

# **Visualizing and Predicting Heart Diseases with an Interactive Dashboard**

**NALAIYA THIRAN PROJECT REPORT  
2022**

*Submitted by*

<b>Joel Moses S</b>	<b>312419205040</b>
<b>Danushraj M</b>	<b>312419205023</b>
<b>Michael Roshan J</b>	<b>312419205058</b>
<b>Harish V</b>	<b>312419205031</b>

**Team ID: PNT2022TMID28392**

# Content

## **1. INTRODUCTION**

- 1.1 Project Overview
- 1.2 Purpose

## **2. LITERATURE SURVEY**

- 2.1 Existing problem
- 2.2 References
- 2.3 Problem Statement Definition

## **3. IDEATION & PROPOSED SOLUTION**

- 3.1 Empathy Map Canvas
- 3.2 Ideation & Brainstorming
- 3.3 Proposed Solution
- 3.4 Problem Solution fit

## **4. REQUIREMENT ANALYSIS**

- 4.1 Functional requirement
- 4.2 Non-Functional requirements

## **5. PROJECT DESIGN**

- 5.1 Data Flow Diagrams
- 5.2 Solution & Technical Architecture

## **6. PROJECT PLANNING & SCHEDULING**

- 6.1 Sprint Planning & Estimation
- 6.2 Sprint Delivery Schedule

## **7. CODING & SOLUTIONING (Explain the features added in the project along with code)**

- 7.1 Feature 1
- 7.2 Feature 2

## **8. TESTING**

- 8.1 Test Cases
- 8.2 User Acceptance Testing

## **9. RESULTS**

- 9.1 Performance Metrics

## **10. ADVANTAGES & DISADVANTAGES**

## **11. CONCLUSION**

## **12. FUTURE SCOPE**

## **13. APPENDIX**

Source Code

GitHub & Project Demo Link

**TEAM ID: PNT2022TMID28392**

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

### **1. Introduction**

#### **1.1 Project Overview**

Today, heart failure diseases affect more people worldwide than other autoimmune conditions. Cardiovascular Diseases (CVDs) affect the heart and obstruct blood flow through the blood vessels. Better technologies like MCGs help in detecting these diseases when they are in an early stage. This project of ours aims to create an interactive Dashboard and dataset to predict which patients are most likely to suffer from a heart disease in the near future using the variables given.

#### **1.2 Purpose**

Several risk factors for manual heart disease prediction may include inactivity in a physical form, unhealthy eating habits, or even the consumption of alcohol. With such well-defined parameters and the rise of data science, a data-driven approach can surely help in heart disease prediction using machine learning technologies. Early identification of heart disease of improved diagnosis and high-risk individuals using a prediction model can be recommended for a fatality rate reduction, and decision-making is improved for further treatment and prevention.

### **2. Literature Survey**

#### **2.1 Existing Problem**

Healthcare industries generate enormous amount of data, so called big data that accommodates hidden knowledge or pattern for decision making. The huge volume of data is used to make decision which is more accurate than intuition. Exploratory Data Analysis (EDA) detects mistakes, finds appropriate data, checks assumptions and determines the correlation among the explanatory variables. In the context, EDA is considered as analysing data that excludes inferences and statistical modelling. Analytics is an essential technique for any profession as it forecast the future and hidden pattern. Data analytics is considered as a cost effective technology in the recent past and it plays an essential role in healthcare which includes new research findings, emergency situations and outbreaks of disease. The use of analytics in healthcare improves care by facilitating preventive care and EDA is a vital step while analysing data..

## 2.2 References

### **“Heart Disease Prediction Using Supervised Machine Learning Algorithms”**

**Narendra Mohan, Vinod Jain, Gauranshi Agrawal**

Predicting and detecting cardiac disease has always been a difficult and time-consuming undertaking for doctors. To treat cardiac disorders, hospitals and other clinics are giving costly therapies and operations. As a result, anticipating cardiac disease in its early stages will be beneficial to people all around the world, allowing them to take required treatment before it becomes serious. Heart disease has been a major issue in recent years, with the primary causes being excessive alcohol use, tobacco use, and a lack of physical activity. Machine learning methods are utilized to forecast cardiac illnesses in this article. For training and testing, a data collection containing diverse human health parameters is used. Many AI&ML algorithms are used to predict cardiac disorders. The performance of the machine learning algorithm is compared after it has been implemented.

### **Prediction of heart disease at early stage using data mining and big data analytics: A survey**

**N. K. Salma Banu, Suma Swamy**

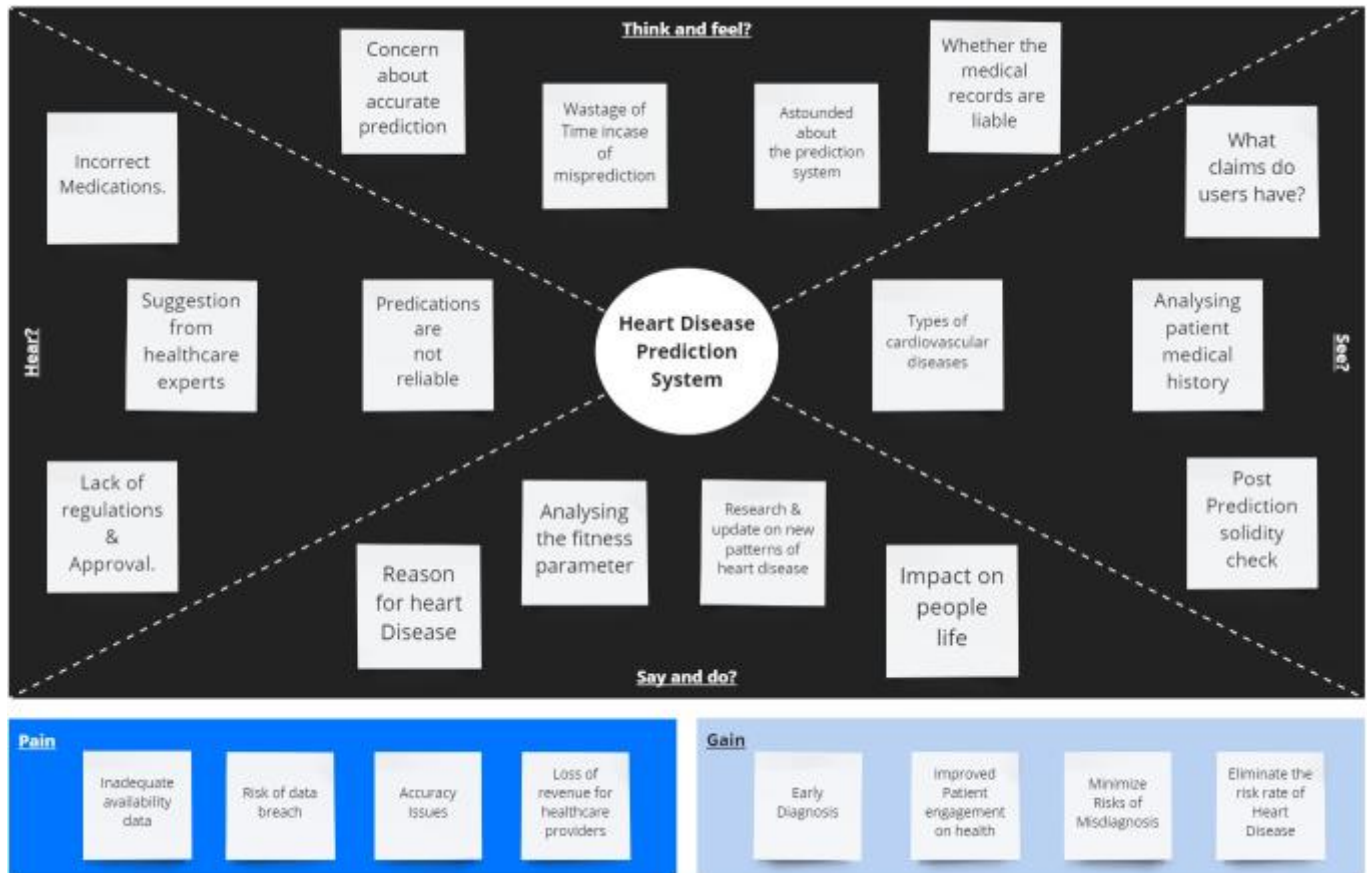
Predicting and detection of heart disease has always been a critical and challenging task for healthcare practitioners. Hospitals and other clinics are offering expensive therapies and operations to treat heart diseases. So, predicting heart disease at the early stages will be useful to the people around the world so that they will take necessary actions before getting severe. Heart disease is a significant problem in recent times; the main reason for this disease is the intake of alcohol, tobacco, and lack of physical exercise. Over the years, machine learning shows effective results in making decisions and predictions from the broad set of data produced by the health care industry.

## 2.3 Problem Statement Definition

The issue occurs for people with unhealthy lifestyles and age above 40. Where is the issue occurring. The issue is originating from an unhealthy lifestyle. It mostly occurs in the blood valves of the heart. If we don't solve the problem, many people will die at a young age. The death rate due to heart disease will increase rapidly. We should predict the problem before giving treatment to the patients. As the problem is predicted early, we can solve it easily and early.


### 3. Ideation and Proposed Solution

#### 3.1 Empathy Map



## 3.2 Ideation and Brainstorming

### Step-1: Team Gathering, Collaboration and Select the Problem Statement



## Brainstorm & idea prioritization

🕒 10 min  
👤 (Dariusz H., Mark V., Joel Rosen S., Michael Rothen J.)

📄 Share template feedback

**1**  
Define your problem statement

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.

**2**  
Brainstorm

Write down any ideas that come to mind that address your problem statement.

🕒 10 minutes

Team member 1	Team member 2	Team member 3	Team member 4
<div>Problem statement: Reduce the risk of heart disease</div> <div>Brainstorm ideas: 1. Increase physical activity 2. Improve diet 3. Reduce stress 4. Quit smoking 5. Regular medical check-ups</div> <div>Feedback: 1. Increase physical activity 2. Improve diet 3. Reduce stress 4. Quit smoking 5. Regular medical check-ups</div>	<div>Problem statement: Reduce the risk of heart disease</div> <div>Brainstorm ideas: 1. Increase physical activity 2. Improve diet 3. Reduce stress 4. Quit smoking 5. Regular medical check-ups</div> <div>Feedback: 1. Increase physical activity 2. Improve diet 3. Reduce stress 4. Quit smoking 5. Regular medical check-ups</div>	<div>Problem statement: Reduce the risk of heart disease</div> <div>Brainstorm ideas: 1. Increase physical activity 2. Improve diet 3. Reduce stress 4. Quit smoking 5. Regular medical check-ups</div> <div>Feedback: 1. Increase physical activity 2. Improve diet 3. Reduce stress 4. Quit smoking 5. Regular medical check-ups</div>	<div>Problem statement: Reduce the risk of heart disease</div> <div>Brainstorm ideas: 1. Increase physical activity 2. Improve diet 3. Reduce stress 4. Quit smoking 5. Regular medical check-ups</div> <div>Feedback: 1. Increase physical activity 2. Improve diet 3. Reduce stress 4. Quit smoking 5. Regular medical check-ups</div>

## Step-2: Brainstorm, Idea Listing and Grouping

3

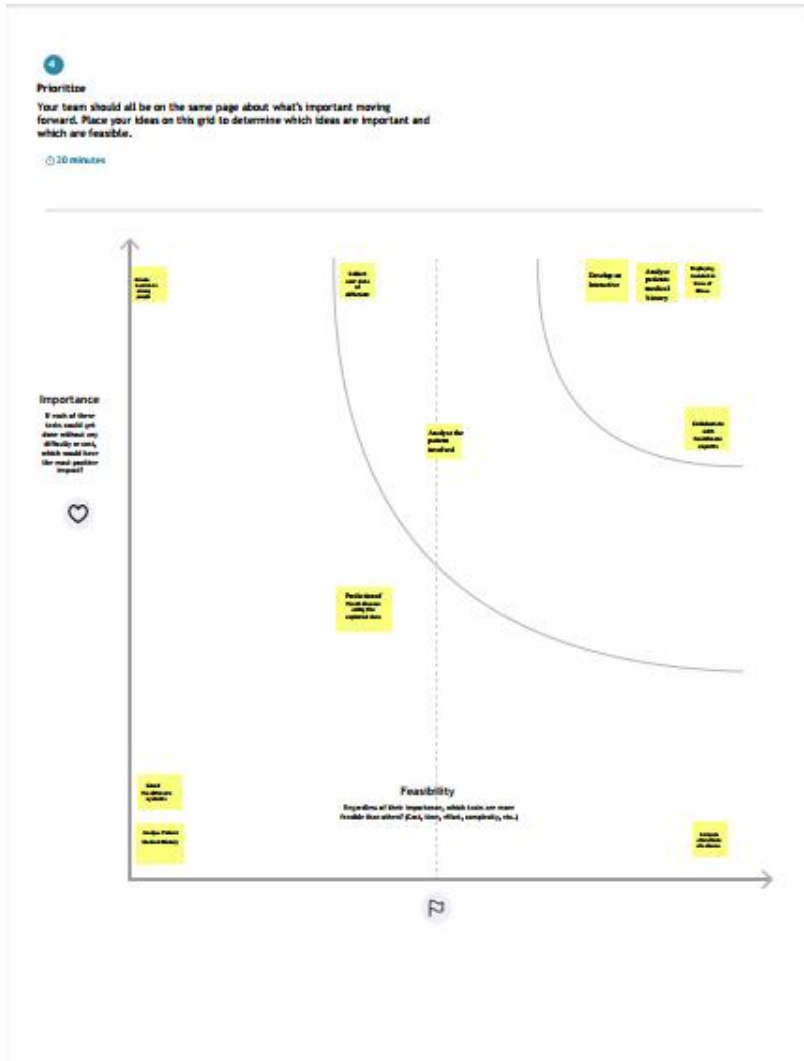
### Group Ideas

Take turns sharing your ideas while clustering similar or related notes as you go. In the last 10 minutes, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you can break it up into smaller sub-groups.

20 minutes

The image shows a digital workspace with four clusters of sticky notes. Each cluster is enclosed in a black rectangular border. The sticky notes are yellow with black text. The clusters contain various ideas related to a project, such as 'Project: Build a new website', 'Project: Create a new app', 'Project: Develop a new feature', 'Project: Improve the user interface', 'Project: Optimize the website', 'Project: Enhance the mobile app', 'Project: Implement a new API', 'Project: Integrate with a third-party service', 'Project: Add a new payment method', 'Project: Implement a new security feature', 'Project: Improve the loading time', 'Project: Optimize the database', 'Project: Enhance the search functionality', 'Project: Implement a new analytics tool', 'Project: Integrate with a new marketing channel', 'Project: Add a new social media link', 'Project: Implement a new email marketing campaign', 'Project: Improve the customer support process', 'Project: Optimize the checkout process', 'Project: Enhance the onboarding process', 'Project: Implement a new feedback mechanism', 'Project: Integrate with a new analytics tool', 'Project: Add a new social media link', 'Project: Implement a new email marketing campaign', 'Project: Improve the customer support process', 'Project: Optimize the checkout process', 'Project: Enhance the onboarding process', 'Project: Implement a new feedback mechanism'.

### Step-3: Idea Prioritization



### 3.3 Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Heart disease is even being highlighted as a silent killer which leads to the death of a person without obvious symptoms. The nature of the disease is the cause of growing anxiety about the disease & its consequences. Hence continued efforts are being done to predict the possibility of this deadly disease in prior. So that various tools & techniques are regularly being experimented with to suit the present-day health needs.
2.	Idea / Solution description	The working of the system starts with the collection of data and selecting the important attributes. Then the required data is pre-processed into the required format. The data is then divided



		into two parts training and testing data. The algorithms are applied and the model is trained using the training data. The accuracy of the system is obtained by testing the system.
3.	Novelty / Uniqueness	The primary purpose of this study is to give clinicians a tool to detect cardiac problems at an early stage.
4.	Social Impact / Customer Satisfaction	This Heart Disease detection system assists a patient based on his/her clinical information of them been diagnosed with a previous heart disease by using these computer aided techniques we can predict the patient fast and better and the cost can be reduced very much.
5.	Business Model (Revenue Model)	Can be deployed by Hospitals or other Health organizations, so that the success rate of prediction is higher.
6.	Scalability of the Solution	It is the property of a system to handle a growing amount of work by adding resources to the system. In order to stay at top positions in market the scalability and adaptability towards changing needs must be met.

### 3.3 Problem Solution Fit

Define CS, fit into CL	<b>1. CUSTOMER SEGMENT(S)</b> <span>CS</span> Customers who use this application should be 18 years old or greater.	<b>6. CUSTOMER LIMITATIONS</b> <span>CL</span> <small>EG. BUDGET, DEVICES</small> <ul style="list-style-type: none"> <li>Worried about safety of the customer data.</li> <li>Wastage of time for customer if misprediction happens.</li> <li>Difficult to use for first time users.</li> </ul>	<b>5. AVAILABLE SOLUTIONS</b> <span>AS</span> <small>PLUSES &amp; MINUSES</small> This application is an healthy way to keep in check the medical condition of the customers. This is a precaution application developed to predict the heart disease at early stages.	Explore AS, differentiate
	<b>2. PROBLEMS / PAINS</b> <span>PR</span> <small>+ ITS FREQUENCY</small> <ul style="list-style-type: none"> <li>It is very difficult to predict the heart disease.</li> <li>Predictions are not reliable.</li> <li>It has higher percentage which leads to misprediction.</li> <li>Lack of regulation and approval.</li> <li>Risk of data breach.</li> <li>Lack of trust leads to reduce in customers using the application.</li> <li>These platforms mostly are paid.</li> </ul>	<b>9. PROBLEM ROOT / CAUSE</b> <span>RC</span> <ul style="list-style-type: none"> <li>Lack of trust or reliability among the users.</li> <li>Presence of many non-approved application leads to downfall of all the health application.</li> <li>Lack of awareness about the seriousness of heart disease.</li> <li>Health applications are not reliable among the users.</li> <li>The health apps work in a profit-making way.</li> </ul>	<b>7. BEHAVIOR</b> <span>BE</span> <small>+ ITS INTENSITY</small> <ul style="list-style-type: none"> <li>Lack of awareness about heart diseases among the customers.</li> <li>The customers are old fashioned and are not convinced by the heart disease prediction.</li> <li>Customers feel insecure about sharing their medical details to the application.</li> <li>Customers are worried about misprediction which leads to wastage of time and effort.</li> </ul>	
Identify strong TR & EM	<b>3. TRIGGERS TO ACT</b> <span>TR</span> <ul style="list-style-type: none"> <li>Seeing social awareness programs about the severity of heart diseases</li> <li>Efficient prediction of results among the patients</li> <li>Ethical functioning of the system</li> </ul>	<b>10. YOUR SOLUTION</b> <span>SL</span> <ul style="list-style-type: none"> <li>This application helps in predicting the heart disease accurately and is reliable.</li> <li>This application maintains the customer medical details safe and secure.</li> <li>This application classifies all the heart disease and alerts the customer based on their prediction.</li> <li>This application is approved by the health institution.</li> <li>The application is free of cost.</li> </ul>	<b>8. CHANNELS of BEHAVIOR</b> <span>CH</span> <b>ONLINE</b> Extract information from the database and update it in real-time.  <b>OFFLINE</b> Extract information from the database and use it for Diagnostic purposes.	Extract online & offline CH of BE
	<b>4. EMOTIONS</b> <span>EM</span> <small>BEFORE / AFTER</small> <b>Before:</b> Apprehension of the disease and Uneasiness about the symptoms. <b>After:</b> Reassurance and Appeasement of diagnosis.			

## 4.Requirement Analysis

### 4.1 Functional Requirement

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Enables user to make registration for the application through Google Account
FR-2	User Confirmation	Once after registration, the user will get confirmation via email.
FR-3	Visualizing Data	User can visualize the trends on the heart disease through Dashboard and see their deliverables.
FR-4	Generation Report	User can view his/her health report and can get treated accordingly

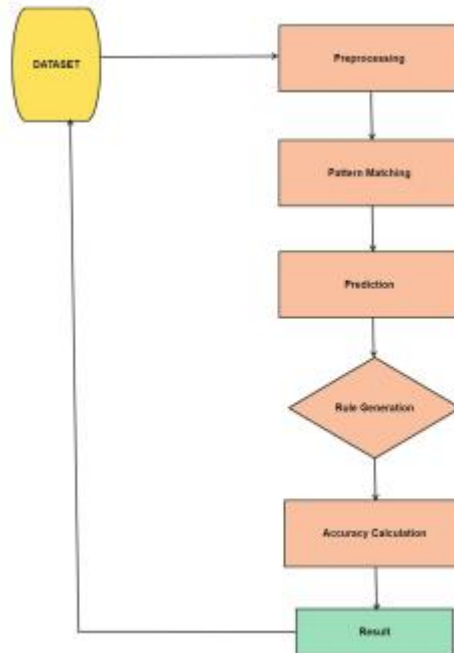
### 4.2 Non-Functional Requirement

Following are the non-functional requirements of the proposed solution.

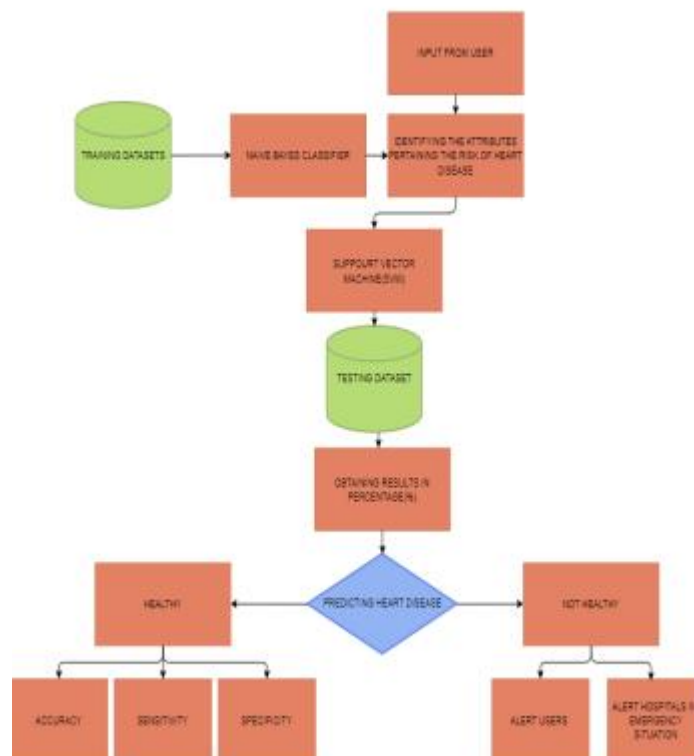
FR No.	Non-Functional Requirement	Description
NFR-1	<b>Usability</b>	The application will have a simple and user-friendly graphical interface. Users will be able to understand and use all the features of the application easily. Any action has to be performed with just a few clicks
NFR-2	<b>Security</b>	to overcome that the data we collected and processed are encrypted with standards that cannot be broken. So, this makes our system to ensure privacy rights to the users.
NFR-3	<b>Reliability</b>	The application has to be consistent at every scenario and has to work without failure in any Environment
NFR-4	<b>Performance</b>	Making the data available in times need is important. We have enclosed a feature of sharing the patients record to authorized facilities in case of emergencies so that right medications can be given on time.
NFR-5	<b>Availability</b>	The application has to be available 24 x 7 for users without any interruption
NFR-6	<b>Scalability</b>	The application can withstand the increase in the no. of users and has to be able to develop Higher versions

## 5 Project Design

### 5.1 Data Flow Diagram



### 5.2 Solution and Technical Architecture



## 6. Project Planning and Scheduling

### 6.1 Script Planning and Execution

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password	7	High	1
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	3	High	3
Sprint-1		USN-3	As a user, I can register for the application through Facebook	3	Medium	1
Sprint-1		USN-4	As a user, I can register for the application through Gmail	3		1
Sprint-2	Login	USN-5	As a user, I can log into the application by entering email & password	20	High	4
Sprint-3	Dashboard	USN-6	As a user, I can see current health status, personalize the interface etc.	10	High	3
Sprint-3		USN-7	As a user, I can update my user profile, medical details and have access to all the medical records.	5	Medium	1
Sprint-3		USN-8	As a user, I can add or customize all the health fit tracking devices	5	Medium	1
Sprint-4	Notification Manager	USN-9	As a user, I can get to know about my health condition at regular intervals via monitoring system	4	High	3
Sprint-4	Appointments	USN-10	As a user, I can book appointments with nearby health care specialist for consultation along with the current medical status report of the user	3	Medium	1
Sprint-4	SOS	USN-11	R In emergency scenario, the automated system must intimate the nearby emergency facility for assistance and also notify their guardian	10	High	4
Sprint-4	Community support service	USN-12	As a patient, I can connect with people of the community to share thoughts and embrace self-confidence to fight against the disease	3	Medium	3

6.2 Sprint Delivery Schedule

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	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

## 7 Coding And Solution

### 7.1 Machine Learning

Learning which model is best for the given Dataset

```
In [65]: final_data = pd.DataFrame({'Models':['LR','SVM','KNN','DT','RF','GB'],  
                                   'ACC':[accuracy_score(y_test,y_pred1)*100,  
                                           accuracy_score(y_test,y_pred2)*100,  
                                           accuracy_score(y_test,y_pred3)*100,  
                                           accuracy_score(y_test,y_pred4)*100,  
                                           accuracy_score(y_test,y_pred5)*100,  
                                           accuracy_score(y_test,y_pred6)*100]})
```

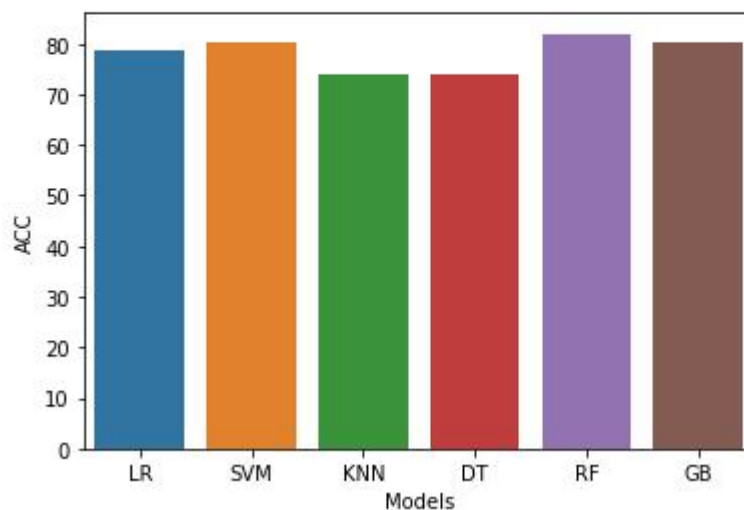
```
In [66]: final_data
```

```
Out[66]:
```

	Models	ACC
0	LR	78.688525
1	SVM	80.327869
2	KNN	73.770492
3	DT	73.770492
4	RF	81.967213
5	GB	80.327869

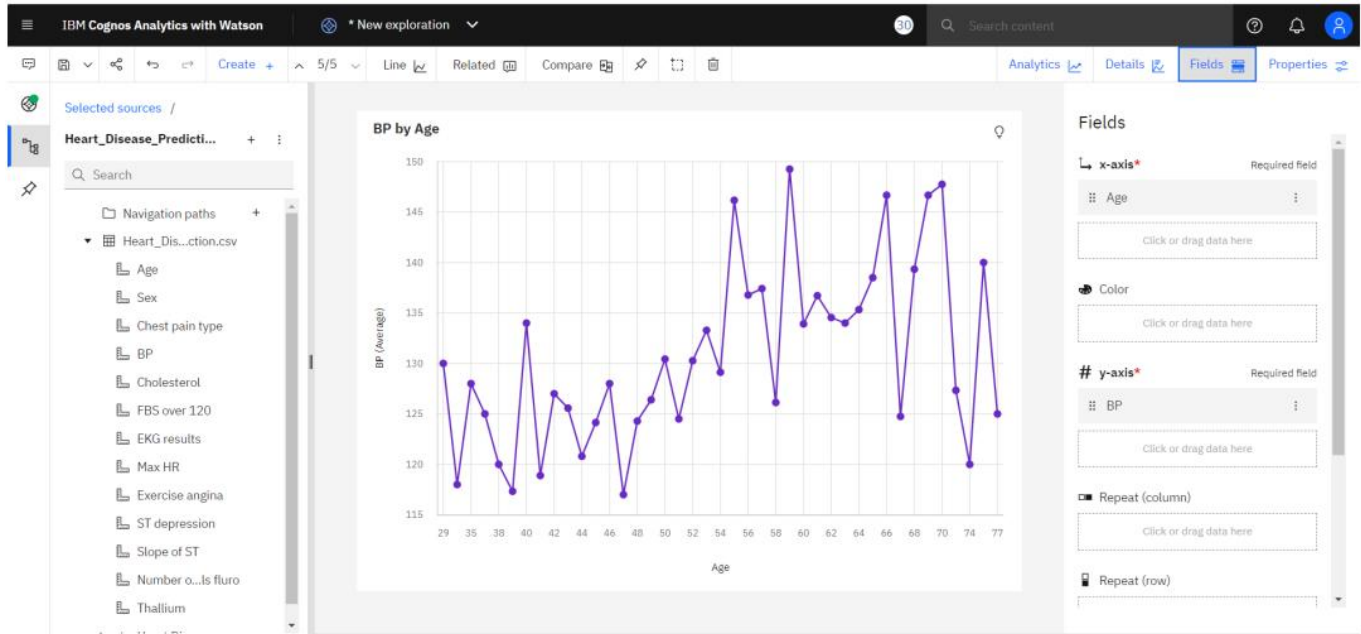
```
In [68]: sns.barplot(final_data['Models'],final_data['ACC'])
```

```
Out[68]:
```

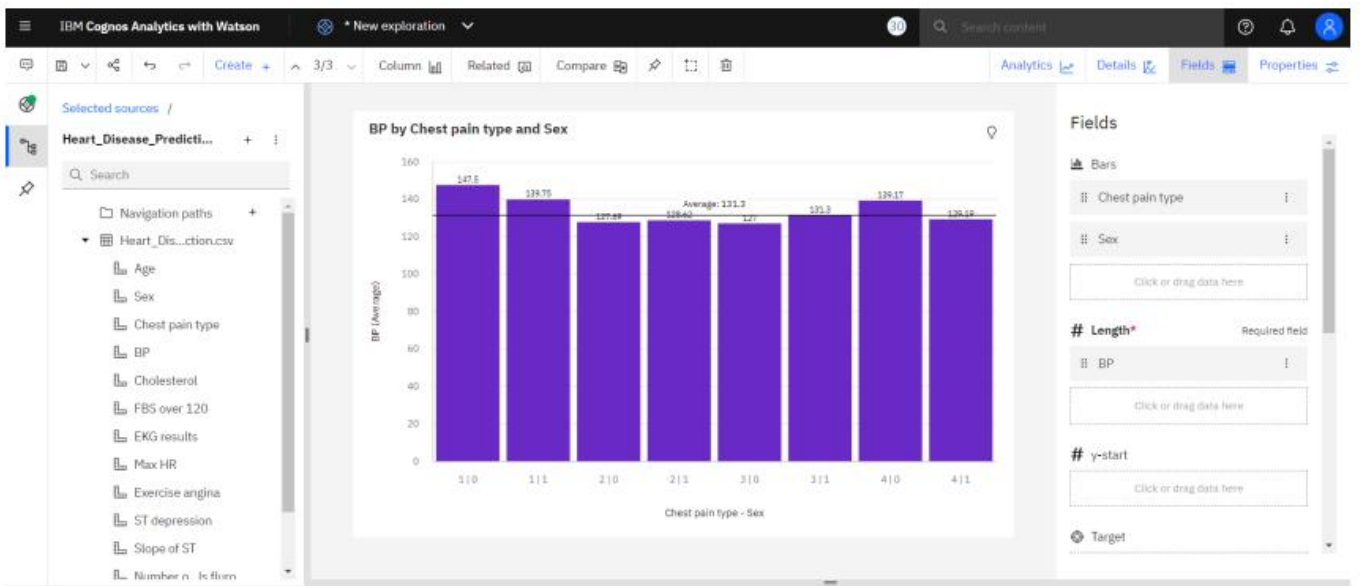


## 7.2 Dashboard

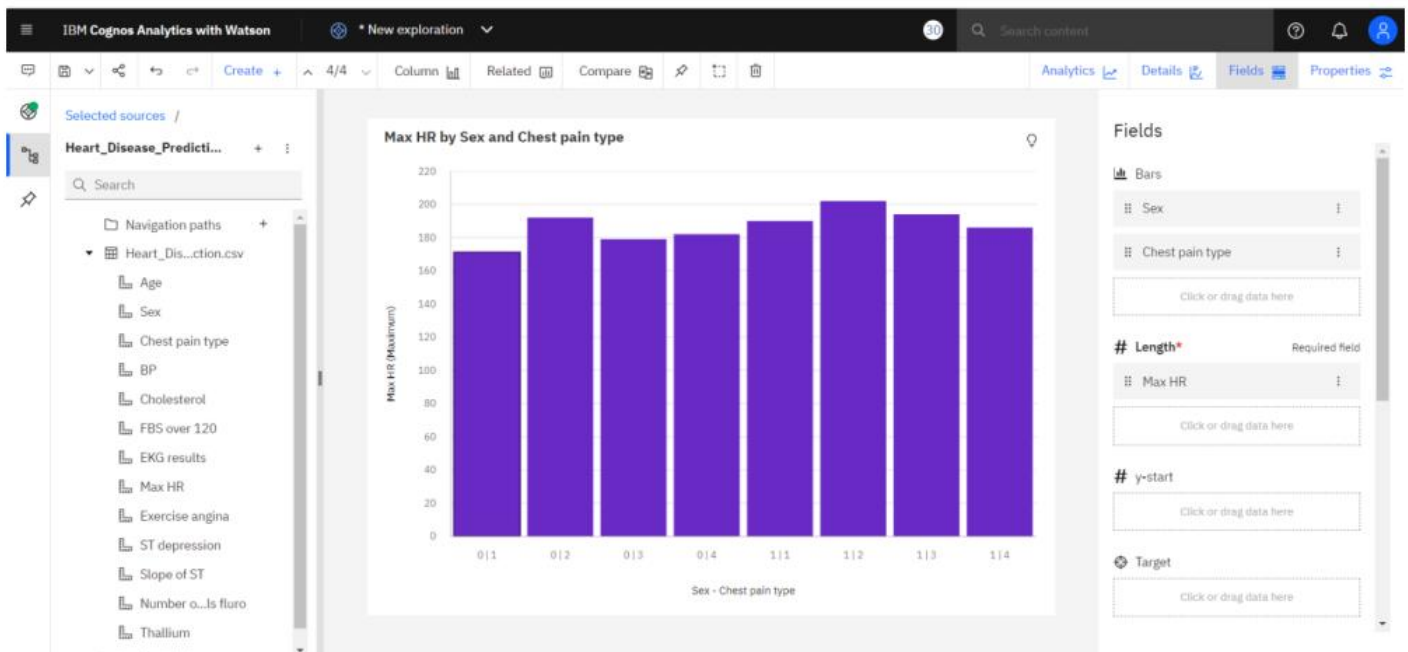
### Average BP during chest pain



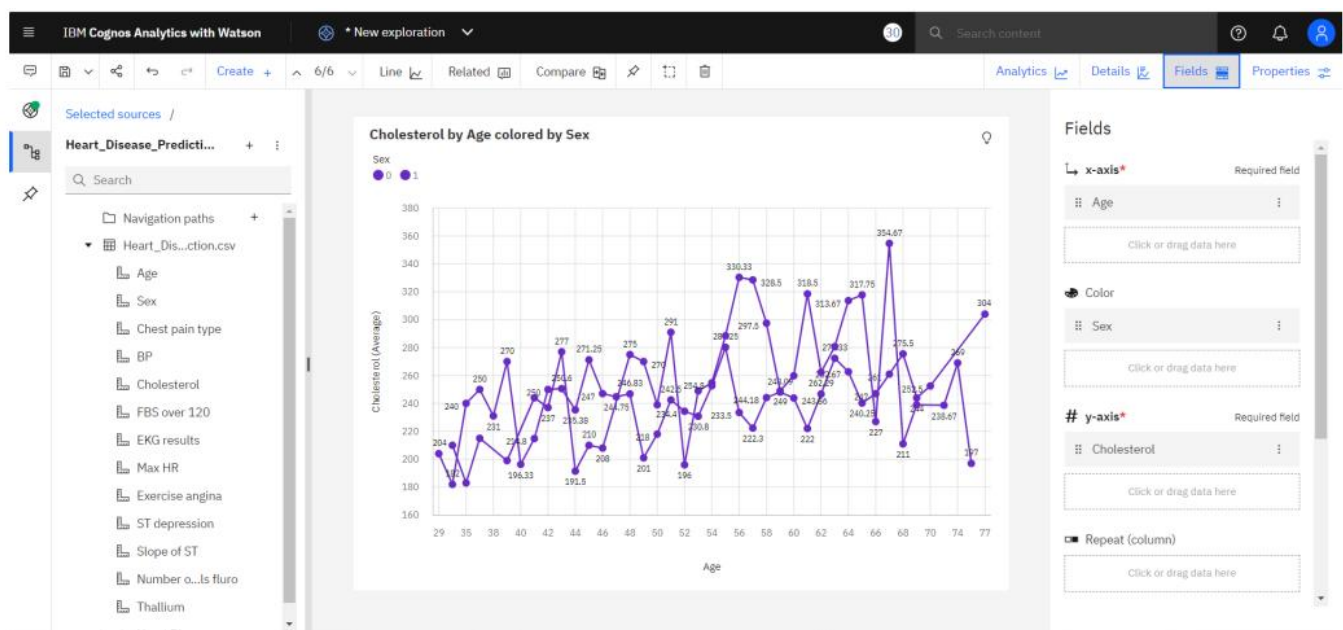
### Exploration Of BP vs Chestpain Type And Gender:



## Exploration Of Max Heart Rate During The Chest Pain:



## Exploration Of Cholesterol by age and Gender:






## 8. Testing

### 8.1 Test Cases

Heart Disease Prediction System !



**Predict**

\*Note: Please verify the hint and enter details. Mostly Integer type Supported

**Age**  
Enter your age

**Gender**  
0-male & 1-female

**Constrictive pericarditis**  
Enter value 0-3


**Resting blood pressure**  
Enter values (eg:120-190)

**Cholesterol**  
Enter value 125 to 200mg/dL

**Fasting Blood Sugar**  
Eg:140 mg/dL

**Resting electrocardiography**

Heart Disease Prediction System !



**Predict**

\*Note: Please verify the hint and enter details. Mostly Integer type Supported

**Maximum heart rate**  
Eg:155 beats/min

**Exercise induced angina**  
(1 = yes; 0 = no)

**Asymptomatic chest pain**  
0-normal,1-fixed defect,2-reversible defect

**ST/heart rate slope**  
Eg:Range (0-2)

**Coronary calcium**  
Eg: 155

**Thalassemia**  
1 or 2 or 3

## 82 User acceptance Testing

Testing a case where user has heart disease

### Heart Disease Prediction System !



You Have The Risk Of Heart Disease! 😟

**Predict**

\*Note: Please verify the hint and enter details. Mostly Integer type Supported

0

Maximum heart rate

115

Exercise induced angina

0

Asymptomatic chest pain

1

ST/heart rate slope

1

Coronary calcium

0

Thalassemia

2

Testing a case where user does not have heart disease

### Heart Disease Prediction System !



You Are Having A Healthy Body! 🥳

**Predict**

\*Note: Please verify the hint and enter details. Mostly Integer type Supported

1

Maximum heart rate

161

Exercise induced angina

1

Asymptomatic chest pain

0

ST/heart rate slope

2

Coronary calcium

1

Thalassemia

3

Default test values for checking the prediction algorithm.

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
723	68	0	2	120	211	0	0	115	0	1.5	1	0	2	1
733	44	0	2	108	141	0	1	175	0	0.6	1	0	2	1
739	52	1	0	128	255	0	1	161	1	0.0	2	1	3	0
843	59	1	3	160	273	0	0	125	0	0.0	2	0	2	0
878	54	1	0	120	188	0	1	113	0	1.4	1	1	3	0

## 9. Result

### 9.1 Performance Metrics

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
2	52	1	0	125	212	0	1	168	0	1	2	2	3	0
3	53	1	0	140	203	1	0	155	1	3.1	0	0	3	0
4	70	1	0	145	174	0	1	125	1	2.6	0	0	3	0
5	61	1	0	148	203	0	1	161	0	0	2	1	3	0
6	62	0	0	138	294	1	1	106	0	1.9	1	3	2	0
7	58	0	0	100	248	0	0	122	0	1	1	0	2	1
8	58	1	0	114	318	0	2	140	0	4.4	0	3	1	0
9	55	1	0	160	289	0	0	145	1	0.8	1	1	3	0
10	46	1	0	120	249	0	0	144	0	0.8	2	0	3	0
11	54	1	0	122	286	0	0	116	1	3.2	1	2	2	0
12	71	0	0	112	149	0	1	125	0	1.6	1	0	2	1
13	43	0	0	132	341	1	0	136	1	3	1	0	3	0
14	34	0	1	118	210	0	1	192	0	0.7	2	0	2	1
15	51	1	0	140	298	0	1	122	1	4.2	1	3	3	0
16	52	1	0	128	204	1	1	156	1	1	1	0	0	0
17	34	0	1	118	210	0	1	192	0	0.7	2	0	2	1
18	51	0	2	140	308	0	0	142	0	1.5	2	1	2	1
19	54	1	0	124	266	0	0	109	1	2.2	1	1	3	0
20	50	0	1	120	244	0	1	162	0	1.1	2	0	2	1
21	58	1	2	140	211	1	0	165	0	0	2	0	2	1

## **10. Advantages Disadvantages**

### **Advantages:**

- One of the fastest and easiest way to determine if a person is likely to suffer from a heart disease or not.
- Useful for all the medical officers to know about the patient's history
- User Friendly
- Easy to understand
- Secure

### **Disadvantages:**

- Need a more dynamic User interface
- Users need to know all the fields
- Does Not take null value as input
- Does not provide suggestions to the user.

## **11. Conclusion**

There are several complications when it comes to heart diseases. We can reduce the risk of complications with early diagnosis and treatment. The solution that we provide from the website will help the patients. It is always to get treated in the early stages when it comes to heart disease.

## **12. Future Scope**

There are many possible improvements that could be explored to improve the scalability and accuracy of this prediction system. 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.

## **13. Appendix**

Source Code:

<https://tinyurl.com/Code-File>

Demo video link:

<https://tinyurl.com/demo-video-url>