Visualizing and Predicting Heart Diseaseswith an Interactive Dashboard

NALAIYA THIRAN PROJECT REPORT 2022

Submitted by

RAJASRI S   **510119104018**

ARCHANA N   **510119104002**

DHENISHA S   **510119104004**

VISHNU PRIYA M   **510119104021**

TEAM ID : PNT2022TMID39416

Content

VISUALIZING AND PREDICTING HEART DISEASES WITH AN INTERACTIVE DASHBOARD

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. This project aims to create an interactive Dashboard using IBM Cognos Tool and dataset to predict which patients are most likely to suffer from a heart disease in the near future using the features given.

1.2 Purpose
    Heart disease (HD) is a major cause of mortality in modern society. Medical diagnosis is an extremely important but complicated task that should be performed accurately and efficiently. Cardiovascular disease is difficult to detect due to several risk factors, including high blood pressure, cholesterol, and an abnormal pulse rate. Based on the analytics we can analyze which patients are most likely to suffer from heart disease in the near future and based on the patient details we will make decisions to cure them.

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 Exploratory Data Analysis" R. Indrakumari, T.Poongodi, Soumya Ranjan Jena
     In this paper, the risk factors that causes heart disease is considered and predicted using K-means algorithm and the analysis is carried out using a publicly available data for heart disease. The dataset holds 209 records with 8 attributes such as age, chest pain type, blood pressure, blood glucose level, ECG in rest, heart rate and four types of chest pain. To predict the heart disease, K means clustering algorithm is used along with data analytics and visualization tool. The paper discusses the pre-processing methods, classifier performances and evaluation metrics. In the result section, the visualized data shows that the prediction is accurate.

Prediction of heart disease at early stage using data mining and big data analytics: A survey N. K. Salma Banu, Suma Swamy
     Several studies have been carried out for developing prediction model using individual technique and also by combining two or more techniques. This paper provides a quick and easy review and understanding of available prediction models using data mining from 2004 to 2016. The comparison shows the accuracy level of each model given by different researchers.

2.3 Problem Statement Definition

Who does the problem affect?

It affects the person who have unhealthy diets, stress, blood pressure, those who have the habit of smoking and id

When does the issue occur?

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.

What would happen if we didn't solve the problem?

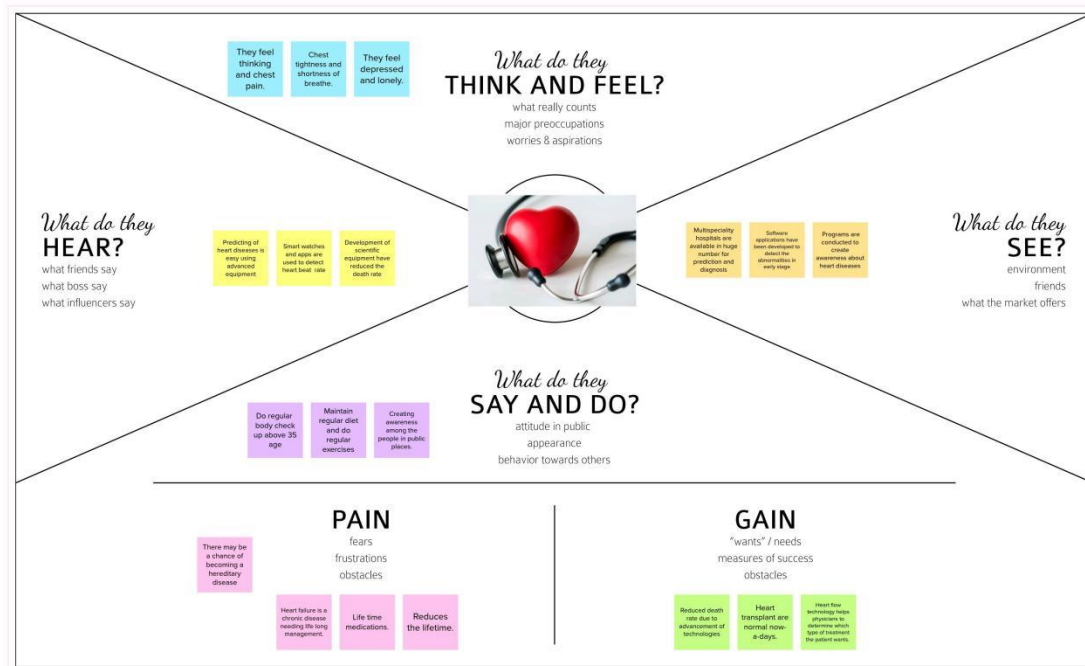If we don't solve the problem, many people will die at a young age. The death rate due to heart disease

will increase rapidly.

Why is it important to fix the problem?

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 Canvas

3.2 Ideation and Brainstorming

Step-1: Team Gathering,

Brainstorm & idea prioritization

Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

- 🕐 **10 minutes** to prepare
- ⧗ **1 hour** to collaborate
- 👤 **2-8 people** recommended

💬 Share template feedback

Before you collaborate

A little bit of preparation goes a long way with this session. Here's what you need to do to get going.

🕐 10 minutes

A Team gathering
Define who should participate in the session and send an invite. Share relevant information or pre-work ahead.

B Set the goal
Think about the problem you'll be focusing on solving in the brainstorming session.

C Learn how to use the facilitation tools
Use the Facilitation Superpowers to run a happy and productive session.

Open article →

1

Define your problem statement

What problem are you trying to solve? Frame your problem as a How Might We statement. This will be the focus of your brainstorm.

🕐 5 minutes

PROBLEM

Analyzing the people with symptoms like fainting,chest tightness,chest pain,shortness of breathe,slow/fast heartbeat and predicting heart disease in them.

Key rules of brainstorming
To run a smooth and productive session

- Stay in topic.
- Defer judgment.
- Go for volume.
- Encourage wild ideas.
- Listen to others.
- If possible, be visual.

Step-2: Brainstorm, Idea Listing and Grouping

2

Brainstorm

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

⏱ 10 minutes

TIP 💡
You can select a sticky note and hit the pencil [switch to sketch] icon to start drawing!

Person 1

Person 2

Person 3

Person 4

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 and break it up into smaller sub-groups.

⏱ 20 minutes

TIP 💡
Add customizable tags to sticky notes to make it easier to find, browse, organize, and categorize important ideas as themes within your mural.

Step-3: Idea Prioritization

Prioritize

Your team should all be on the same page about what's important moving
forward. Place your ideas on this grid to determine which ideas are important and
which are feasible.

⏱ 20 minutes

Importance

If each of these
tasks could get
done without any
difficulty or cost,
which would have
the most positive
impact?

Heart disease can
be predicted based
on signs, symptoms
and physical test of
the patient.

Heart disease can
be predicted and
diagnosed with the
help of K-means
clustering.

Person with
smoking history
may have
chance of heart
disease.

TIP 💡

Participants can use their
cursors to point at where
sticky notes should go on
the grid. The facilitator can
confirm the spot by using
the laser pointer holding the
H key on the keyboard.

Feasibility

Regardless of their importance, which tasks are more
feasible than others? (Cost, time, effort, complexity, etc.)

3.3 Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	The goal is to accurately create a data set about the Heart patients so that the hospital can use this information to easily visualize and predict the patient details.
2.	Idea / Solution description	The parameters in the data set helps hospitals to identify the patient heart condition and their health condition.A informative and creative dashboard can be created to present the data and utilize it for future use.
3.	Novelty / Uniqueness	Treatment can be easy for the doctors on the basis of the patient heart condition. Time can be saved.
4.	Social Impact / Customer Satisfaction	It help the hospitals to know the health records of the heart patient. It will make the hospital to work efficiently.
5.	Business Model (Revenue Model)	Ad based revenue model - Awareness can be created among the patient through ads.
6.	Scalability of the Solution	Easy prediction of the patient details with heart disease. Maintains best user experience.

3.4 Problem Solution Fit

1. CUSTOMER SEGMENT(S) `CS`

Who is your customer?
i.e. working parents of 0-5 y.o. kids

A person suffering from Cardiovascular disease.

6. CUSTOMER CONSTRAINTS `CC`

What constraints prevent your customers from taking action or limit their choices of solutions? i.e. spending power, budget, no cash, network connection, available devices.

The patient needs to visualize and predict the heart disease by blood tests or health monitoring diseases.

5. AVAILABLE SOLUTIONS `AS`

Which solutions are available to the customers when they face the problem
or need to get the job done? What have they tried in the past? What pros & cons do these solutions have? i.e. pen and paper is an alternative to digital notetaking

Take heart healthy diet, maintain a healthy weight, get good quality sleep, manage stress and don't smoke or use tobacco.

2. JOBS-TO-BE-DONE / PROBLEMS `J&P`

Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one; explore different sides.

The main problems related to heart disease are high blood pressure, high blood cholesterol, chest pain, shortness of breathe, angina, etc...

9. PROBLEM ROOT CAUSE `RC`

What is the real reason that this problem exists?
What is the back story behind the need to do this job?
i.e. customers have to do it because of the change in regulations.

The root cause for cardiovascular disease are unhealthy foods, lack of exercise, smoking, poor diet, high cholesterol and obesity.

7. BEHAVIOUR `BE`

What does your customer do to address the problem and get the job done?
i.e. directly related: find the right solar panel installer, calculate usage and benefits; indirectly associated: customers spend free time on volunteering work (i.e. Greenpeace)

Try to limit saturates fats, foods high in sodium and added sugars. Eat plenty of fresh fruit, vegetables and whole grains. Don't smoke or use tobacco.

3. TRIGGERS `TR`

What triggers customers to act? i.e. seeing their neighbour installing solar panels, reading about a more efficient solution in the news.

Symptoms like heavy heartbeat rate, shortness of breathe triggers them to take tests related to heart disease. People are also triggered to experiment tests by what they see and hear.

4. EMOTIONS: BEFORE / AFTER `EM`

How do customers feel when they face a problem or a job and afterwards?
i.e. lost, insecure > confident, in control – use it in your communication strategy & design.

The patient often results in depression, anxiety, isolation, loneliness, sadness diminished self-esteem, etc.......

10. YOUR SOLUTION `SL`

If you are working on an existing business, write down your current solution first, fill in the canvas, and check how much it fits reality.
If you are working on a new business proposition, then keep it blank until you fill in the canvas and come up with a solution that fits within customer limitations, solves a problem and matches customer behaviour.

Create an Interactive dashboards about the patient, which help in categorizing them under unique classification such as age, heartbeat range, etc....

8. CHANNELS of BEHAVIOUR `CH`

8.1 ONLINE
What kind of actions do customers take online? Extract online channels from #7

8.2 OFFLINE
What kind of actions do customers take offline? Extract offline channels from #7 and use them for customer development.

Based on informations in dashboard we can predict which patients are most likely to suffer from heart disese.

4. RequirementAnalysis

4.1 Functional Requirement

Following are the functional requirements of the proposed solution.

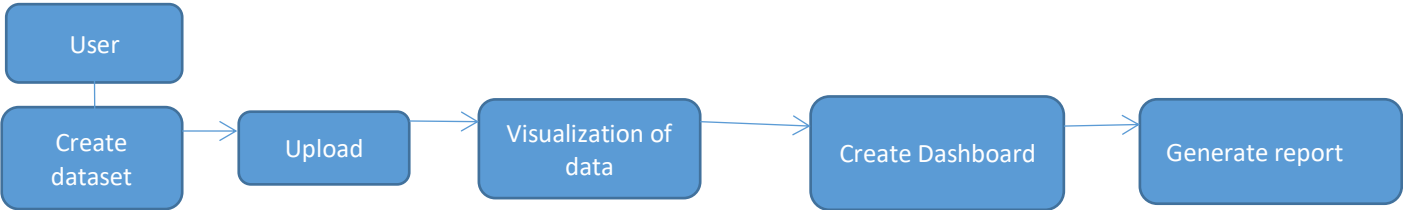| FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	User will be Register the application via email
FR-2	User Confirmation	User gets the confirmation mail via email, once they successfully registered
FR-3	Visualizing data	User can visualize the trends in the heart disease analysis through Dashboard created using Cognos Analytics.
FR-4	Analyzing data	User can analyze the data present in the Dashboard
FR-5	Generate Report	User can view his/her report through analysis and take futher steps accordingly.

Non-functional Requirements:

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

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	This is an user friendly application and simple to use. This shows graphical interface which is very easy to understand visually. User can see relationship between various data with just few clicks.
NFR-2	**Security**	For security of the application database replication and duplication should be used. For instance if the system may not work we can easily backup and recover the data. Personal details of user cannot be shared anymore.
NFR-3	**Reliability**	The application will give exact information. It shows exact relationship between data during visualization.
NFR-4	**Performance**	Performance of the application depends upon the response time of the data submission which is depends on the efficiency of implemented algorithm.
NFR-5	**Availability**	The application should be available 24x7 for users without any interruption.

NFR-6	Scalability	The application achieves saclability, because it can withstand when no of data is increased. We can upload many dataset in the sametime.

5. Project Design

5.1 Data Flow Diagram

```
User
Create dataset  →  Upload  →  Visualization of data  →  Create Dashboard  →  Generate report
```

5.2 Solution and Technical Architecture

```
Method 1

Dataset  →(Method 1)→  Classification  →  Performance Evaluation  →  Results
Dataset  →(Method 1)→  Feature sSection  →  Prediction  →  Performance Evaluation
```
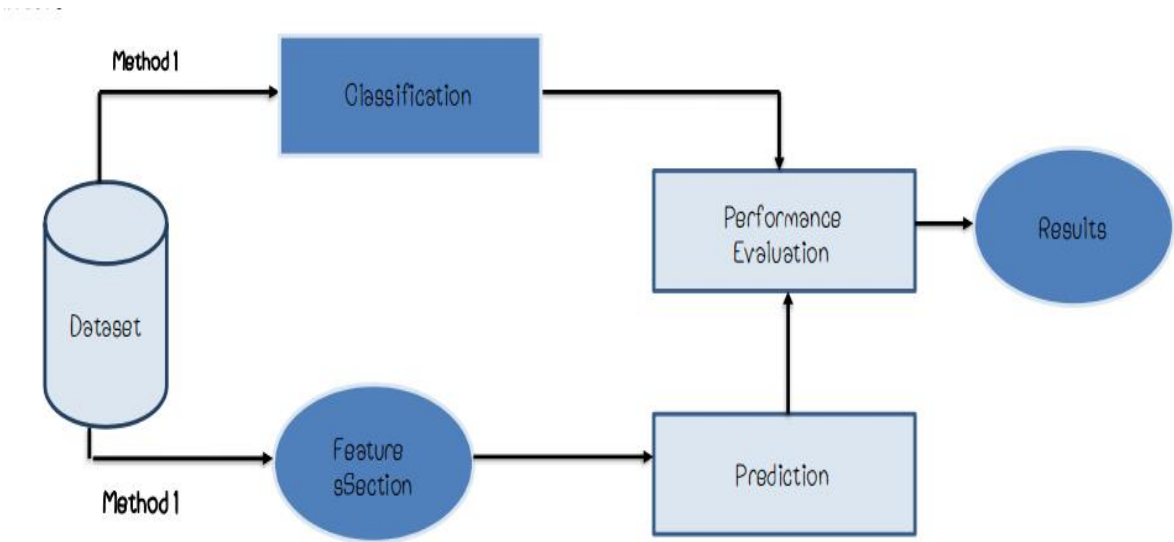
5.3 User Stories:

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria
Customer (Web user)	Registration	USN-1	As a user, I can register for the dashboard by entering my email, password, and confirming my password.	I can access my account / dashboard
		USN-2	As a user, I will receive confirmation email once I have registered for the dashboard.	I can receive confirmation email & click confirm
	Login	USN-3	As a user, I can log into the dashboard by entering email & password	I can register & access the dashboard
	Dashboard	USN-4	As a user, they can view complete medical history.	I can view my medical history in the dashboard.
		USN-5	As a user, they can predict the ocuurence of heart disease.	I can view the accuracy of heart disese
Customer Care Executive	Helpdesk	USN-6	As a customer care executive, they can view the patient problems.	I can send my problem in the dashboard
		USN-7	As a customer care executive, they can cure the patient problems.	I can get help from the helpdesk
Administrator	User profile	USN-8	As an admin, they can add or delete the patient details.	I can acess my dashboard
		USN-9	As an admin, they can update the health issues of the patient.	I can view my updated health details
		USN-10	As an admin, they can manage the patient details	I can view my complete the health details

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	Database Connectivity and Upload of data set	USN-1	We need to collect and fetch the dataset from the external API and connect it with database using IBM Cognos and upload the dataset	2	High	RAJASRI.S ARCHANA.N DHENISHA.S VISHNU PRIYA.M
Sprint-1	Data Modules & Data Exploration	USN-2	With the uploaded dataset we create a data module and perform data exploration	1	Medium	RAJASRI.S ARCHANA.N DHENISHA.S VISHNU PRIYA.M
Sprint-2	Dashboard	USN-3	Create an interactive dashboard after performing the data exploration	2	High	RAJASRI.S ARCHANA.N DHENISHA.S VISHNU PRIYA.M
Sprint-3	Report and Stories	USN-4	Create Report and User stories based on the dashboard	2	High	RAJASRI.S ARCHANA.N DHENISHA.S VISHNU PRIYA.M
Sprint-4	Web Application	USN-5	Create a web application for dashboard, report and user stories	1	High	RAJASRI.S ARCHANA.N DHENISHA.S VISHNU PRIYA.M

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

6.3 Reports from JIRA
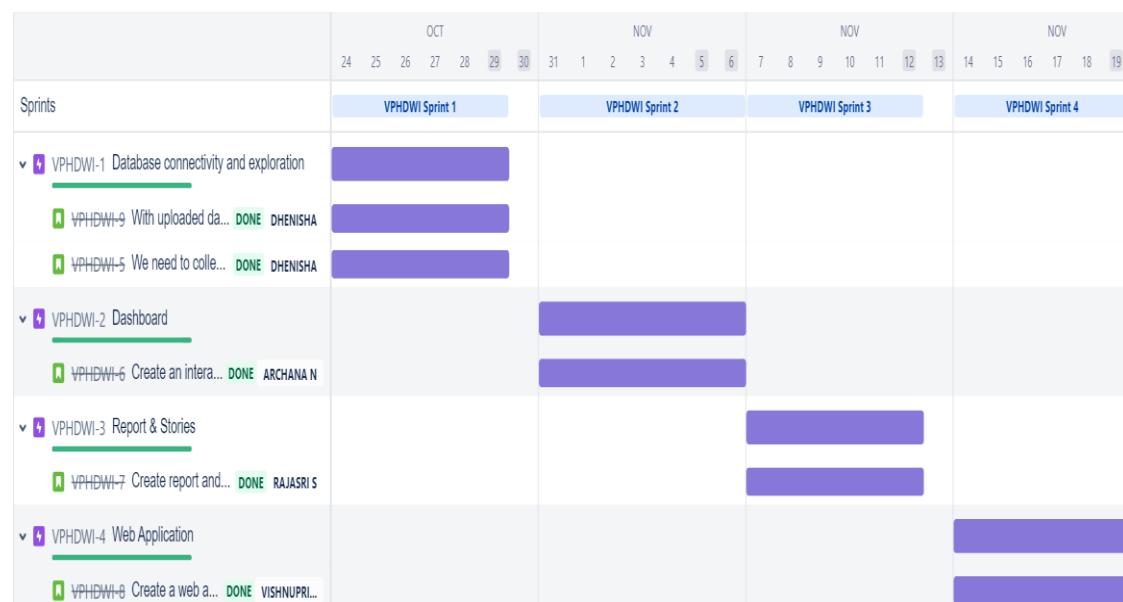
7 Coding And Solutioning

Coding for the design of web page to represent the Dashboard,Report and Story.

INDEX.html

```
<html>
<head>
<title> HEART DISEASE PREDICTION </title>
<link rel="stylesheet" href="Style.css">
<style>
a:hover {
  background-color: crimson;
}
</style>
</head>
<body>
```

```html
<div class="menu-bar">
<ul>
<li class="active"><a href="#">Home</a></li>
<li><a
href="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2F
Heart%2BDisease%2BPrediction%2Bdashboard&action=view&mode=dashboard&subView
=model0000018460dbcb2d_00000002">Dashboard</a></li>
<li><a
href="https://us3.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FHeart%2BDisease%2BP
rediction%2Breport&action=run&format=HTML&prompt=false">Report</a></li>
<li><a
href="https://us3.ca.analytics.ibm.com/bi/?perspective=story&pathRef=.my_folders%2FHear
t%2Bdisease%2Bprediction%2Bstory&action=view&sceneId=model0000018474ff5d5c_000
00000&sceneTime=0">Story</a></li>
</ul>
</div>
<div class="content">
<h1>HEART DISEASE PREDICTION</h1>
<p>Let us have a look on Dashboard,Report and Story that is used for predicting the heart
disease</p>
</div>
</body>
</html>
```

STYLE.css

```html
<html>
<head>
<title> HEART DISEASE PREDICTION </title>
<link rel="stylesheet" href="Style.css">
<style>
a:hover {
  background-color: crimson;
}
</style>
</head>
<body>
<div class="menu-bar">
<ul>
<li class="active"><a href="#">Home</a></li>
<li><a
href="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2F
Heart%2BDisease%2BPrediction%2Bdashboard&action=view&mode=dashboard&subView
=model0000018460dbcb2d_00000002">Dashboard</a></li>
<li><a
href="https://us3.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FHeart%2BDisease%2BP
rediction%2Breport&action=run&format=HTML&prompt=false">Report</a></li>
<li><a
href="https://us3.ca.analytics.ibm.com/bi/?perspective=story&pathRef=.my_folders%2FHear
t%2Bdisease%2Bprediction%2Bstory&action=view&sceneId=model0000018474ff5d5c_000
00000&sceneTime=0">Story</a></li>
</ul>
</div>
<div class="content">
<h1>HEART DISEASE PREDICTION</h1>
```
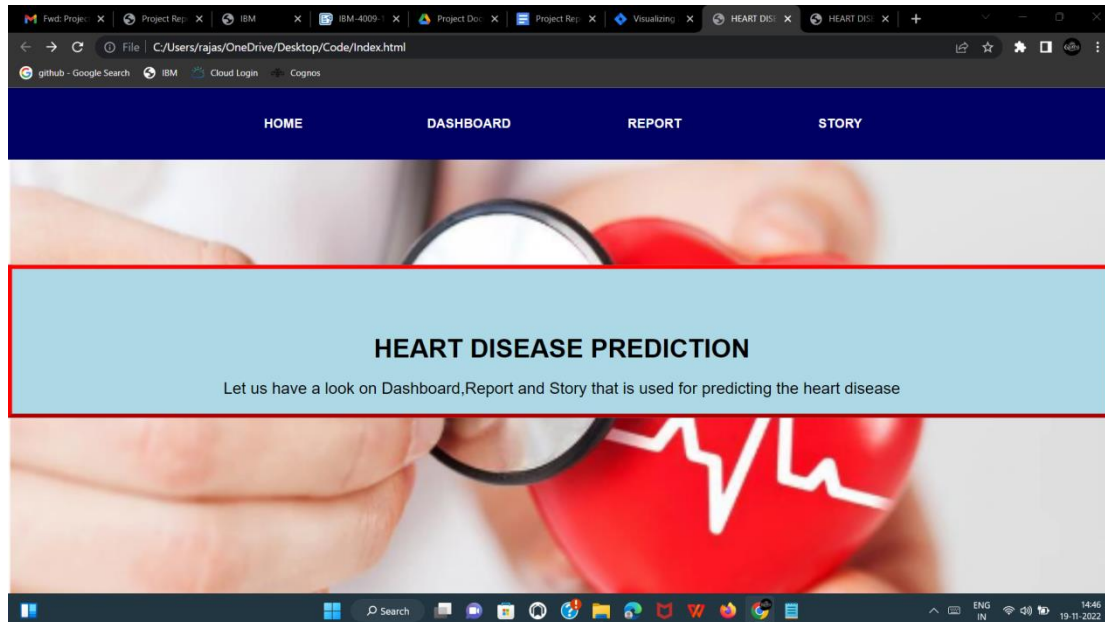
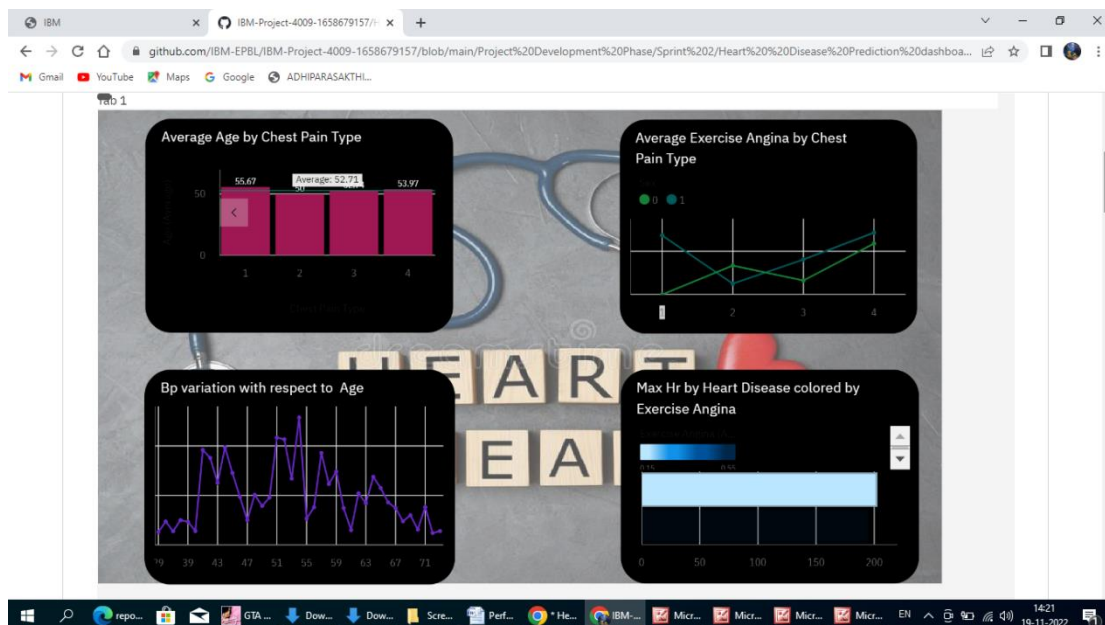<p>Let us have a look on Dashboard,Report and Story that is used for predicting the heart disease</p>
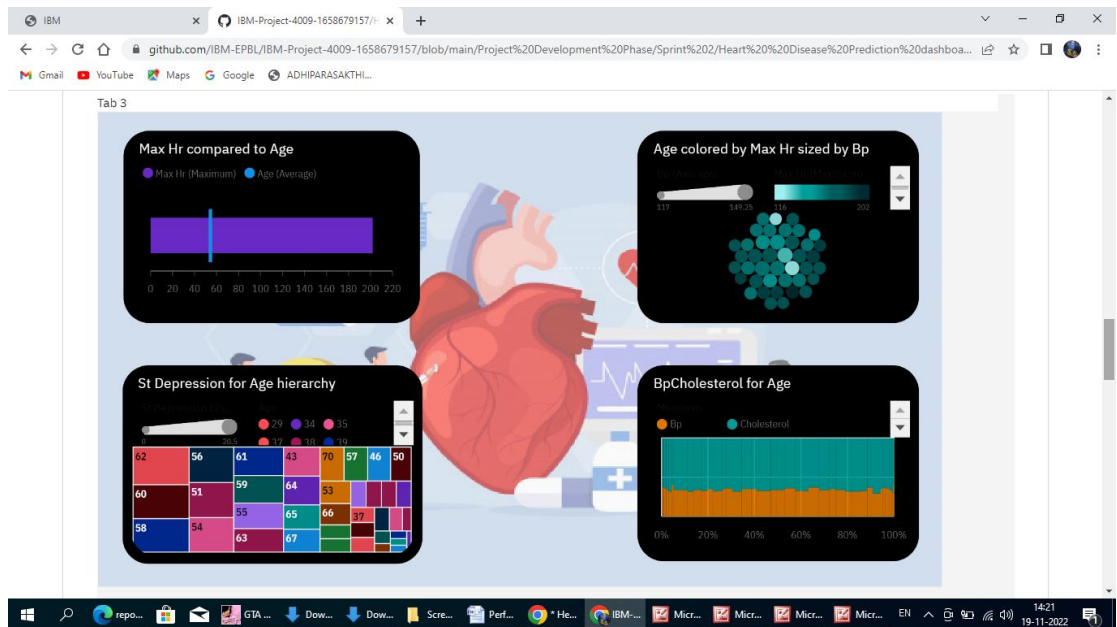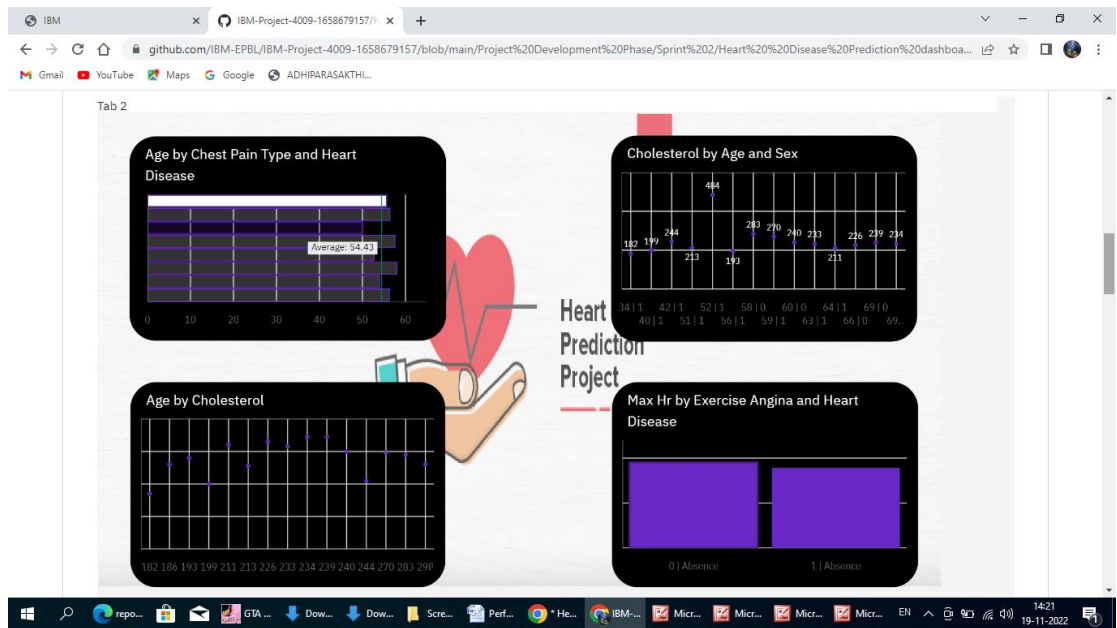</div>
</body>
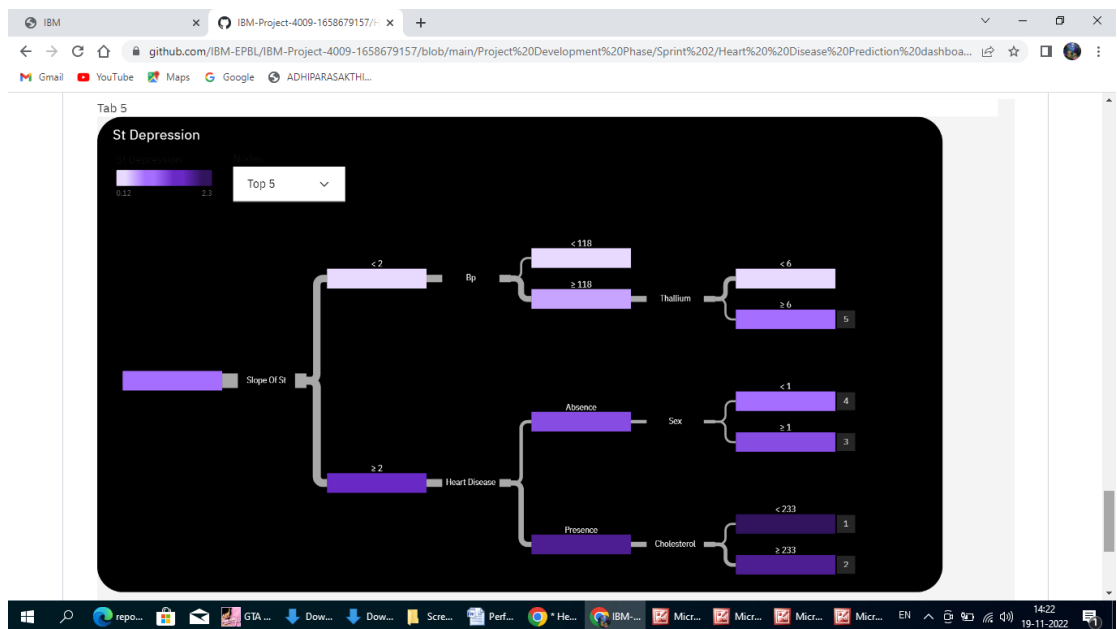</html>

WEB PAGE

DASHBOARD

If we click on the dashboard in the above webpage we get the following explorations.

Tab 4 — Heart Disease Prediction dashboard showing Bp by Max Hr, EKG Results by Age and Sex, St Depression hierarchy colored by Age, and Cholesterol by Age colored by Sex.

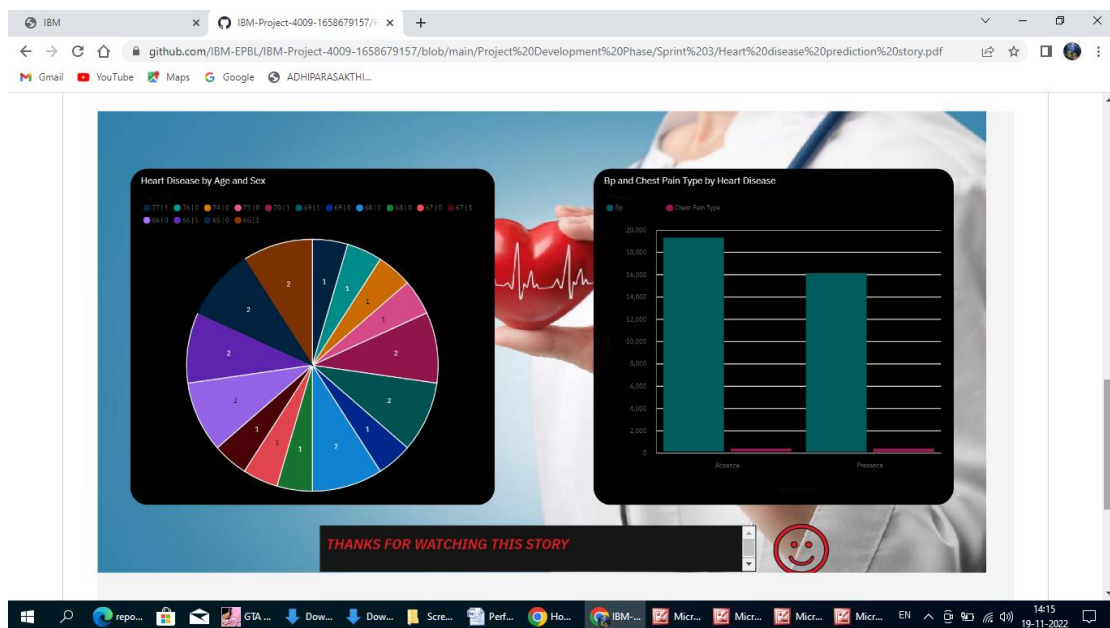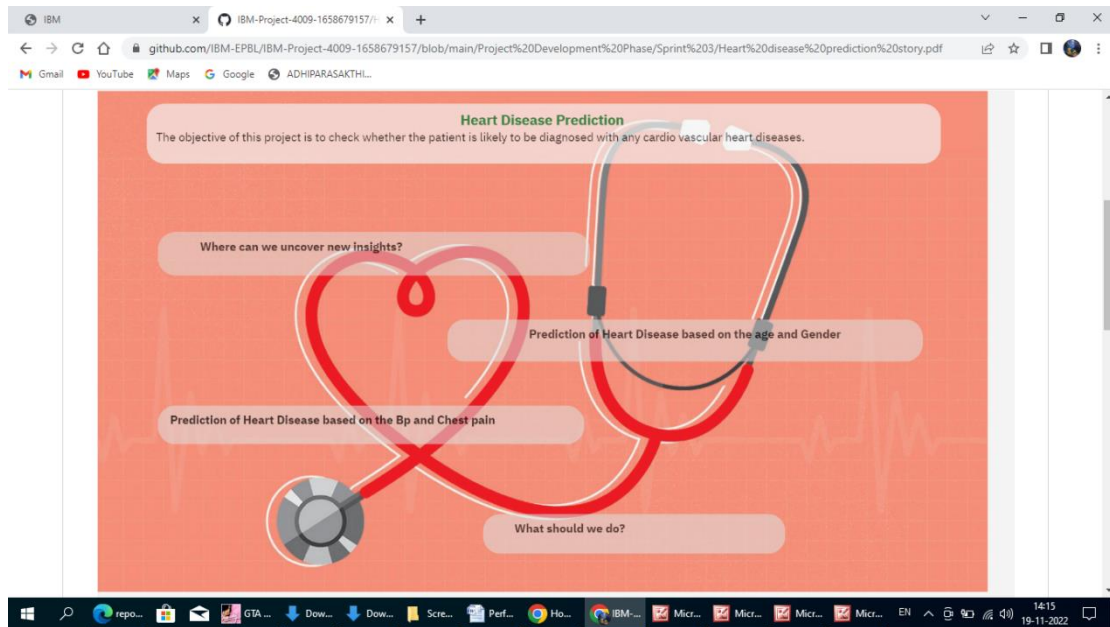Tab 5 — St Depression decision tree (Top 5).
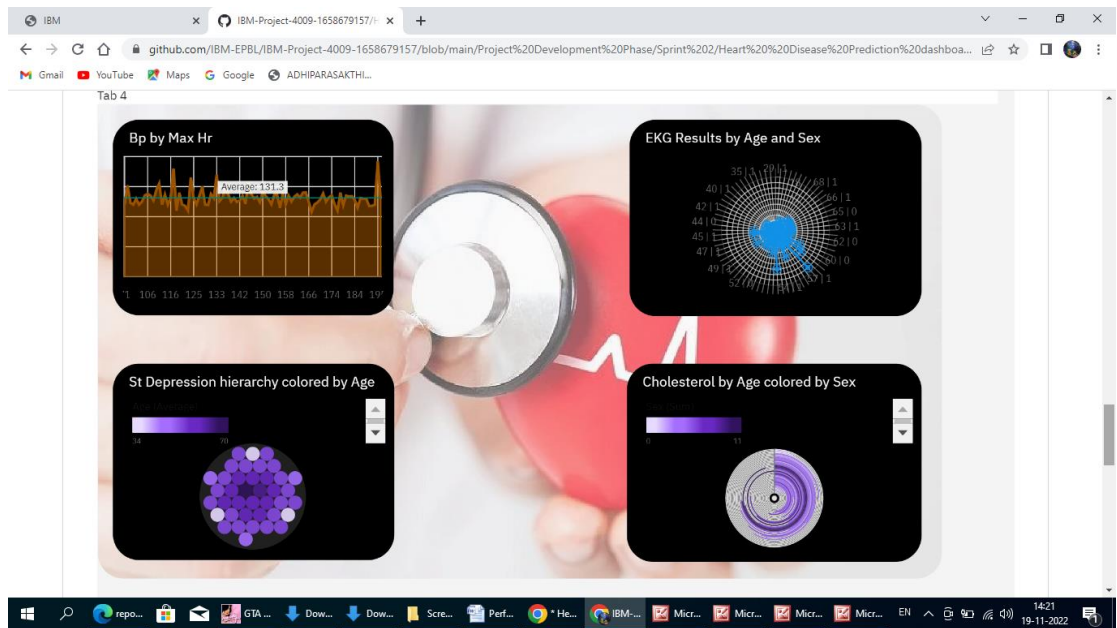
REPORT

If we click on the report in the webpage we get the following explorations.

STORY

If we click on the story in the given webpage we get the following scenes.

8. TESTING
8.1 User Acceptance Testing

Purpose of Document

The purpose of this document is to briefly explain thetestcoverageandopenissuesofthe [ProductName] project at the time of the release to User Acceptance Testing (UAT).

DefectAnalysis

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

Resolution	Severity1	Severity 2	Severity3	Severity4	Subtotal
By Design	10	4	2	3	20
Duplicate	1	0	3	0	4
External	2	3	0	1	6
Fixed	11	2	4	20	37
Not Reproduced	0	0	1	0	1
Skipped	0	0	1	1	2
Won'tFix	0	5	2	1	8
Totals	24	14	13	26	77

TestCaseAnalysis

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

Section	TotalCases	Not Tested	Fail	Pass
Dashboard	7	0	0	7
Report	2	0	0	2
Story	2	0	0	2
Webpage	3	0	0	3
Exception Reporting	9	0	0	9
Final Report Output	4	0	0	4
VersionControl	2	0	0	2

9. ADVANTAGES & DISADVANTAGES

ADVANTAGES

- It reduces the time complexity of doctors.

- It is very cost effective for patients.

- This is one of the fastest ways to determine if a person is likely to suffer from a heart disease or not.

- The dashboard provides insightful informations.

DISADVANTAGES

- Prediction of cardiovascular disease results is not accurate.

- It cannot handle enormous datasets for patient records.

- It does not provide suggestions to user.

10.CONCLUSION

Chest pain is the major key to predict the heart disease. In this project, the heart diseases are predicted by considering the major factors with four types of chest pain.Complications of heart disease include heart attack and stroke. You can reduce the risk of complications with early diagnosis and treatment. So the suggestion that we get from the website might help save patients. It is always to get treated in the early stage itself.

11.FUTURE SCOPE

Like the saying goes "Prevention is better than cure". We have to look into methods to prevent heart diseases altogether other than just predicting it in early stages.In the future more machine learning approaches will be used for the prediction of heart disease in the early stage itself so that the number of deaths can be reduced. The goal of our heart disease prediction project is to determine if a patient should be diagnosed with heart disease or not.

12. APPENDIX

PROJECT DEMO LINK:
https://drive.google.com/file/d/1JrD6aMXq3yuyHVqCkka3HAq3-3O7WoAl/view?usp=drivesdk

GITHUB LINK: https://github.com/IBM-EPBL/IBM-Project-4009-1658679157