PROFESSIONAL READINESS FOR INNOVATION, EMPLOYABILITYAND ENTREPRENEURSHIP

VISUALIZING AND PREDICTING HEART DISEASE WITH AN INTERACTIVE DASHBOARD Team ID - PNT2022TMID17743

PROJECT REPORT FORMAT

1. INTRODUCTION

1.1 Project Overview

This project aims to visualize and predict the presence of heart disease based on various factors that might cause heart disease. We will analyze the heart disease dataset and gain insights which then helps to know the factor which causes high severity to patients. Based on that, we will be choosing a suitable machine learning model to detect the probability of a patient having heart disease. Machine Learning models helps us to make better decisions. Even though, heart disease might occur in variant forms, there are several factors which are quite common. Taking those into account, we propose an effective method to predict the heart disease by analyzing and visualizing the causes of heart disease.

1.2 Purpose

At all times, doctors might not be available due to some reasons. So, a patient will not be able to consult doctors in case of emergency. In order to eliminate these problems, Heart Disease Prediction application, an online consultant helps user to know if they have heart disease by providing immediate results. Our research will help students and doctors better understand the causes of heart attacks and has led to improved treatment options for this condition. This helps find ways to prevent and treat heart attacks in an earlier manner.

2. LITERATURE SURVEY

2.1 Existing Problem

[a] Heart Disease Prediction using Exploratory Data Analysis (2020)

R.Indrakumari, et al.(2020), worked with K-means unsupervised algorithm on heart disease prediction by exploiting features of publicly available dataset for heart disease. The study finds necessary hidden patterns for predicting heart diseases and the proposed method is robust to noisy training data. But, KNN works well with small number of input variables but as the numbers of variables grow K-NN algorithm struggles to predict the output of new data point.

[b] Weight optimized neural network for heart disease prediction using hybrid lion plus particle swarm algorithm (2020)

Renji P. Cherian, et al. (2020), proposed a new heart disease prediction model with the inclusion of specific processes like Feature Extraction, Record, Attribute Minimization and Classification. The proposed method has improved accuracy in the prediction and works well on dimensionally reduced features. But, it is computationally expensive than traditional algorithms.

[c] Heart disease prediction using supervised machine learning algorithms (2021)

Md Mamun Ali, et.al (2021) worked with machine learning classifiers for highest accuracy of such diagnostic purposes. The RF method achieved 100% accuracy along with 100% sensitivity and specificity. But, the data quantity on heart disease data provided by this dataset was not large enough to adequately address all issues.

[d] Heart disease prediction system using machine learning (2021)

K.Aruljothi, et.al(2021) proposed a system which predicts the chances of getting heart disease in advance using KNN and Decision tree algorithm for guessing the risk level of heart disease. The proposed work can be used for the automated work analysis and reduces the manual work like other existing systems. This paper didn't work on multiclass classification problem.

[e] Heart disease prediction based on pre-trained deep neural networks combined with principal component analysis (2022)

Dimman Hassan, et al. (2022) proposed a new approach utilizing a pre-trained Deep Neural Network (DNN) for feature extraction, Principal Component Analysis (PCA) for dimensionality reduction, and Logistic Regression (LR) for prediction. The method provides high accuracy rate not only on the training data but also on the testing data, resulting in outstanding generalization capabilities. No sign of overfitting or underfitting in the proposed model, which is the main problem that plagues the DNN-like methods.

2.2 References

1. R. Indrakumari , T. Poongodi , Soumya Ranjan Jena, "**Heart Disease Prediction using Exploratory Data Analysis**", ELSEVIER – Smart Sustainable Intelligent Computing, (2020).
2. Renji P. Cherian, Noby Thomas, Sunder Venkitachalam, "**Weight optimized neural network for heart disease prediction using hybrid lion plus particle swarm algorithm**", ELSEVIER – Journal of Biomedical Informatics, Vol.110, (2020).
3. Md Mamun Ali, Bikash Kumar Paul, Kawsar Ahmed, Francis M. Bui, Julian M. W. Quinn, Mohammad Ali Moni, "**Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison**", ELSEVIER – Computers in Biology and Medicines, Vol.136, (2021).
4. K. Arul Jothi, S. Subburam, V. Umadevi, K. Hemavathy, "**Heart disease prediction system using machine learning**", ELSEVIER – Materials Today: Proceedings, (2021).
5. Diman Hassan, Haval I. Hussein, Masoud M. Hassan, "**Heart disease prediction based on pre-trained deep neural networks combined with principal component analysis**", ELSEVIER – Biomedical Signal Processing and Control, (2022).

2.3 Problem Statement Definition

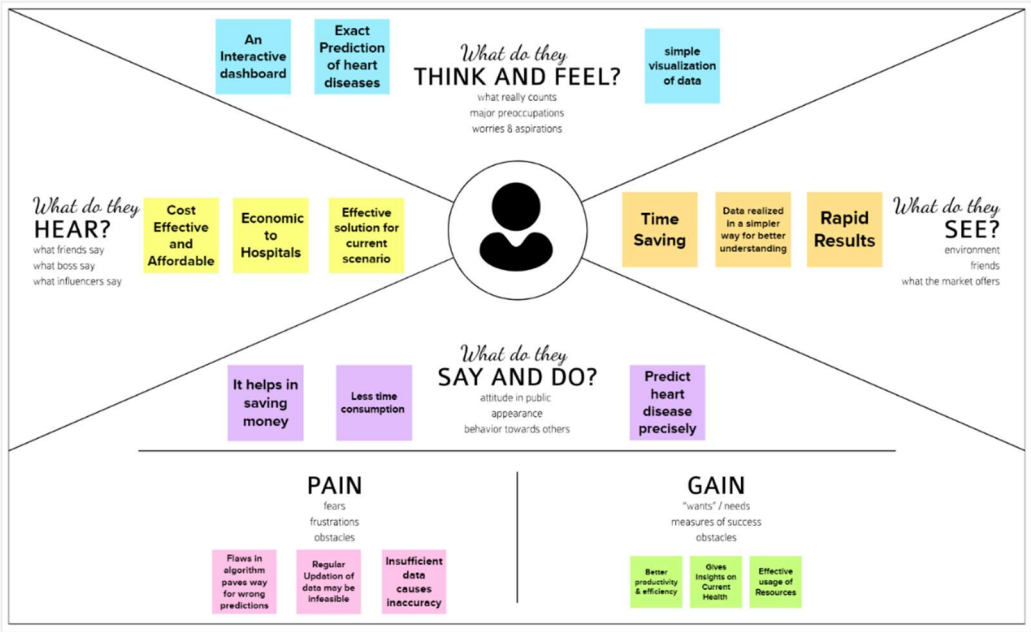
Heart disease refers to several types of abnormalities in heart conditions. The leading cause of death in the developed world is heart disease. It is infeasible for a common man to frequently undergo costly tests for ECG and so on. Hence, there needs a replacement to be done which must be handy and reliable. Thus, we propose an interactive dashboard for visualizing and predicting heart diseases in which user can view his medical report analysis and the predicted final result.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas:

An empathy map is a simple, easy-to-digest visual that captures knowledge about a user's behavior and attitudes. It is a useful tool to helps teams better understand their users. Creating an effective solution requires understanding the true problem and the person

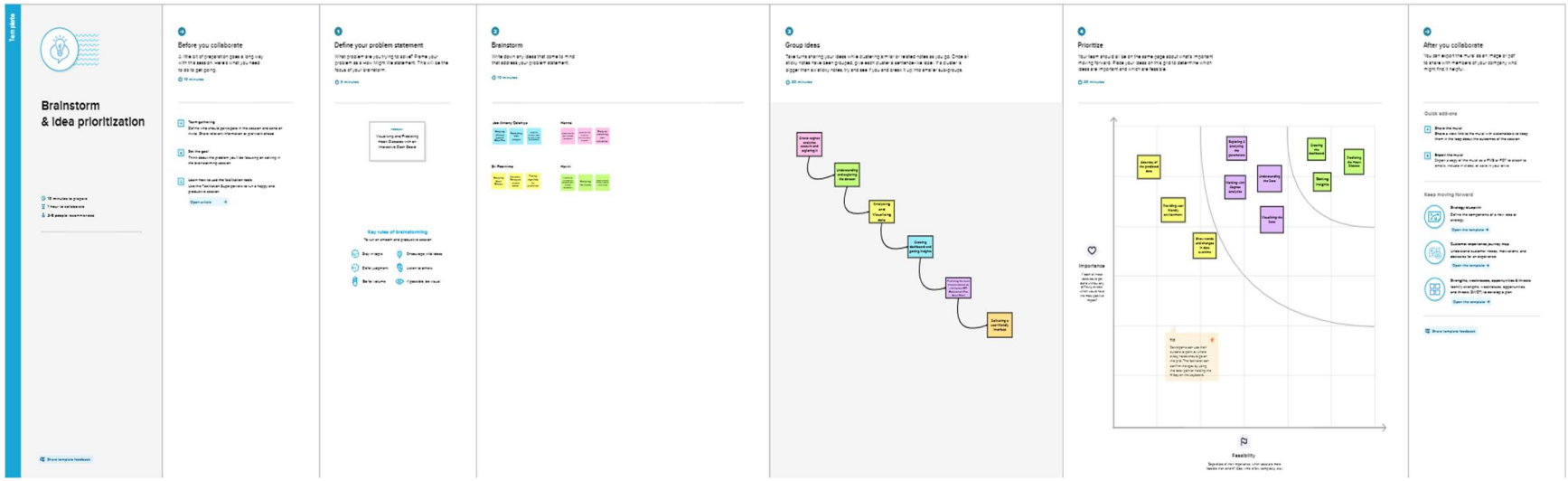
who is experiencing it. The exercise of creating the map helps participants consider things from the user's perspective along with his or her goals and challenges.



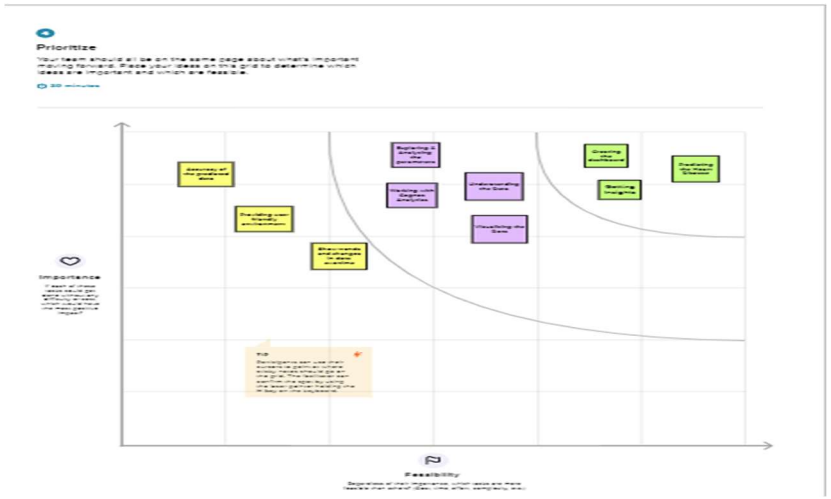
3.2 Ideation and Brainstorming:

3.2.1 Brainstorm, Idea Listing and Grouping

Brainstorming provides a free and open environment that encourages everyone within a teamto participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants areencouraged to collaborate, helping each other develop a rich amount of creative solutions.



3.2.2 Idea Prioritization



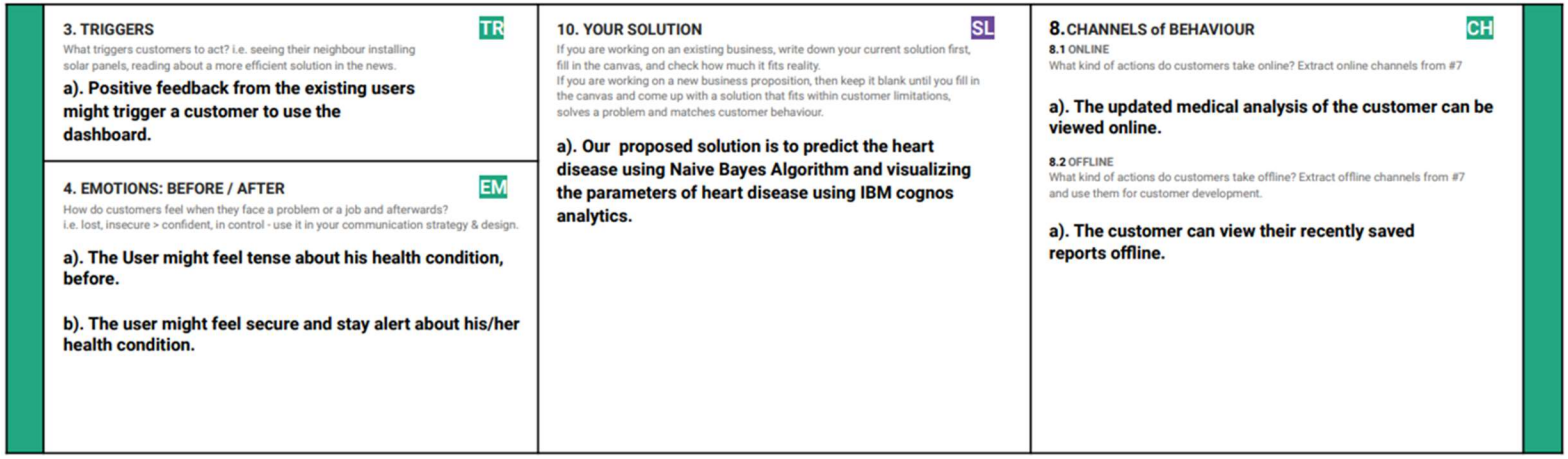
3.3 Proposed Solution

Project team shall fill the following information in proposed solution template.

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Heart disease refers to several types of abnormalities in heart conditions. The leading cause of death is heart disease. It is infeasible for a common man to frequently undergo tests for ECG and so on. Hence, there needs a replacement for this, which must be handy and reliable.
2.	Idea / Solution description	The idea behind the proposed solution is to propose an interactive dashboard for visualizing and predicting heart diseases in which user can view his/her medical report analysis and the predicted final result. The dashboard will be generated using IBM Cognos. The heart disease will be predicted using Decision Tree Algorithm.
3.	Novelty / Uniqueness	The novelty behind the proposed system is to provide suggestions to the user based on his/her medical analysis. It will provide the preventive measures to take care of the user himself.
4.	Social Impact / Customer Satisfaction	The system helps the user as well as the doctor to make better decisions to predict heart disease. It is useful in predicting the disease in an earlier stage and makes the user alert about his current condition periodically.

5.	Business Model (Revenue Model)	This interactive dashboard for heart disease prediction can be deployed in Health care centres and Hospitals, so that it makes the analysis in a fast manner.
6.	Scalability of the Solution	The proposed solution will work efficiently in both smaller and larger datasets in a similar manner. In future, it can be changed to predict some other diseases with more accuracy.

3.4 Problem Solution fit



4. Requirement Analysis

4.1 Functional requirements

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 Gmail
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 created using IBM Cognos Analytics
FR-4	Generating Report	User can view his/her health report and can make decisions accordingly

4.2 Non-Functional requirements

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

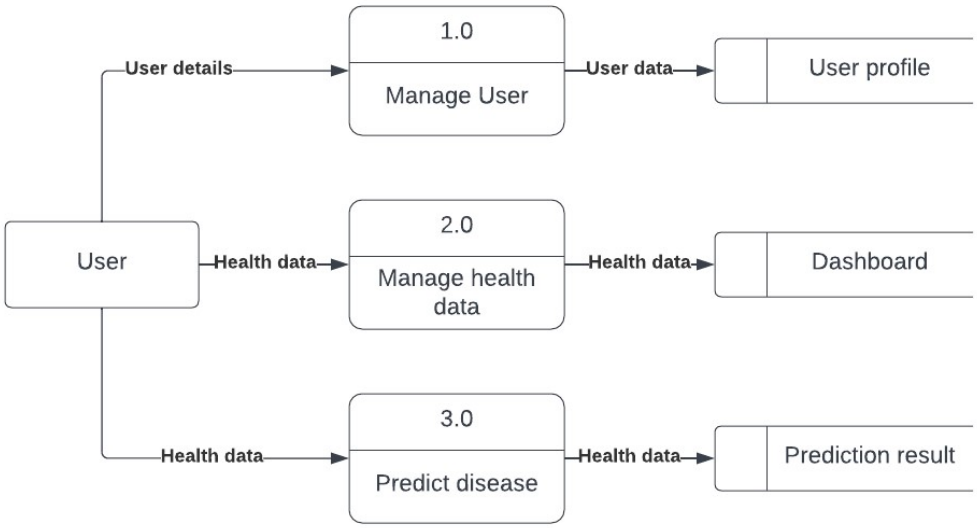
NFR No.	**Non-Functional Requirement**	**Description**
NFR-1	Usability	The application will have a simple and user-friendly graphical interface. Users will be able to understand and use all the features of the application easily. Any action has to be performed with just a few clicks
NFR-2	Security	For security of the application the technique known as database replication should be used so that all the important data should be kept safe. In case of crash, the system should be able to backup and recover the data

NFR-3	Reliability	The application has to be consistent at every scenario and has to work without failure in any environment
NFR-4	Performance	Performance of the application depends on the response time and the speed of the data submission. The response time of the application is direct and faster which depends on the efficiency of implemented algorithm
NFR-5	Availability	The application has to be available 24 x 7 for users without any interruption
NFR-6	Scalability	The application can withstand the increase in the no. of users and has to be able to develop higher versions

5.PROJECT DESIGN

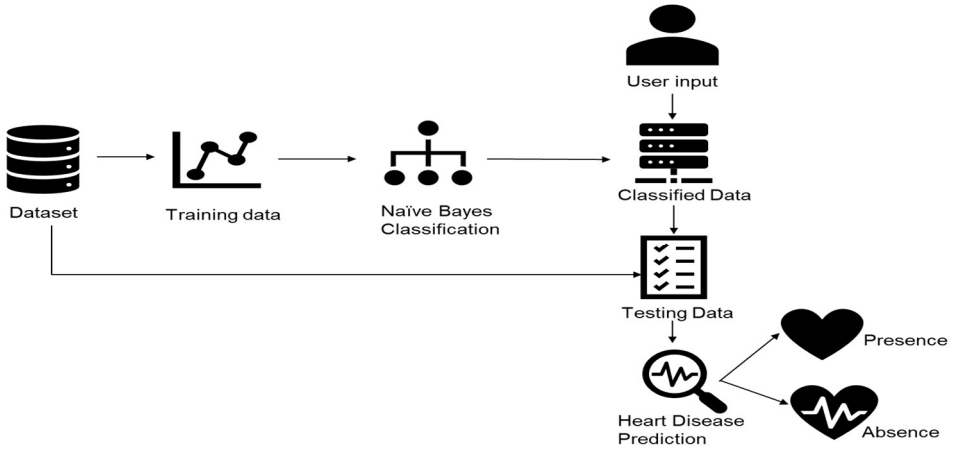
5.1 Data Flow Diagrams

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the rightamount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

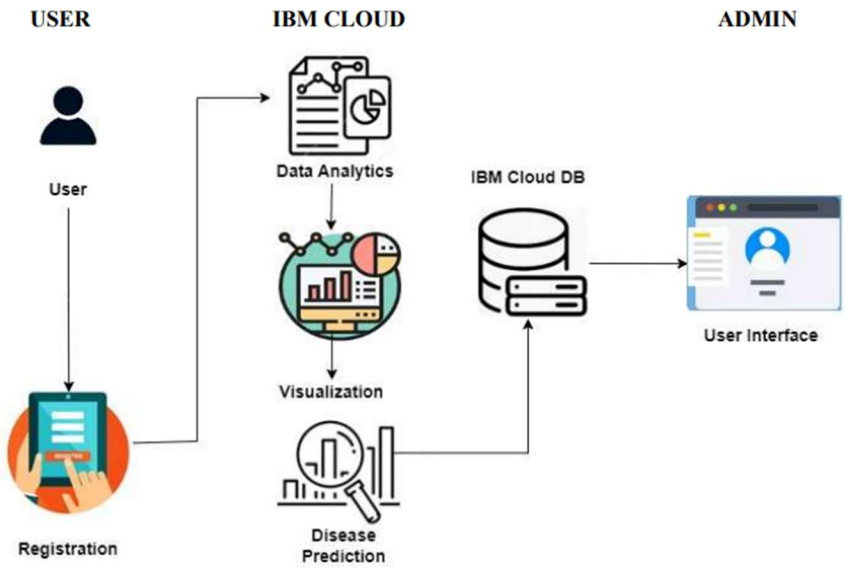


5.2 Solution & Technical Architecture

The below diagram represents the solution architecture.



The below diagram represents the technical architecture.



5.3 User Stories

User Story No	User Story/Talk
USN-1	As a user, I can view the details of data.
USN-2	As an analyst, I will explore the data and clean the data for building the efficient model.
USN-3	As an analyst, I will create a data module
USN-4	As an analyst, I will create an exploratory data analysis to identify the significant factors for disease prediction.
USN-5	As an analyst, I will visualize the various causes for the disease using various chart types.
USN-6	As an analyst, I will generate different models to find the model with high accuracy and efficiency.

USN-7	As an analyst, I will choose a suitable model for my solution.
USN-8	As an analyst, I will create a dashboard showing the analysis of the user.
USN-9	As an analyst, I will generate the result of heart disease prediction.
USN-10	As an analyst, I will import my model into a suitable framework.

6.PROJECT PLANNING & SCHEDULING

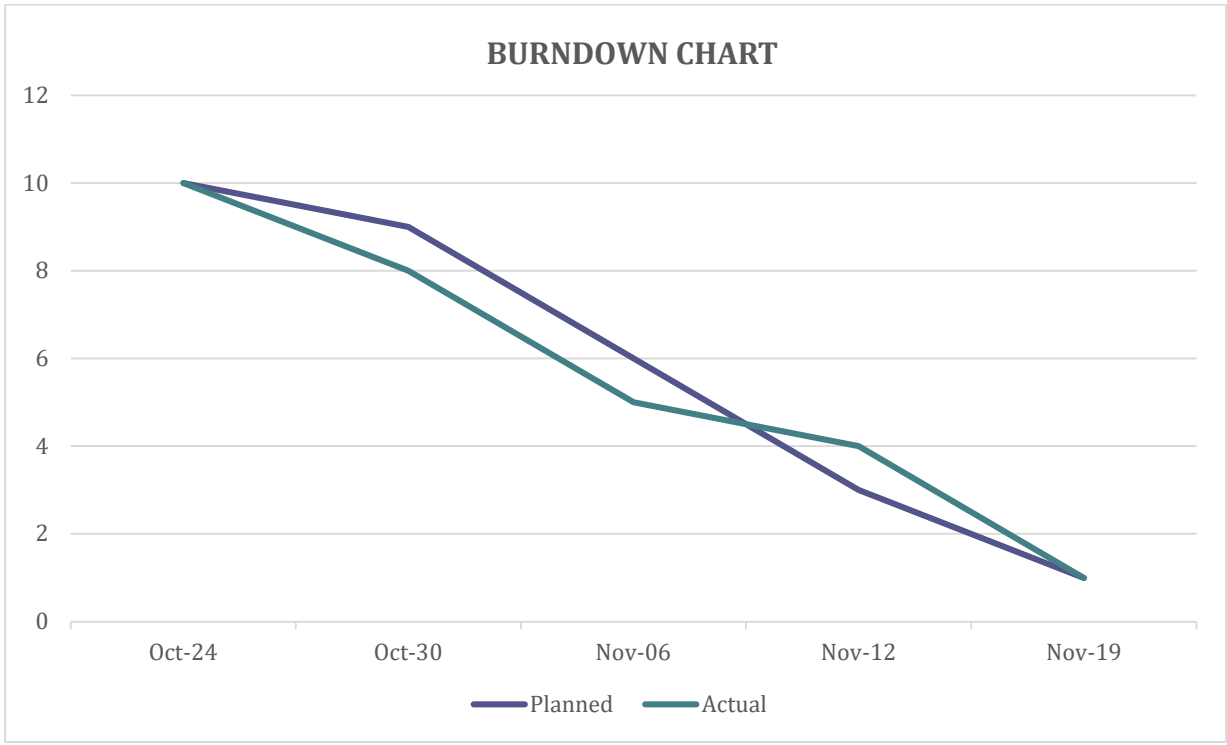
6.1 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Exploring the dataset	USN-1	As a user, I can view the details of data.	3	High	Joe Antony Celshiya JA
		USN-2	As an analyst, I will explore the data and clean the data for building the efficient model.	3	High	Harini G
		USN-3	As an analyst, I will create a data module.	4	High	Sripoornima T
Sprint-2	Data Visualization	USN-4	As an analyst, I will create an exploratory data analysis to identify the significant factors for disease prediction.	5	High	Hanna S, Sripoornima T
		USN-5	As an analyst, I will visualize the various causes for the disease using various chart types.	5	High	Hanna S, Sripoornima T
Sprint-3	Disease Prediction	USN-6	As an analyst, I will generate different models to find the model with high accuracy and efficiency	5	Medium	Joe Antony Celshiya JA, Harini G
		USN-7	As an analyst, I will choose a suitable model for my solution.	5	High	Joe Antony Celshiya

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
						JA, Harini G
Sprint-4	Dashboard Creation	USN-8	As an analyst, I will create a dashboard showing the analysis of the user	3	High	Harini G
		USN-9	As an analyst, I will generate the result of heart disease prediction.	3	High	Sripoornima T
		USN-10	As an analyst, I will import my model into a suitable framework.	4	High	Joe Antony Celshiya JA, Hanna S

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

BurnDown Chart:



6.2 Sprint Delivery Schedule

A milestone schedule, or milestone chart, is a timeline that uses milestones to divide a project schedule into majorphases. A milestone chart is a way to visualize the most important steps of our project. Each milestone the team achieves brings us closer to completing the project. As a result, milestones provide a sense of accomplishment andshow the team how the work they're doing contributes to the overarching project objective.

Activity Name	Activity Number	Activity Description	Tasks Assigned	Status
Preparation Phase	1	a) Access the resources in project dashboard b) Explore the dataset provided in workspace c) Create GitHub account & collaborate with Project Repository in project workspace d) Set-up the prerequisites for the project	SRIPOORNIMA T HANNA S HARINI G JOE ANTONY CELSHIYA JA	Completed
Ideation Phase	2	a) Literature survey relevant to the selected project b) Preparation of Empathy Map to identify the user pros and cons c) List the ideas by organizing the brainstorming session and prioritize the top 3 ideas based on the feasibility & importance	SRIPOORNIMA T HANNA S HARINI G JOE ANTONY CELSHIYA JA	Completed
Project Design Phase-I	3			
Proposed Solution	3.1	Preparation of proposed solution document, which includes the novelty, feasibility of idea, business model, social impact and solution scalability solution	HANNA S	Completed
Problem Solution Fit	3.2	Prepared problem solution fit which provides effective solutions for the problem	JOE ANTONY CELSHIYA JA	Completed
Solution Architecture	3.3	Develop effective architecture for the proposed solution	SRIPOORNIMA T HARINI G	Completed

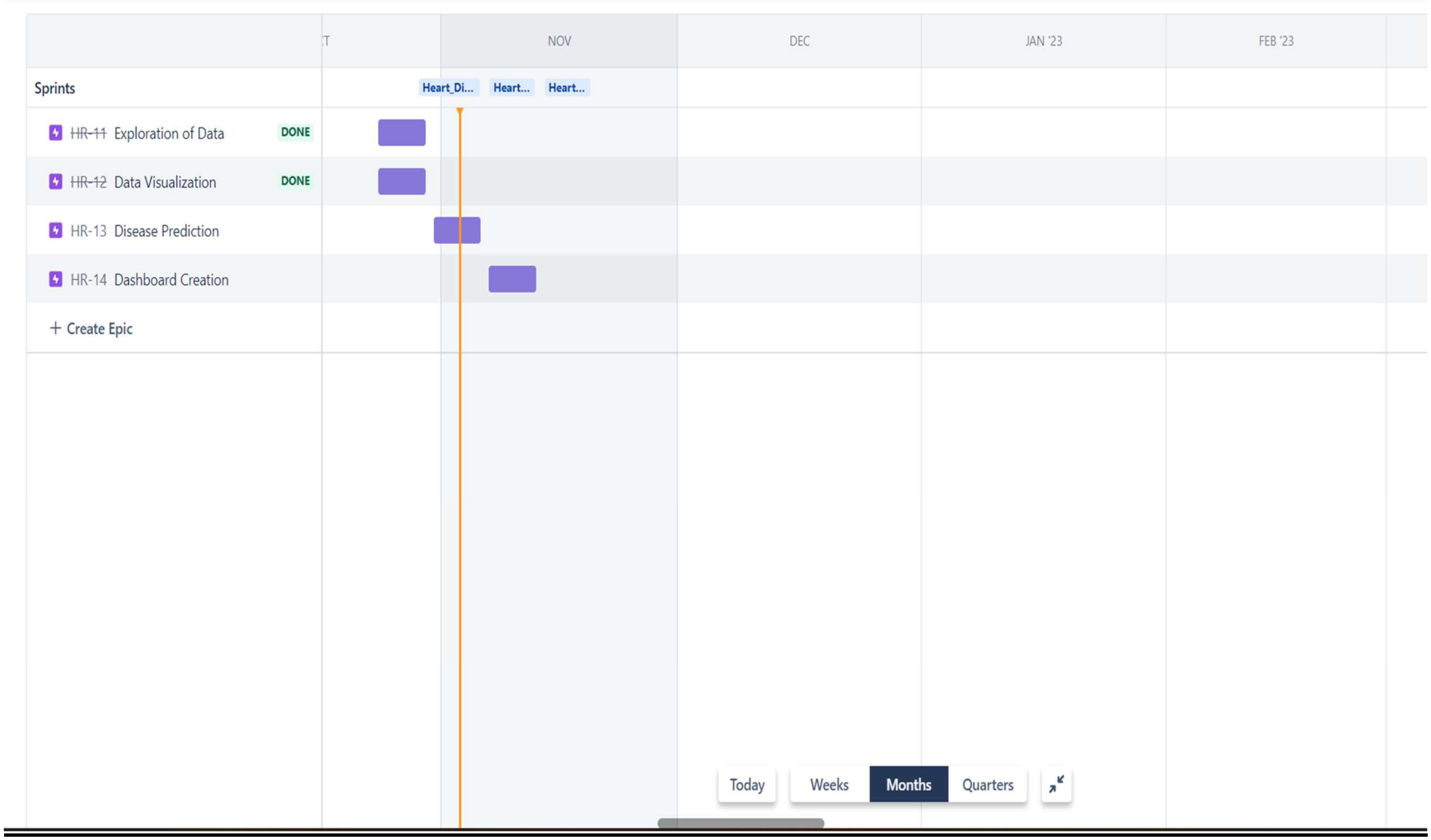
Project Design Phase-II	4			
Requirement Analysis	4.1	Identify the Functional and Non-Functional requirements	JOE ANTONY CELSHIYA JA	Completed
Customer Journey	4.2	Preparation of customer journey map to understand the user interactions & experiences with the application from the entry level to exit level	HANNA S	Completed
Data Flow Diagram and User stories	4.3	Generate Data flow diagram of the project	SRIPOORNIMA T	Completed
Technical Architecture	4.4	Develop effective technical architecture for the proposed solution	HARINI G	Completed
Project Planning Phase	5			
Milestones & Activity List	5.1	Prepare Milestone and Activity list of the project	SRIPOORNIMA T HANNA S HARINI G JOE ANTONY CELSHIYA JA	Completed
Sprint Plan	5.2	Prepare Sprint Delivery plan of the project	SRIPOORNIMA T HANNA S HARINI G JOE ANTONY CELSHIYA JA	Completed
Project Development Phase	6			
Delivery of Sprint-1	6.1	Implement the coding phase of Sprint-1	SRIPOORNIMA T HANNA S HARINI G JOE ANTONY CELSHIYA JA	In Progress
Delivery of Sprint-2	6.2	Implement the coding phase of Sprint-2	JOE ANTONY CELSHIYA JA	In Progress
Delivery of Sprint-3	6.3	Implement the coding phase of Sprint-3	SRIPOORNIMA T HANNA S	In Progress

			HARINI G JOE ANTONY CELSHIYA JA	
Delivery of Sprint-4	6.4	Implement the coding phase of Sprint-4	SRIPOORNIMA T HANNA S HARINI G JOE ANTONY CELSHIYA JA	In Progress

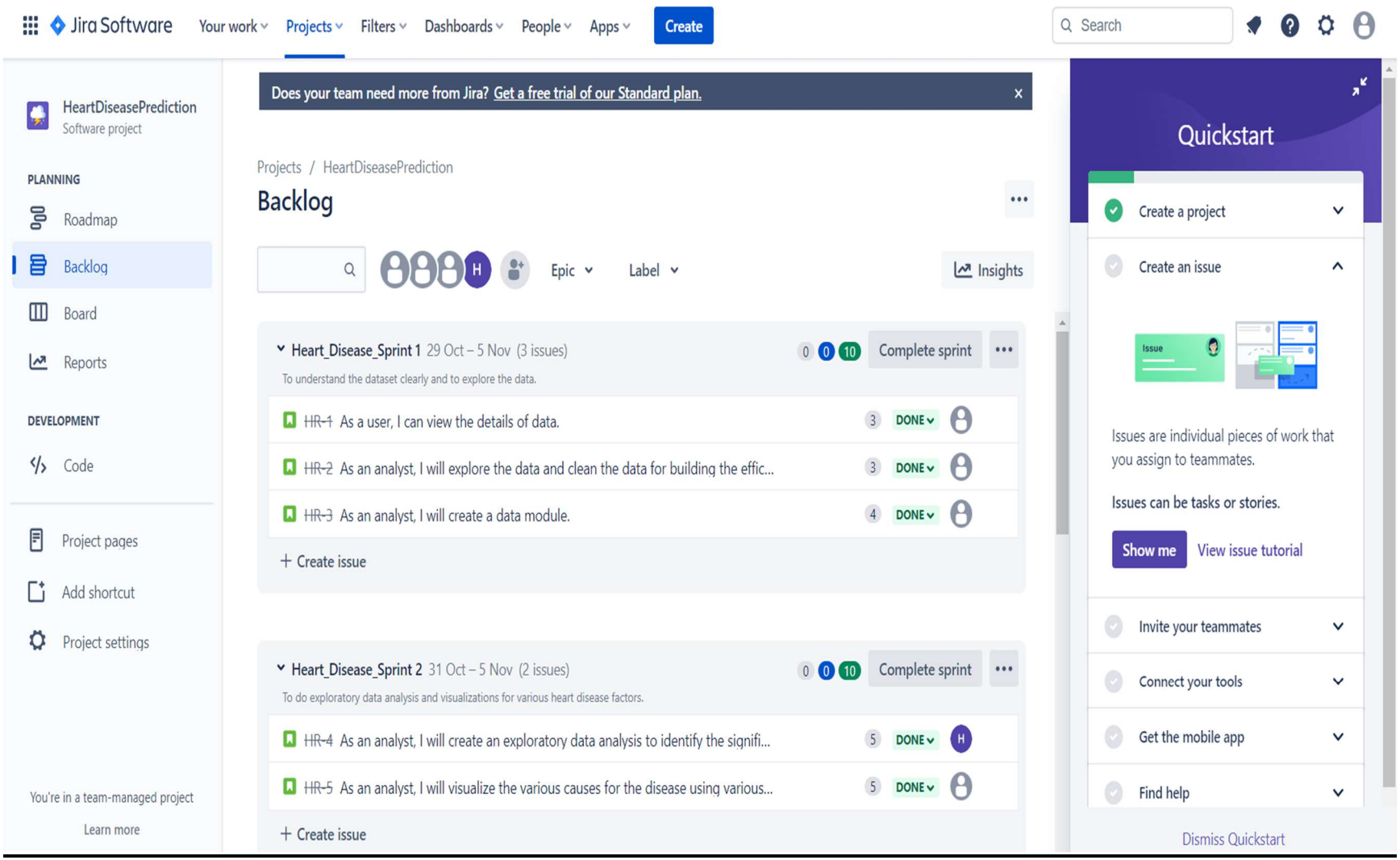
Milestone Timeline Chart:

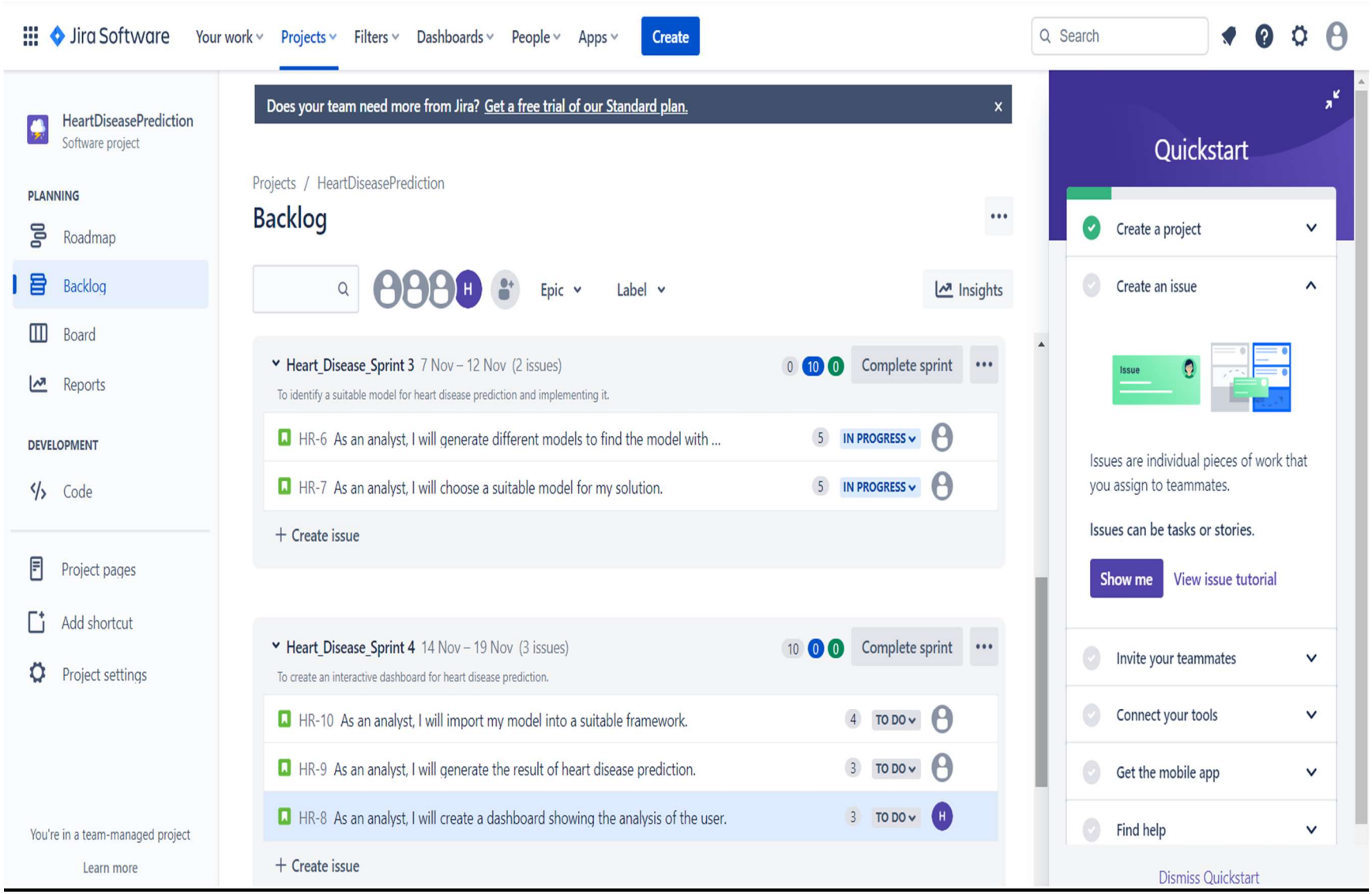
6.3 Reports from JIRA

Project Roadmap:

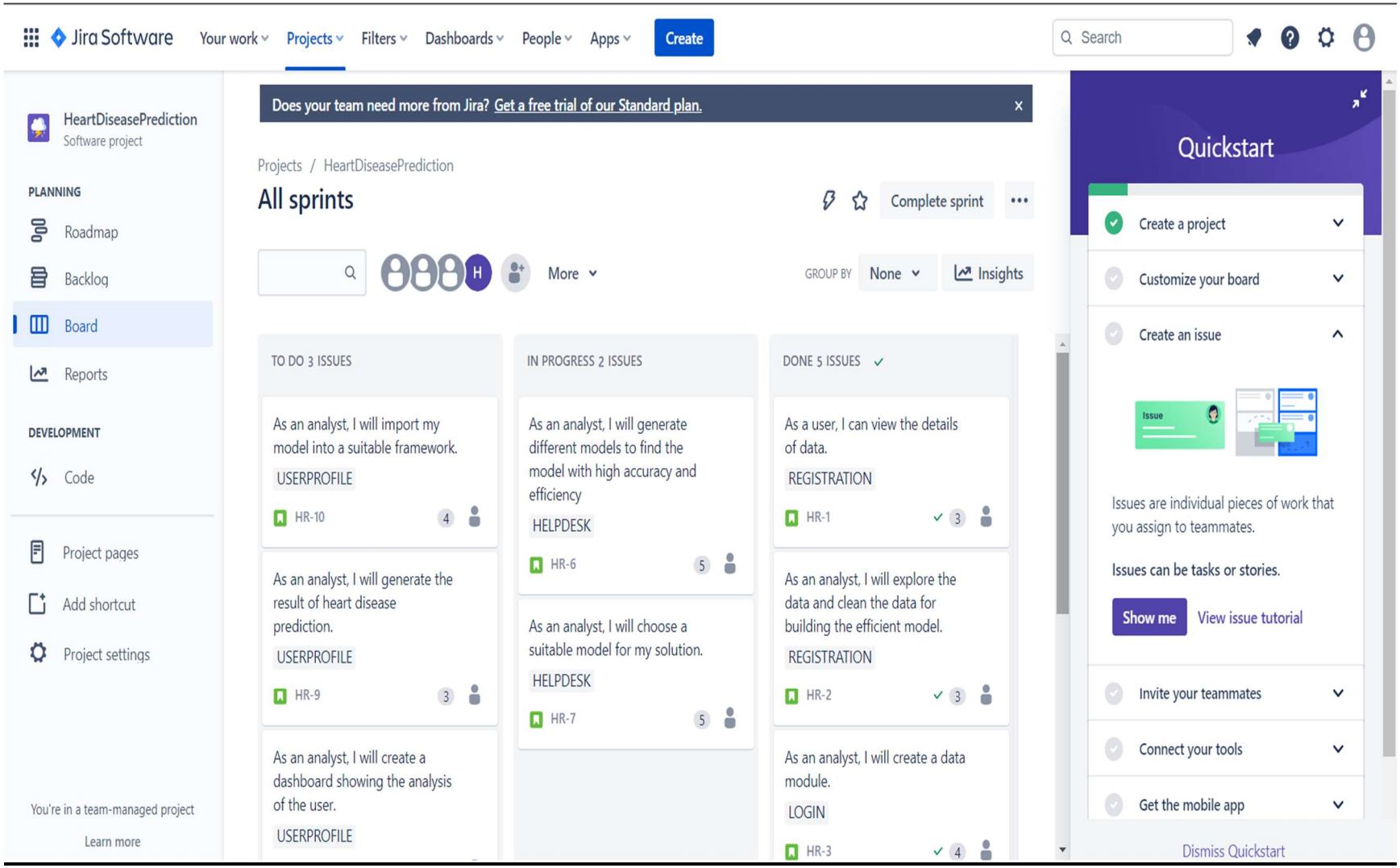


Project Backlog:




Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
HeartDiseasePrediction
Software project
PLANNING
Roadmap
Backlog
Board
Reports
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / HeartDiseasePrediction
Backlog
Epic
Label
Insights
Heart_Disease_Sprint 3 7 Nov – 12 Nov (2 issues)
0 10 0
Complete sprint
To identify a suitable model for heart disease prediction and implementing it.
HR-6 As an analyst, I will generate different models to find the model with ...
5
IN PROGRESS
HR-7 As an analyst, I will choose a suitable model for my solution.
5
IN PROGRESS
Create issue
Heart_Disease_Sprint 4 14 Nov – 19 Nov (3 issues)
10 0 0
Complete sprint
To create an interactive dashboard for heart disease prediction.
HR-10 As an analyst, I will import my model into a suitable framework.
4
TO DO
HR-9 As an analyst, I will generate the result of heart disease prediction.
3
TO DO
HR-8 As an analyst, I will create a dashboard showing the analysis of the user.
3
TO DO
Create issue
Quickstart
Create a project
Create an issue
Issue
Issues are individual pieces of work that you assign to teammates.
Issues can be tasks or stories.
Show me
View issue tutorial
Invite your teammates
Connect your tools
Get the mobile app
Find help
Dismiss Quickstart

Sprint Creation:



7. CODING & SOLUTIONING

7.1 Embedded_Dashboard.html

```
<html>

<head>

<title>Visualizing and Predicting HeartDisease</title>

<meta charset="utf-8">

<meta name="viewport" content="width=device-width, initial-scale=1">

<link href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.1/dist/css/bootstrap.min.css" rel="stylesheet">

<script src="https://cdn.jsdelivr.net/npm/bootstrap@5.2.1/dist/js/bootstrap.bundle.min.js">

</script>

</head>

<body>

<div class="container-fluid p-5 bg-info text-secondary text-center">

<h1> Visualizing and Predicting HeartDisease with an Interactive Dashboard</h1>

<p>

<iframe

src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&amp;pathRef=.my_folders%2FHeart_Disease_Prediction%2FD
ashboard%2BVisuals&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=false&amp;shareMode=embedde
```

d&action=view&mode=dashboard&subView=model00000183705ee369_00000003" width="1200" height="700"

frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen="">

</iframe>

</p>

</body>

</html>

Embedded_Story.html:

<!DOCTYPE html>

<html lang="en">

<head>

<title>Visualizing and Predicting HeartDisease</title>

<meta charset="utf-8">

<meta name="viewport" content="width=device-width, initial-scale=1">

<link href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.1/dist/css/bootstrap.min.css"

rel="stylesheet">

<script

```
src="https://cdn.jsdelivr.net/npm/bootstrap@5.2.1/dist/js/bootstrap.bundle.min.js"></script>
</head>
<body>
<div class="container-fluid p-5 bg-warning text-success text-center"> <h1>Visualizing and Predicting HeartDisease with an Interactive Dashboard</h1>
<p>
<iframe src="https://us1.ca.analytics.ibm.com/bi/?perspective=story&amp;pathRef=.public_folders%2FAssignment-1%2FStory&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=view&amp;sceneId=model000001847b70980f_00000002&amp;sceneTime=0" width="1200" height="600" frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen="">
</iframe>
</p>
</div>
</body>
</html>
```

7.2 Heart_Disease_Prediction_Model_Building.ipynb:

IMPORT DATASETS:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

READ DATASET:

```
dt=pd.read_csv("Heart_Disease_Prediction.csv")
```

ENCODING LABELS:

```
from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
label = le.fit_transform(dt['Heart_Disease'])
label
```

EXPLORATION:

```
dt.drop("Heart_Disease", axis=1, inplace=True)
dt["Heart_Disease"] = label
dt

dt.describe()

dt.isnull().sum()
```

TRAIN AND TEST DATA:

```
from sklearn.model_selection import train_test_split
x = dt.drop("Heart_Disease",axis=1)
y = dt["Heart_Disease"]
```

```
X_train,X_test,Y_train,Y_test = train_test_split(x,y,test_size=0.30,random_state=0)
```

NAÏVE BAYES CLASSIFIER:

```
from sklearn.naive_bayes import GaussianNB
nb = GaussianNB()
nb.fit(X_train,Y_train)
Y_pred_nb = nb.predict(X_test)

from sklearn.metrics import accuracy_score
score_nb = round(accuracy_score(Y_pred_nb,Y_test)*100,2)
print("The accuracy score achieved using Naive Bayes is: "+str(score_nb)+" %")

from sklearn.metrics import confusion_matrix
plt.figure(figsize=(8, 8))
CM=confusion_matrix(Y_test,Y_pred_nb)
sns.heatmap(CM, annot=True)
```

DECISIONTREE CLASSIFIER:

```
from sklearn.tree import DecisionTreeClassifier
decTree = DecisionTreeClassifier(max_depth=6, random_state=0)
decTree.fit(X_train,Y_train)
y_pred_decTree = decTree.predict(X_test)

import sklearn.metrics as metrics
score_dectree = round(accuracy_score(y_pred_decTree,Y_test)*100,2)
print("The accuracy score achieved using Decision Tree Classifier is: " , metrics.accuracy_score(Y_test,y_pred_decTree))

plt.figure(figsize=(8, 8))
```

```
CM=confusion_matrix(Y_test,y_pred_decTree)
sns.heatmap(CM, annot=True)
```

LOGISTIC REGRESSION:

```
from sklearn.linear_model import LogisticRegression
lr = LogisticRegression()
lr.fit(X_train,Y_train)
Y_pred_lr = lr.predict(X_test)

score_lr = round(accuracy_score(Y_pred_lr,Y_test)*100,2)
print("The accuracy score achieved using Logistic Regression is: "+str(score_lr)+" %")

plt.figure(figsize=(8, 8))
CM=confusion_matrix(Y_test,Y_pred_lr)
sns.heatmap(CM, annot=True)
```

RANDOM FOREST CLASSIFER:

```
from sklearn.ensemble import RandomForestClassifier
model=RandomForestClassifier()
model.fit(X_train,Y_train)
Y_pred_rfc=model.predict(X_test)

score_rfc = round(accuracy_score(Y_pred_rfc,Y_test)*100,2)
print("The accuracy score achieved using Random Forest Classifier is: "+str(score_rfc)+" %")

plt.figure(figsize=(8, 8))
CM=confusion_matrix(Y_test,Y_pred_rfc)
sns.heatmap(CM, annot=True)
```

SUPPORT VECTOR MACHINE:

```
from sklearn.svm import SVC
classifier = SVC(kernel='rbf', random_state = 1)
classifier.fit(X_train,Y_train)
Y_pred_svm = classifier.predict(X_test)

score_svm = round(accuracy_score(Y_pred_svm,Y_test)*100,2)
print("The accuracy score achieved using Support Vector Machine is: "+str(score_svm)+" %")

plt.figure(figsize=(8, 8))
CM=confusion_matrix(Y_test,Y_pred_svm)
sns.heatmap(CM, annot=True)
```

ACCURACY SCORES:

```
scores = [score_nb,score_dectree,score_lr,score_rfc,score_svm]
models = ["Naive Bayes Classifier","Decision Tree Classifier","Logistic Regression","Random Forest Classifier","Support Vector
Machine"]

for i in range(len(models)):
    print("The accuracy score achieved using "+models[i]+" is: "+str(scores[i])+" %")
```

8.TESTING

8.1 User Acceptance 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 Dashboard 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
BP	10	4	2	3	19
Cholesterol	9	3	3	0	15
EKG Results	2	3	0	1	6
ST Depression	1	2	4	7	14
Exercise Angina	9	6	1	0	16
Thallium	3	4	1	1	9
Totals	34	22	11	12	79

3. Test Case Analysis

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

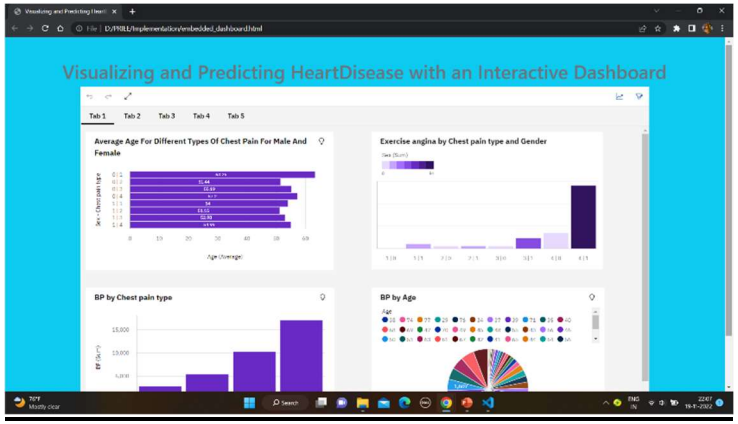
Section	Total Cases	Not Tested	Fail	Pass
BP	15	0	0	15
Cholesterol	20	0	0	20
EKG Results	3	0	0	3
ST Depression	4	0	0	4
Exercise Angina	2	0	0	2
Thallium	3	0	0	3

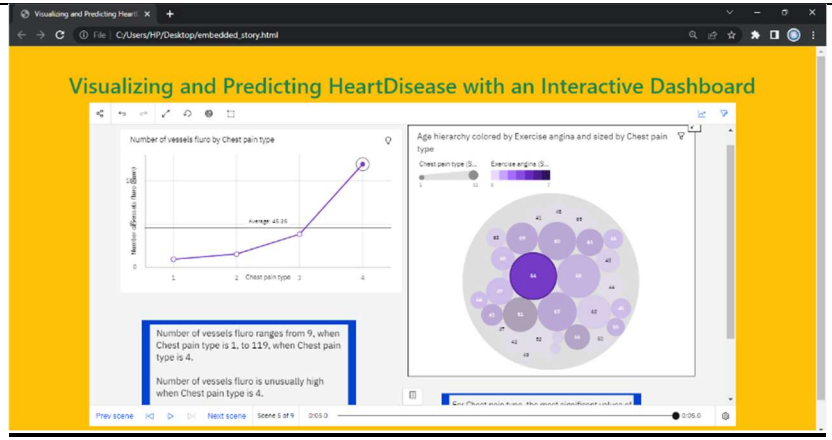
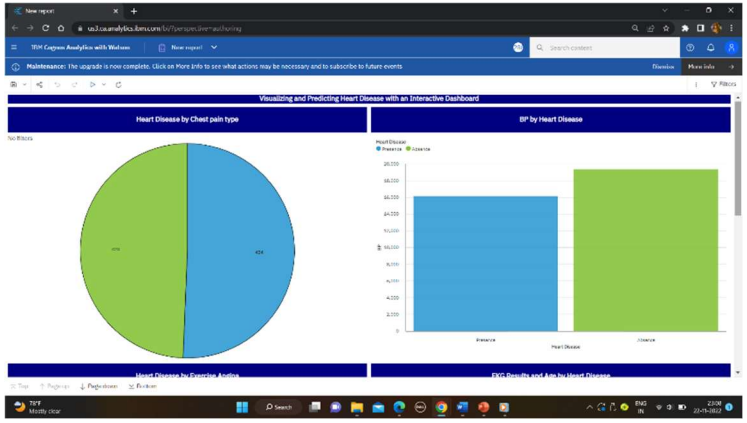
8.2 Testcases Report

Date	21-Nov-22
Team ID	PNT2022TMID17743
Project Name	Project - Visualizing and Predicting
Maximum Marks	4 marks

Test case ID	Feature Type	Component	Test Scenario	Pre-Requisite	Steps To Execute	Test Data	Expected Result	Actual Result	Status	Commnets	TC for Automation(Y/N)	BUG ID	Executed By
BP	Dashboard, Report and Story	IBM Cognos Analytics	Verify if the dataset is clean and preprocessed	A Preprocessed dataset	1. Upload the dataset 2. Explore the dataset 3. Visualize the features 4. Create Dashboard, Story and report	https://github.com/IBM-EPBL/IBM-Project-2206-1658466325/blob/main/Final%20Deliverables/Heart_Disease_Prediction.csv	Accurate Prediction	Working as expected	Pass		Yes		Sripoornima T,Hanna S
Cholesterol	Dashboard, Report and Story	IBM Cognos Analytics	Verify if the dataset is clean and preprocessed	A Preprocessed dataset	1. Upload the dataset 2. Explore the dataset 3. Visualize the features 4. Create Dashboard, Story and report	https://github.com/IBM-EPBL/IBM-Project-2206-1658466325/blob/main/Final%20Deliverables/Heart_Disease_Prediction.csv	Accurate Prediction	Working as expected	Pass	Steps are not clear to follow	Yes	BUG-1234	Hanna S, Harini G
EKG Results	Dashboard, Report and Story	IBM Cognos Analytics	Verify if the dataset is clean and preprocessed	A Preprocessed dataset	1. Upload the dataset 2. Explore the dataset 3. Visualize the features 4. Create Dashboard, Story and report	https://github.com/IBM-EPBL/IBM-Project-2206-1658466325/blob/main/Final%20Deliverables/Heart_Disease_Prediction.csv	Accurate Prediction	Working as expected	Pass		Yes		Joe Antony Celshiya J A,Harini G
ST Depression	Dashboard, Report and Story	IBM Cognos Analytics	Verify if the dataset is clean and preprocessed	A Preprocessed dataset	1. Upload the dataset 2. Explore the dataset 3. Visualize the features 4. Create Dashboard, Story and report	https://github.com/IBM-EPBL/IBM-Project-2206-1658466325/blob/main/Final%20Deliverables/Heart_Disease_Prediction.csv	Accurate Prediction	Working as expected	Pass		Yes		Sripoornima T,Hanna S
Exercise Angina	Dashboard, Report and Story	IBM Cognos Analytics	Verify if the dataset is clean and preprocessed	A Preprocessed dataset	1. Upload the dataset 2. Explore the dataset 3. Visualize the features 4. Create Dashboard, Story and report	https://github.com/IBM-EPBL/IBM-Project-2206-1658466325/blob/main/Final%20Deliverables/Heart_Disease_Prediction.csv	Accurate Prediction	Working as expected	Pass		Yes		Sripoornima T,Hanna S
Thallium	Dashboard, Report and Story	IBM Cognos Analytics	Verify if the dataset is clean and preprocessed	A Preprocessed dataset	1. Upload the dataset 2. Explore the dataset 3. Visualize the features 4. Create Dashboard, Story and report	https://github.com/IBM-EPBL/IBM-Project-2206-1658466325/blob/main/Final%20Deliverables/Heart_Disease_Prediction.csv	Accurate Prediction	Working as expected	Pass		Yes		Joe Antony Celshiya J A,Harini G

8.3 PerformanceTesting

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	No of Visualizations / Graphs – 20 Visualizations in 5 Dashboard tabs (4 Visualizations in each Dashboard tab) 
2.	Data Responsiveness	The visualizations show the relationship between the features. It shows how each factor affects the target “Heart Disease”.
3.	Utilization of Data Filters	Filters are used to aggregate or group common data.
4.	Effective User Story	No of Scene Added – 14 Stories with 2 visualizations each.

		
5.	Descriptive Reports	No of Visualizations / Graphs – 4 reports with 2/4 Visualizations each. 

9.RESULTS

9.1 Performance Metrics

Accuracy Score Of Machine Learning Models

+ Code + Markdown

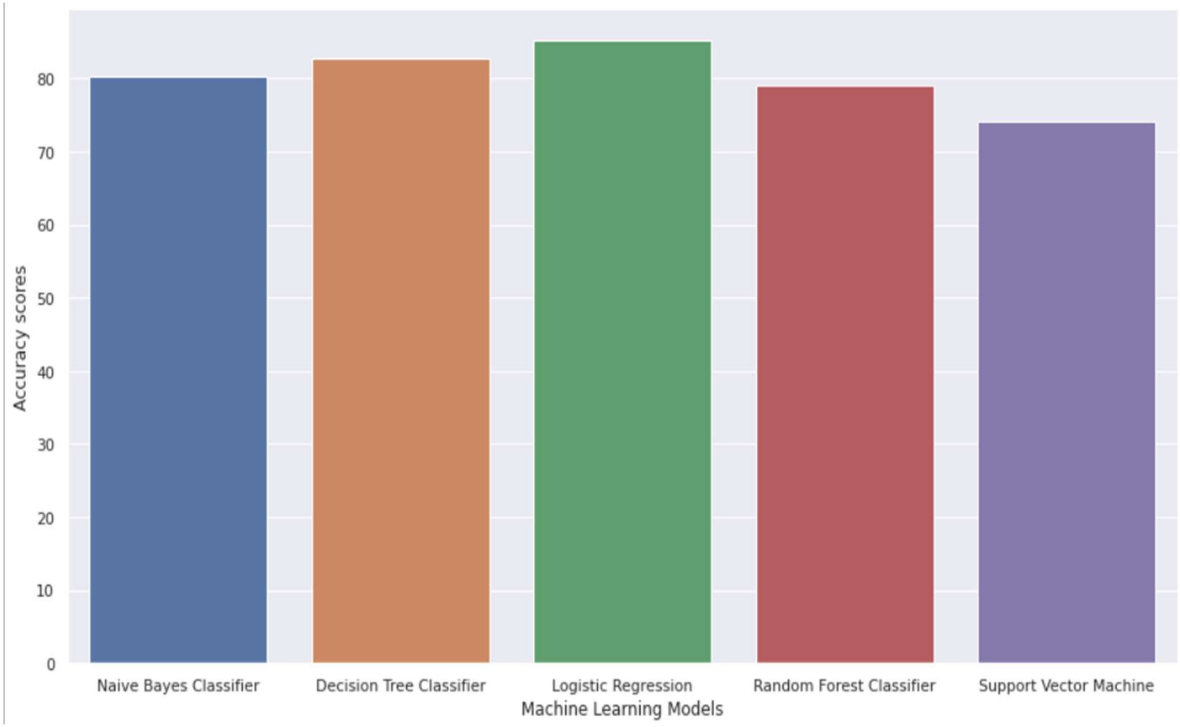
```
scores = [score_nb,score_dectree,score_lr,score_rfc,score_svm]
models = ["Naive Bayes Classifier","Decision Tree Classifier","Logistic Regression","Random Forest Classifier","Support Vector Machine"]

for i in range(len(models)):
    print("The accuracy score achieved using "+models[i]+" is: "+str(scores[i])+" %")
```

[38]

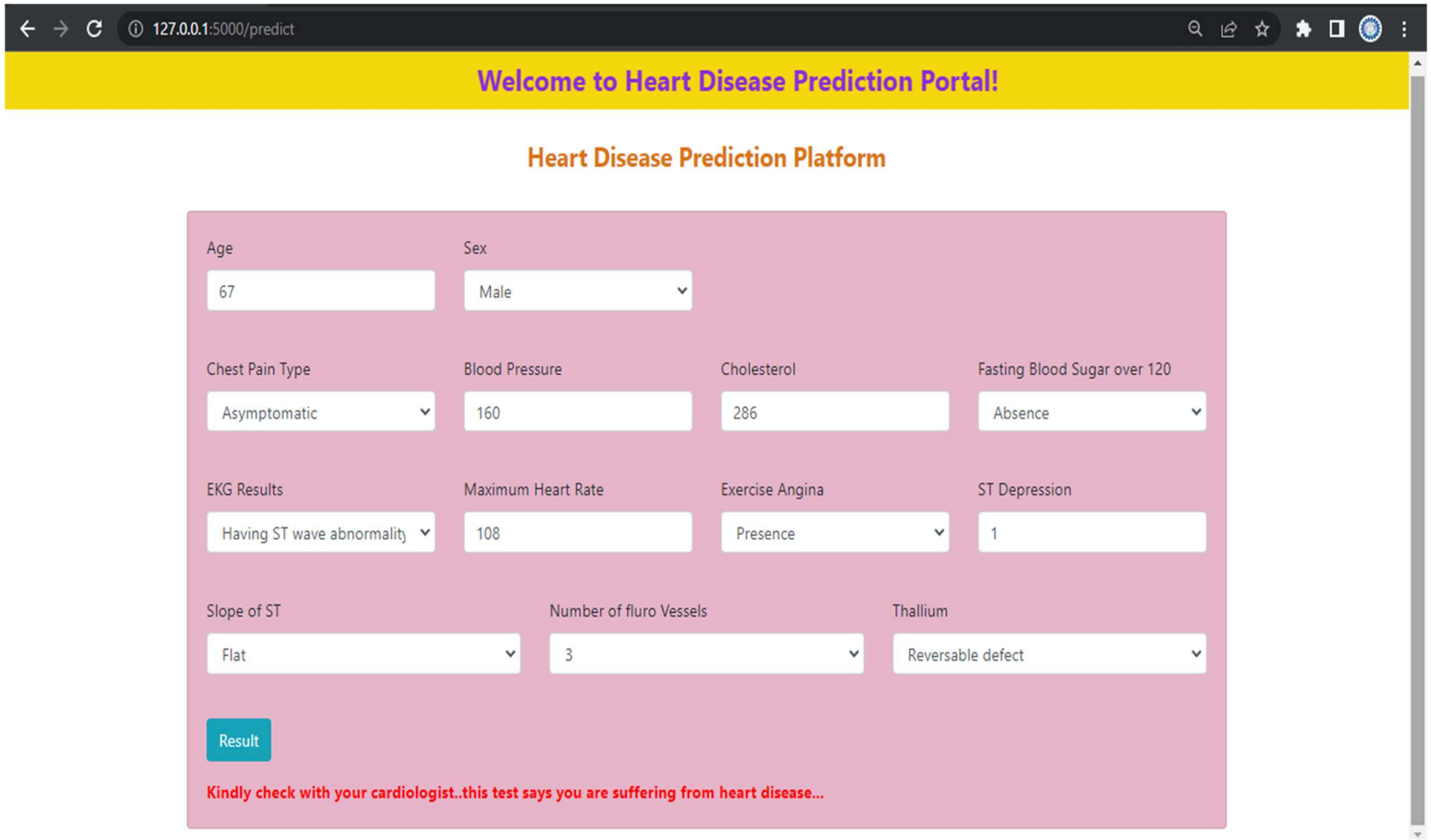
```
The accuracy score achieved using Naive Bayes Classifier is: 80.25 %
The accuracy score achieved using Decision Tree Classifier is: 82.72 %
The accuracy score achieved using Logistic Regression is: 85.19 %
The accuracy score achieved using Random Forest Classifier is: 79.01 %
The accuracy score achieved using Support Vector Machine is: 74.07 %
```

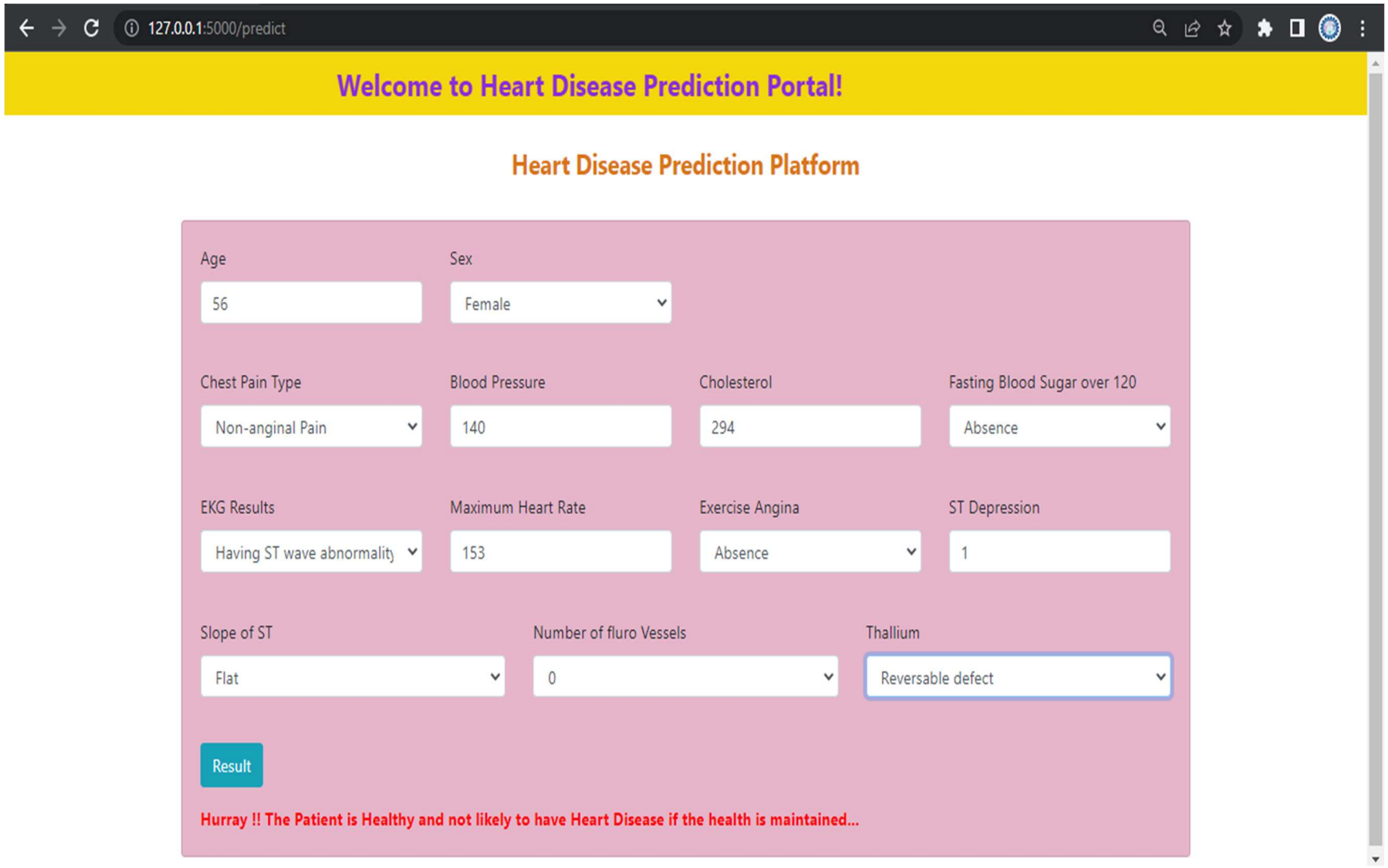
ML Model	NaïveBayes Classifier	DecisionTree Classifier	Logistic Regression	RandomForest Classifier	SupportVector Machine
Accuracy	80.25%	82.72%	85.19%	79.01%	74.07%

Comparative Analysis:



9.2 Result Screenshots




127.0.0.1:5000/predict
Welcome to Heart Disease Prediction Portal!
Heart Disease Prediction Platform
Age
56
Sex
Female
Chest Pain Type
Non-anginal Pain
Blood Pressure
140
Cholesterol
294
Fasting Blood Sugar over 120
Absence
EKG Results
Having ST wave abnormality
Maximum Heart Rate
153
Exercise Angina
Absence
ST Depression
1
Slope of ST
Flat
Number of fluro Vessels
0
Thallium
Reversable defect
Result
Hurray !! The Patient is Healthy and not likely to have Heart Disease if the health is maintained...

10. ADVANTAGES & DISADVANTAGES

10.1 Advantages

Following are some of the advantages of the web app,

1) User can check for heart disease prediction at any point of time.
2) User can get immediate results for prediction of heart disease.
3) The System helps user to know about their current health conditions.
4) The system alerts user if they have heart disease.

10.2 Disadvantages

Following are some of the disadvantages of the web app,

1) Sometimes the results being predicted are not accurate.
2) Thus, better decisions cannot be made.
3) The system does not provide deeper insights about the disease rather it just predicts presence or absence of disease.

11. CONCLUSION

Heart stroke and vascular disease are the major cause of disability and premature death. In this work, the heart diseases are predicted by considering major factors. The factors are analyzed and visualized using IBM Cognos Analytics. Various Machine Learning algorithms were applied to the heart disease dataset and we obtained high accuracy using Decision Tree Classifier. Using this model, we predict whether a person has heart disease or not.

12. FUTURE SCOPE

The heart disease prediction accuracy can be improved by tuning the parameters of the dataset. It can be extended to an independent mobile app by predicting the heart disease and provide solutions to keep their health fit. Based on the prediction results, it may provide online consultation with doctors. Then, it can be deployed into an android app to make heart disease prediction by our own. It can also be deployed to hospitals to make a quick prediction in emergency situation.

13. APPENDIX

13.1 Source Code

flask_app.py

```
import numpy as np
from flask import Flask,render_template,request
import pickle
app = Flask(__name__)
model = pickle.load(open('decmodel.pkl', 'rb'))
@app.route('/')
def home():
    return render_template('index.html')
@app.route('/predict',methods=['POST'])
def predict():
    int_features = [float(x) for x in request.form.values()]
    features = [np.array(int_features)]
    prediction = model.predict(features)
    output = prediction
```

```
        if output == 1:
            return render_template('index.html',result="Kindly check with your cardiologist..this test says you are suffering from heart disease...")
        else:
            return render_template('index.html',result="Hurray !! The Patient is Healthy and not likely to have Heart Disease if the health is maintained...")
        if __name__ == "__main__":
        app.run(debug=True)
```

model.py

```
#!/usr/bin/env python
# coding: utf-8
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
import pickle
import pandas as pd
import numpy as np
dt=pd.read_csv("F:/IBM/Heart_Disease_Prediction.csv")
```

```
from sklearn.preprocessing import LabelEncoder

le = LabelEncoder()

label = le.fit_transform(dt['Heart_Disease'])

label

dt.drop("Heart_Disease", axis=1, inplace=True)

dt["Heart_Disease"] = label

dt

x = dt.drop("Heart_Disease",axis=1)

y = dt["Heart_Disease"]

X_train,X_test,Y_train,Y_test = train_test_split(x,y,test_size=0.30,random_state=0)

decTree = DecisionTreeClassifier(max_depth=6, random_state=0) #Decision Tree Classifier

decTree.fit(X_train,Y_train)

pickle.dump(decTree, open('decmodel.pkl', 'wb'))
```

index.html:

```
<html>

<head>

    <title>Heart Disease Prediction</title>
```

```
    <link rel="stylesheet" href="https://stackpath.bootstrapcdn.com/bootstrap/4.5.2/css/bootstrap.min.css" integrity="sha384-
JcKb8q3iqJ61gNV9KGb8thSsNjpSL0n8PARn9HuZOnIxN0hoP+VmmDGMN5t9UJ0Z" crossorigin="anonymous">
    <script src="https://code.jquery.com/jquery-3.5.1.slim.min.js" integrity="sha384-
DfXdz2htPH0lsSSs5nCTpuj/zy4C+OGpamoFVy38MVBnE+IbbVYUew+OrCXaRkfj" crossorigin="anonymous"></script>
    <script src="https://cdn.jsdelivr.net/npm/popper.js@1.16.1/dist/umd/popper.min.js" integrity="sha384-
9/reFTGAW83EW2RDu2S0VKaIzap3H66lZH81PoYlFhbGU+6BZp6G7niu735Sk7lN" crossorigin="anonymous"></script>
    <script src="https://stackpath.bootstrapcdn.com/bootstrap/4.5.2/js/bootstrap.min.js" integrity="sha384-
B4gt1jrGC7Jh4AgTPSdUtOBvfO8shuf57BaghqFfPlYxofvL8/KUEfYiJOMMV+rV" crossorigin="anonymous"></script>
  </head>
  <body background-image="cardiac_blues.jpg" background-repeat="no-repeat" background-size="cover">
  <script src="https://code.jquery.com/jquery-3.5.1.slim.min.js" integrity="sha384-
DfXdz2htPH0lsSSs5nCTpuj/zy4C+OGpamoFVy38MVBnE+IbbVYUew+OrCXaRkfj" crossorigin="anonymous"></script>
  <script src="https://cdn.jsdelivr.net/npm/popper.js@1.16.1/dist/umd/popper.min.js" integrity="sha384-
9/reFTGAW83EW2RDu2S0VKaIzap3H66lZH81PoYlFhbGU+6BZp6G7niu735Sk7lN" crossorigin="anonymous"></script>
  <script src="https://stackpath.bootstrapcdn.com/bootstrap/4.5.2/js/bootstrap.min.js" integrity="sha384-
B4gt1jrGC7Jh4AgTPSdUtOBvfO8shuf57BaghqFfPlYxofvL8/KUEfYiJOMMV+rV" crossorigin="anonymous"></script>
```

```
<nav class="navbar navbar-dark" style="background-color: rgb(243, 216, 13);">
<marquee style="font-size:20pt;color:blueviolet"behavior="alternate" direction="left"><b>Welcome to Heart Disease Prediction Portal!</b></marquee>
</nav>
<div class="container">
<br>
<form action = "{{url_for('predict')}}" method ="POST" >
<fieldset>
<legend style="color: rgb(219, 117, 21); text-align: center;"><b>Heart Disease Prediction Platform</b></legend><br>
<div class="card card-body" style="background-color: rgba(224, 162, 188, 0.795);">
<div class="form-group row">
<div class="col-sm-3">
<label for="age">Age</label>
<input type="number" class="form-control" id="age" name="age" min="18" max="100" required>
</div>
<div class="col-sm-3">
<label for="sex">Sex</label>
```

```
<select class="form-control" id="sex" name="sex" required>
<option disabled selected value> Choose Field </option>
<option value = "0">Female</option>
<option value = "1">Male</option>
</select>
</div>
</div>
<br>
<div class="form-group row">
<div class="col-sm">
<label for="cp">Chest Pain Type</label>
<select class="form-control" id="cp" name = "cp" required>
<option disabled selected value> Choose Field </option>
<option value = "1">Typical Angina</option>
<option value = "2">Atypical Angina</option>
<option value = "3">Non-anginal Pain</option>
<option value = "4">Asymptomatic</option>
```

```
</select>
</div>
<div class="col-sm">
<label for="trestbps">Blood Pressure</label>
<input type="number" class="form-control" id="trestbps" name="trestbps" min="60" max="200" required>
</div>
<div class="col-sm">
<label for="chol">Cholesterol</label>
<input type="number" class="form-control" id="chol" name="chol" min="100" max="600" required>
</div>
<div class="col-sm">
<label for="fbs">Fasting Blood Sugar over 120</label>
<select class="form-control" id="fbs" name="fbs" required>
<option disabled selected value> Choose Field </option>
<option value = "0">Absence</option>
<option value = "1">Presence</option>
</select>
```

```
</div>
</div>
 <br>
<div class="form-group row">
<div class="col-sm">
<label for="restecg">EKG Results </label>
<select class="form-control" id="restecg" name="restecg" required>
<option disabled selected value> Choose Field </option>
<option value = "0">Normal </option>
<option value = "1">Having ST wave abnormality </option>
<option value = "2">HAving High abnormality</option>
</select>
</div>
<div class="col-sm">
<label for="thalach">Maximum Heart Rate</label>
<input type="number" class="form-control" id="thalach" name="thalach" min="100" max="200" required>
</div>
```

```
<div class="col-sm">
<label for="exang">Exercise Angina </label>
<select class="form-control" id="exang" name="exang" required>
<option disabled selected value> Choose Field </option>
<option value = "0">Absence</option>
<option value = "1">Presence</option>
</select>
</div>
<div class="col-sm">
<label for="oldpeak">ST Depression</label>
<input type="number" step="any" class="form-control" id="oldpeak" name="oldpeak" min="0" max="7" required>
</div>
</div>
<br>
<div class="form-group row">
<div class="col-sm">
<label for="slope">Slope of ST</label>
```

```
<select class="form-control" id="slope" name="slope" required>
<option disabled selected value> Choose Field </option>
<option value = "1">Upslope</option>
<option value = "2">Flat</option>
<option value = "3">Downslope</option>
</select>
</div>
<div class="col-sm">
<label for="ca">Number of fluro Vessels</label>
<select class="form-control" id="ca" name = "ca" required>
<option disabled selected value> Choose Field </option>
<option value = "0">0</option>
<option value = "1">1</option>
<option value = "2">2</option>
<option value = "3">3</option>
</select>
</div>
```

```
<div class="col-sm">
<label for="thal">Thallium</label>
<select class="form-control" id="thal" name = "thal" required>
<option disabled selected value> Choose Field </option>
<option value = "3">Normal</option>
<option value = "6">Fixed defect</option>
<option value = "7">Reversable defect</option>
</select>
</div>
</div>
<br>
<div class="form-group">
<input class="btn btn-info" type="submit" value="Result">
</div>
<div id ="result">
<strong style="color:red">{{result}}</strong>
</div>
```

```
</div>

</fieldset>

</form>

</div>

</body>

</html>
```

13.2 GitHub & Project Demo Link

https://github.com/IBM-EPBL/IBM-Project-2206-58466325/

https://drive.google.com/file/d/1GGLIPcFnEtTiedZq5KAxKsk6bqvx6K1J/view?usp=sharing