    return render_template('Heart_Disease_Classifier.html')

@app.route('/predict', methods =['POST'])
def predict():

    features = [float(i) for i in request.form.values()]
    #Convert features to array
    array_features = [np.array(features)]
    #Predict features
    prediction = model.predict(array_features)
    output = prediction
    if output == 1:
        return render_template('Heart_Disease_Classifier.html', result = 'The
patient is not likely to have heart disease!')
    else:
        return render_template('Heart_Disease_Classifier.html', result = 'The
patient is likely to have heart disease!')

if __name__ == '__main__':
    app.run(debug=True)
```



## **8. ADVANTAGES & DISADVANTAGES**

### **8.1. Advantages**

- a. The proposed work predicts the chances of Heart Disease and classifies patient's risk level
- b. It is implementing different data mining techniques such as Naive Bayes, Decision Tree, Logistic Regression and Random Forest.
- c. User friendly

### **18.2. Disadvantages**

- a. Data analytics techniques do not help to. provide effective decision making.
- b. Cannot handle enormous datasets
- c. Prediction of cardiovascular disease results is not accurate

## **9.TESTING**

### **1. Purpose of Document**

The purpose of this document is to briefly explain the test coverage and open issues of the Visualizing and Predicting Heart Disease with an Interactive Dash Board project at the time of the release to User Acceptance Testing (UAT).

### **2. Defect Analysis**

This report shows the number of resolved or closed bugs at each severity level, and how they were resolved.

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	5	0	2	2	9
Duplicate	0	0	1	0	1
External	2	1	0	0	3
Fixed	5	0	0	13	18
Not Reproduced	0	0	1	0	1
Skipped	0	1	1	1	3
Won't Fix	0	0	0	0	0
Totals	12	2	5	16	35

### 3. Test Case Analysis

This report shows the number of test cases that have passed, failed, and untested.

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	7	0	0	7
Client Application	51	0	0	51
Security	2	0	0	2
Outsource Shipping	3	0	0	3
Exception Reporting	9	0	0	9
Final Report Output	4	0	0	4
Version Control	2	0	0	2

## 10. CONCLUSION

The long-term preservation of human life and the early identification of irregularities in heart problems will benefit from the identification of the processing of raw healthcare data related to the heart. In this study, raw data was processed using machine learning techniques to produce a brand-new understanding of cardiac disease. In the medical field, heart disease prediction is difficult and crucial. The death rate, however, can be significantly reduced if the disease is identified in its early stages and preventative measures are put in place as soon as feasible. To move the investigations from simply theoretical frameworks and simulations to actual datasets, further elaboration of this study is extremely desirable. The model's ability to be employed to increase the precision of heart attack prediction in any individual was regulated using a very helpful technique.

When compared to the previously employed classifiers, such as naive bayes, etc., the proposed model's strength was quite satisfying. It was able to predict signs of having a heart illness in a specific individual by applying KNN and Logistic Regression, which demonstrated good accuracy. Therefore, by utilizing the provided model to determine the likelihood that the classifier will correctly and reliably detect the heart illness, a large amount of pressure has been reduced. The Given heart disease prediction system improves and lowers the cost of medical care. This project gives us significant knowledge that can help us predict the patients with heart disease It is implemented on the .pynb format

## **11. FUTURE SCOPE**

This study discusses the issue of constricting and summarizing various data mining strategies utilized in the field of medical forecasting. For intelligent and successful heart attack prediction via data mining, the emphasis is on combining various methods and combinations of numerous target attributes. Significantly, 15 attributes are specified for predicting heart attacks, and using simple data mining techniques, other approaches, including ANN, time series, clustering and association rules, soft computing approaches, etc., can also be included. The results of predictive data mining on the same dataset show that Decision Tree outperforms and, occasionally, Bayesian classification has accuracy levels comparable to those of decision tree, but other predictive methods, such as KNN, Neural Networks, and Classification based on clustering, are not performing well. The second finding is that using a genetic algorithm to lower the actual data quantity and obtain the ideal subset of attributes suitable for heart disease prediction increases the decision tree and Bayesian classification's accuracy. For the automation of heart disease prediction, the proposed work can be expanded and improved. Real data from healthcare institutions and agencies must be gathered, and all methods must be compared for the highest level of accuracy.

## **12.APPENDIX**

GitHub & Project Demo Link:

<https://github.com/IBM-EPBL/IBM-Project-37235-1660302156>