PROJECT REPORT

VISUALIZING AND PREDICTING HEART DISEASES WITH

AN INTERACTIVE DASHBOARD

Team ID – PNT2022TMID34115

- Submitted by

TEAM MEMBERS NAME	REGISTER NUMBER
HEMANTH KUMAR P	713919104012
VASANTH E	713919104040
PREETHI S	713919104021
SNEGA R	713919104030

Project Report

1. **INTRODUCTION**
    1. Project Overview
    2. Purpose

2. **LITERATURE SURVEY**
    1. Existing problem
    2. References
    3. Problem Statement Definition

3. **IDEATION & PROPOSED SOLUTION**
    1. Empathy Map Canvas
    2. Ideation & Brainstorming
    3. Proposed Solution
    4. Problem Solution fit

4. **REQUIREMENT ANALYSIS**
    1. Functional requirement
    2. Non-Functional requirements

5. **PROJECT DESIGN**
    1. Data Flow Diagrams
    2. Solution & Technical Architecture
    3. User Stories

6. **PROJECT PLANNING & SCHEDULING**
    1. Sprint Planning & Estimation
    2. Sprint Delivery Schedule
    3. Reports from JIRA

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

    1. Feature 1
    2. Feature 2
    3. Database Schema (if Applicable)

8. **TESTING**

    1. Test Cases
    2. User Acceptance Testing

9. **RESULTS**

    1. Performance Metrics

10. **ADVANTAGES & DISADVANTAGES**

11. **CONCLUSION**

12. **FUTURE SCOPE**

13. **APPEND**

INTRODUCTION

1.1 PROJECT OVERVIEW

The terms "heart disease" and "cardiovascular disease" are frequently used interchangeably. Heart disease is a general term that covers a wide range of heart related medical conditions. The irregular health state that directly affects the heart and all of its components is characterized by these medical conditions.

In order to forecast cardiac disease, this study discusses various data mining, big data, and machine learning techniques. Building an important model for the medical system to forecast heart disease or cardiovascular illness requires the use of data mining and machine learning. Our application helps the user in finding out if they have heart disease or not.

They can find out by entering details such as their heart rate, cholesterol, blood pressure etc. A dashboard is also attached along with the results for better understanding where they can compare their blood pressure and similar metrics with other users. This project focuses on Random Forest Classifier. The accuracy of our project is 87% for which is better than most other systems in terms of achieving accuracy quickly.

1.2 PURPOSE

This project's goal is to determine, depending on the patient's medical characteristics—such as gender, age, chest pain, and fasting blood sugar level, etc.—whether they are likely to be diagnosed with any cardiovascular heart illnesses. The leading cause of death in the developedworld is heart disease. Heart disease cases are rising quickly every day, thus it's crucial and worrisome to predict any potential illnesses in advance. This diagnosis is a challenging task that requires accuracy and efficiency.

Therefore, there needs to be work done to help prevent the risks of having a heart attack or It can assist inidentifying disease with less medical tests and effective therapies, so that patients can be treated appropriately. It can identify anyone who is experiencing any heart disease symptoms,such as chest pain or high blood pressure.

Around the world, machine learning is applied in many different fields. There is no exception in the healthcare sector. Machine learning may be crucial in determining whether locomotors disorders, heart illnesses, and other conditions are present or absent. If foreseen well in advance, such information can offer valuable insights to doctors, who can then customize their diagnosis and course of care for each patient.

LITERATURE SURVEY

2.1 EXISTING PROBLEM

A quiet significant amount of works related to the diagnosis of heart disease using Machine Learning algorithms have been made. An efficient heart disease prediction has beenmade by using various algorithms some of them include Logistic Regression, KNN, Random Forest Classifier etc. It can be seen in results that each algorithm has its strength to register the defined objectives.

The model incorporating IHDPS had the ability to calculate the decision boundary using the previous and new model of machine learning and deep learning. It facilitated the important and the most basic factors/knowledge such as family history connected with any heart disease. But the accuracy that was obtained in such IHDPS model was far more less than the new upcoming model such as detecting coronary heart disease using artificial neural networkand other algorithms of machine and deep learning.

2.2 REFERENCES

[1] Ali, Liaqat, et al, "An optimized stacked support vector machines based expert system for the effective prediction of heart failure." IEEE Access 7 (2019): 54007-54014. www.ijcrt.org © 2020 IJCRT | Volume 8, Issue 8 August 2020 | ISSN: 2320-2882 IJCRT2008170 International Journal of Creative Research Thoughts (IJCRT) www.ijcrt.org 1606

[2] Mohan, Senthilkumar, Chandrasegar Thirumalai, and Gautam Srivastava, "Effective heart disease prediction using hybrid machine learning techniques." IEEE Access 7 (2019): 81542-81554.

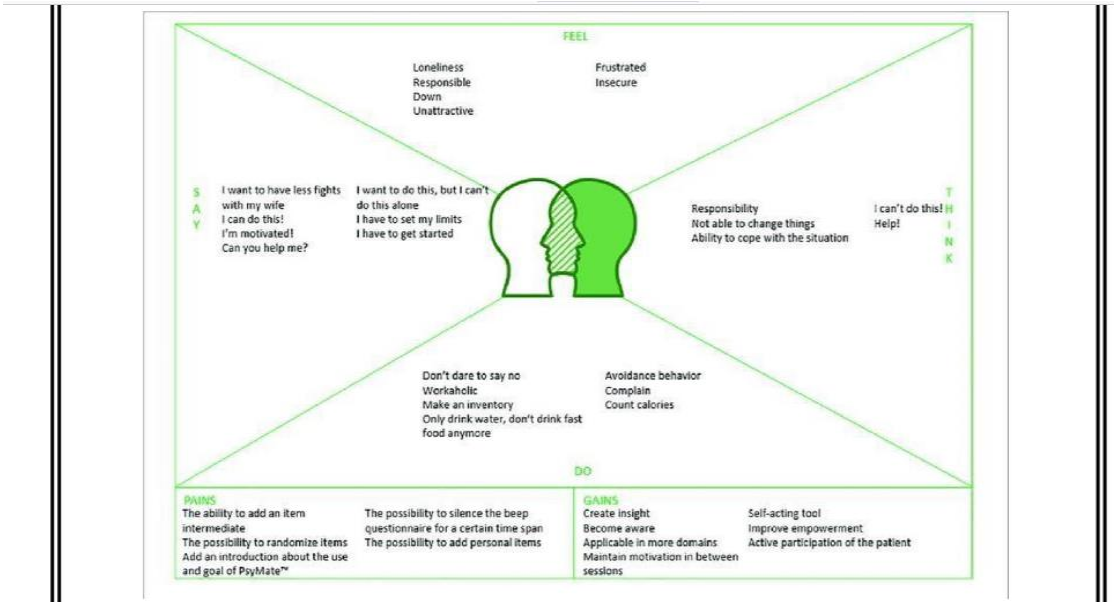
[3] Purushottam, Kanak Saxena and Richa Sharma, "Efficient heart disease prediction system." Procedia Computer Science 85 (2016): 962-969.

[4] Singh, Yeshvendra K., Nikhil Sinha, and Sanjay K. Singh, "Heart Disease Prediction System Using Random Forest", International Conference on Advances in Computing and Data Sciences. Springer, Singapore, 2016.

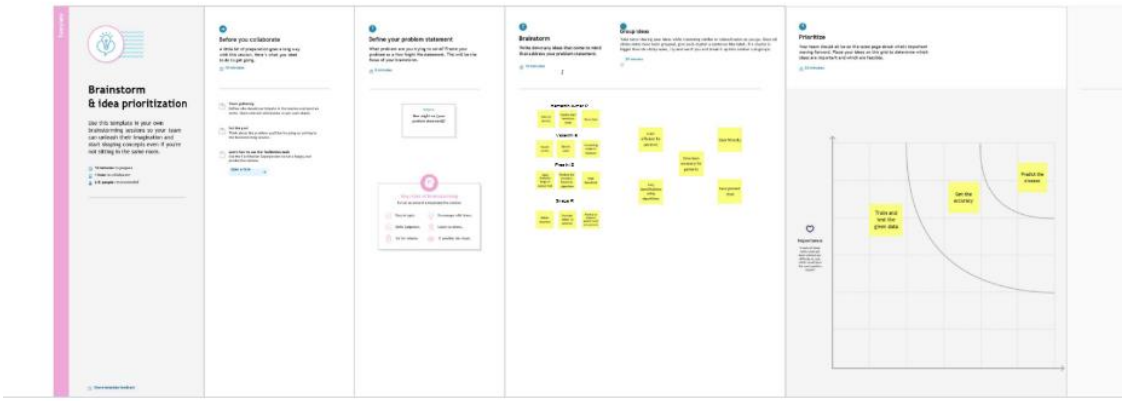
[5] Heart Disease Prediction using Exploratory Data Analysis R. Indrakumari, T.Poongodi, Soumya Ranjan Jena

IDEATION & PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS



3.2 IDEATION AND BRAINSTORMING



3.3PROBLEM STATEMENT:

Heart disease can be managed effectively with a combination of lifestyle changes, medicine and, in some cases, surgery. With the right treatment, the symptoms of heart disease can be reduced and the functioning of the heart improved. The predicted results can be used to prevent and thus reduce cost for surgical treatment and other expensive. The overall objective of our work will be to predict accurately with few tests and attributes the presence of heart disease. Attributes considered form the primary basis for tests and give accurate results more or less. Many more input attributes can be taken but our goal is to predict with few attributes and faster efficiency the risk of having heart disease. Decisions are often made based on doctors' intuition and experience rather than on the knowledge rich data hidden in the data set and databases. This practice leads to unwanted biases, errors and excessive medical costs which affects the quality of service provided to patients. There is a wealth of data available within the healthcare systems. However, there is a lack of effective analysis tools to discover hidden relationships and trends in the data for African genres

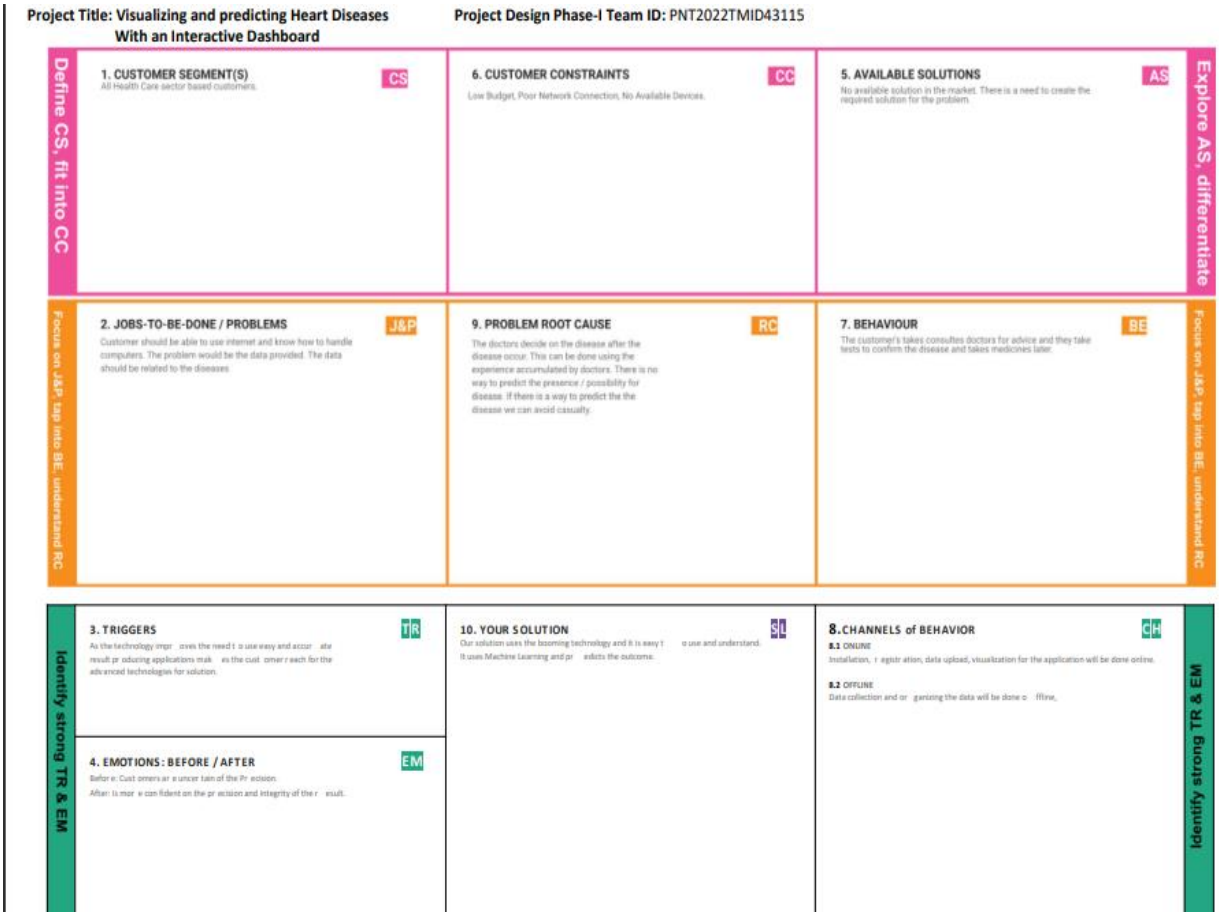
I am	The person to analyze the data on the record of health data using data analytics.
I'm trying to	Use the recent technologies to predict the Analytics for Heart disease Data.
But	I am unaware of the existing technology that can help me a lot to analyze the health care data and I don't know to use the correct technology.
Because	I don't want to waste the cost and time.
Which makes me feel	I want a best accuracy which can analyze the health care data so that the people can move with their necessary treatments.

3 . 4 PROPOSED SOLUTION

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	To develop an interactive dashboard to predict the heart disease accurately with few tests and attributes the presence of heart disease.
2.	Idea / Solution description	Analyzing data and identifying the heart disease using Cognos analysis.
3.	Novelty / Uniqueness	Hoping to achieve maximum accuracy to provide prior treatment to the patients and reduce the fatality rate.
4.	Social Impact / Customer Satisfaction	Saving lives, User friendly interactive dashboard. Reduces the exorbitant medical cost of the patients. Reduces the biases and mistakes caused by the decisions of doctors based on their intuitions and experiences
5.	Business Model (Revenue Model)	➢ Data security. ➢ Easy to use. ➢ Constant updates according to necessity.
6.	Scalability of the Solution	➢ Can be used in any platform (Windows, mac, etc.,) ➢ Adding new feature doesn't affect the performance of the system. ➢ Scalable dataset.

3.5 PROBLEM SOLUTION FIT

The Problem-Solution Fit simply means that we have found a problem with our customer and that the solution we have realized for it actually solves the customer's problem. It helps entrepreneurs, marketers and corporate innovators identify behavioral patterns and recognize what would work and why. The purpose is to solve complex problems in a way that fits the state of your customers and succeed faster and increase your solution adoption by tapping into existing mediums and channels of behavior



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.

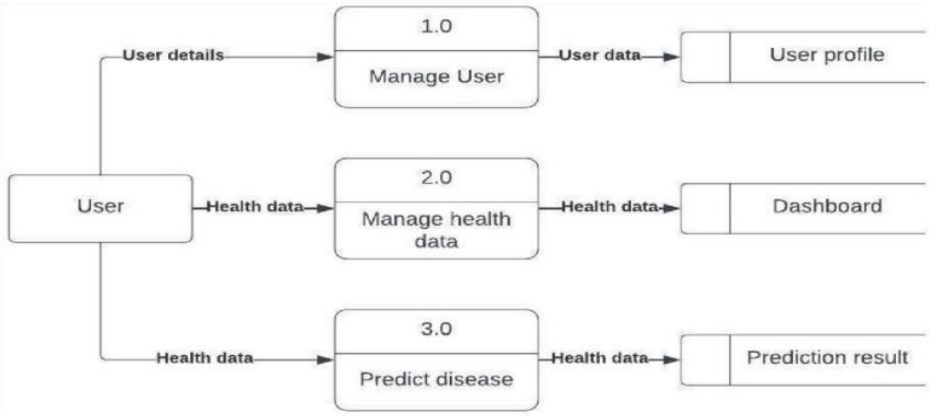
NFR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The application will have a simple and user-friendly graphical interface. Users will be ableto understand and use all the features of the application easily. Any action has to be performed with just a few clicks
NFR-2	**Security**	For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

PROJECT DESIGN

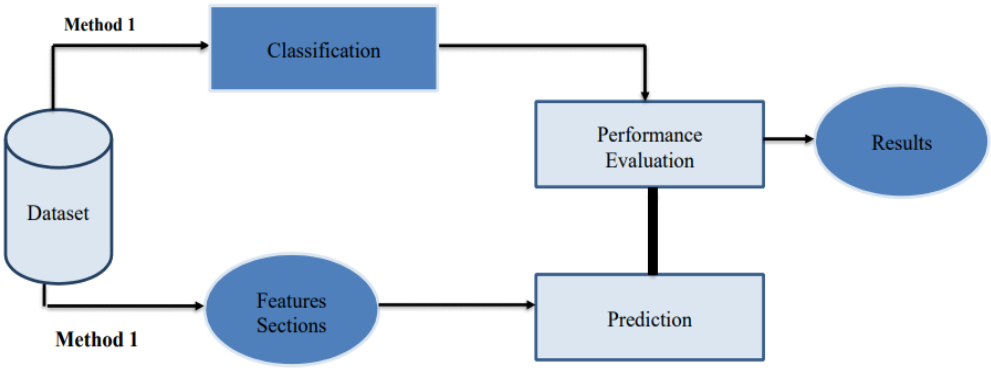
Data Flow Diagrams, Solution & Technical Architecture

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 right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.



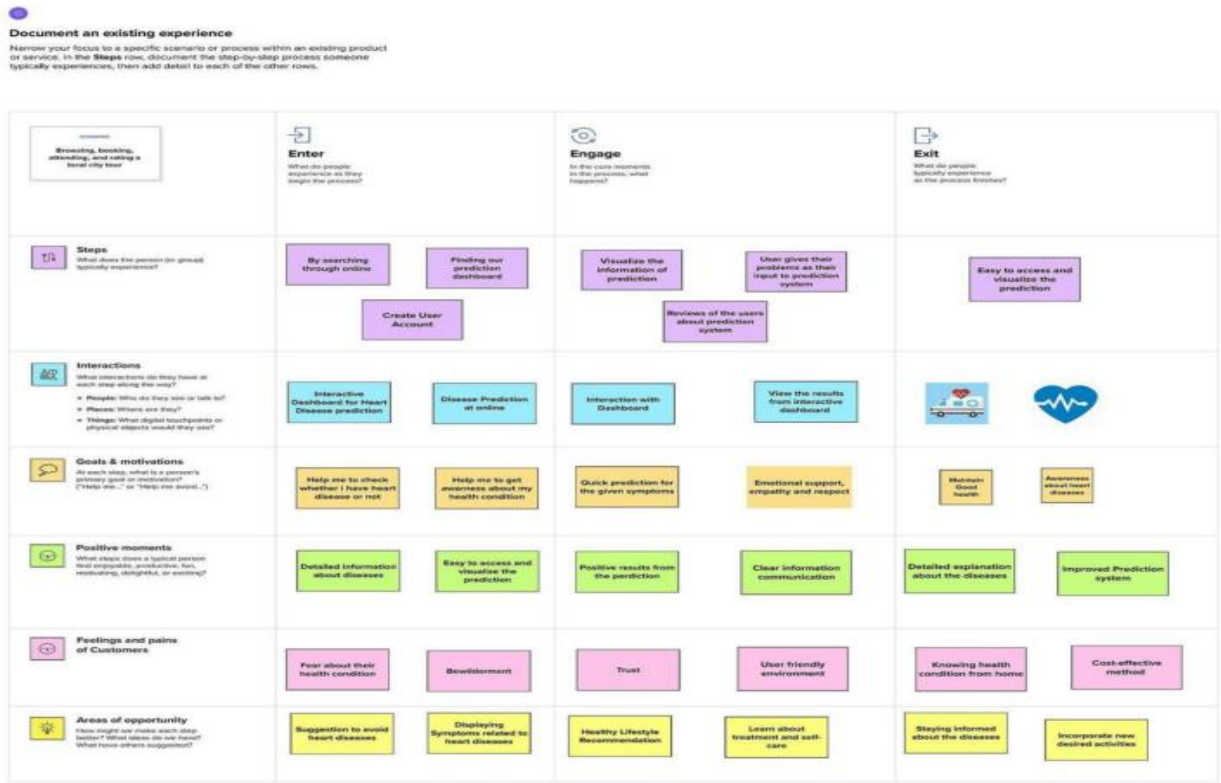
5.2 SOLUTION & TECHNICAL ARCHITECTURE:



5.3 USER STORIES:

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Web user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / Dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
	Login	USN-3	As a user, I can log into the application by entering email & password	I can access my account / Dashboard whenlogged in	High	Sprint-1
Customer (Web user)	Dashboard	USN-4	User can view his/her complete medical analysis and accuracy of disease prediction	I can view my medical analysis in the dashboard	High	Sprint-2
		USN-5	User can view the accuracy of occurrence of heart disease	I can view the accuracy of heart disease in the dashboard	High	Sprint-2
Customer Care Executive	Helpdesk	USN-6	As a customer care executive, he/she can view the customer queries.	I can post my queries in the dashboard	Medium	Sprint-3
		USN-7	As a customer care executive, he/she can answer the customer queries.	I can get support from helpdesk	High	Sprint-3

5.4 CUSTOMER JOURNEY MAP:



PROJECT PLANNING & PHASE

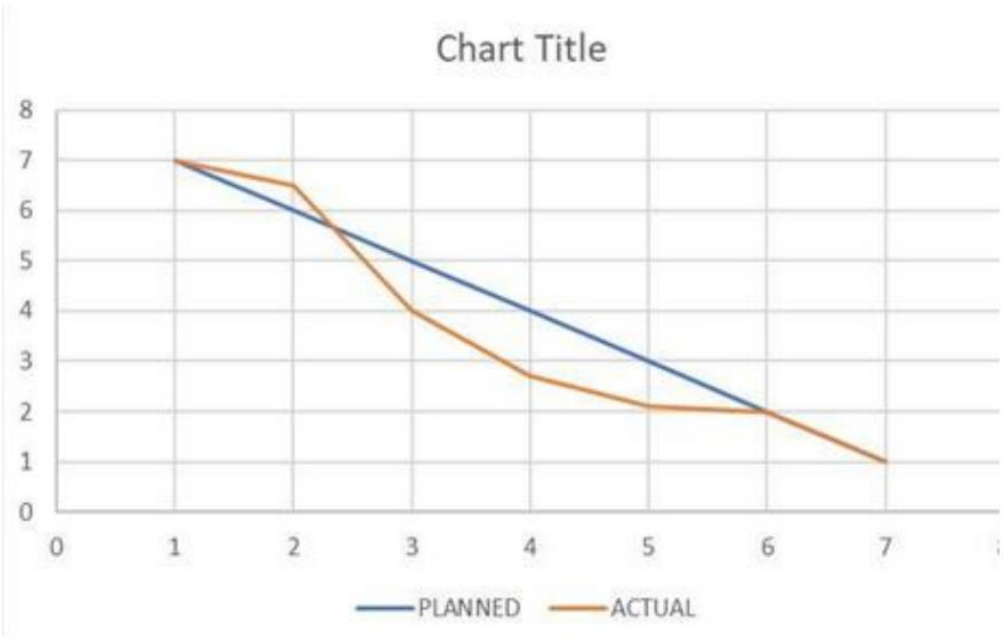
6.1 PRODUCT BACKLOG, SPRINT SCHEDULE, AND ESTIMATION

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	10	High	Hemanth Kumar P
Sprint-1	Login	USN-2	As a user, I will receive confirmation email once I have registered for the application	10	High	Preethi S, Snega R
Sprint-2	Login	USN-3	As a user, I can register for the application through Gmail	20	Medium	Vasanth E
Sprint-3	Dashboard	USN-4	As a user I can visualize the trends on the heart disease through Dashboard	20	High	Hemanth Kumar P, Preethi S,
Sprint-4	Dashboard	USN-5	I can view my health records and can make the decisions accordingly	20	High	Vasanth E, Snega R

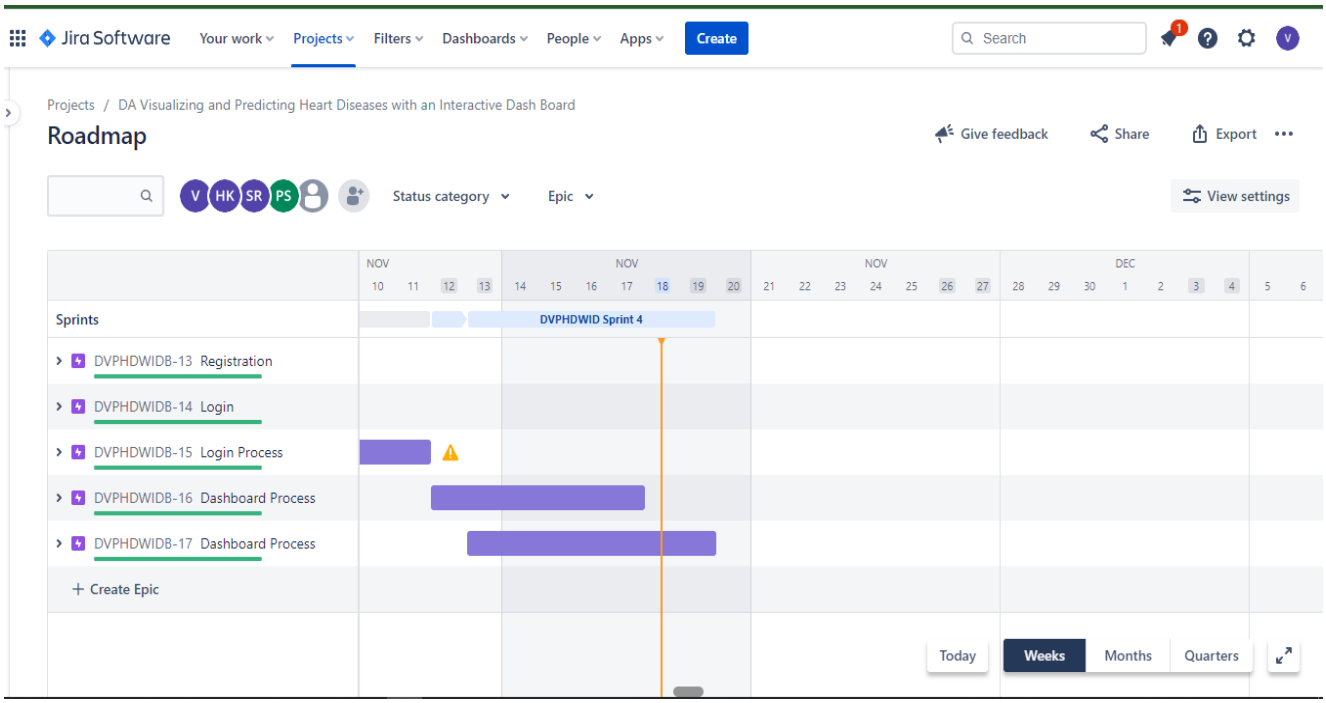
6.2PROJECT TRACKER, VELOCITY & BURNDOWN CHART:

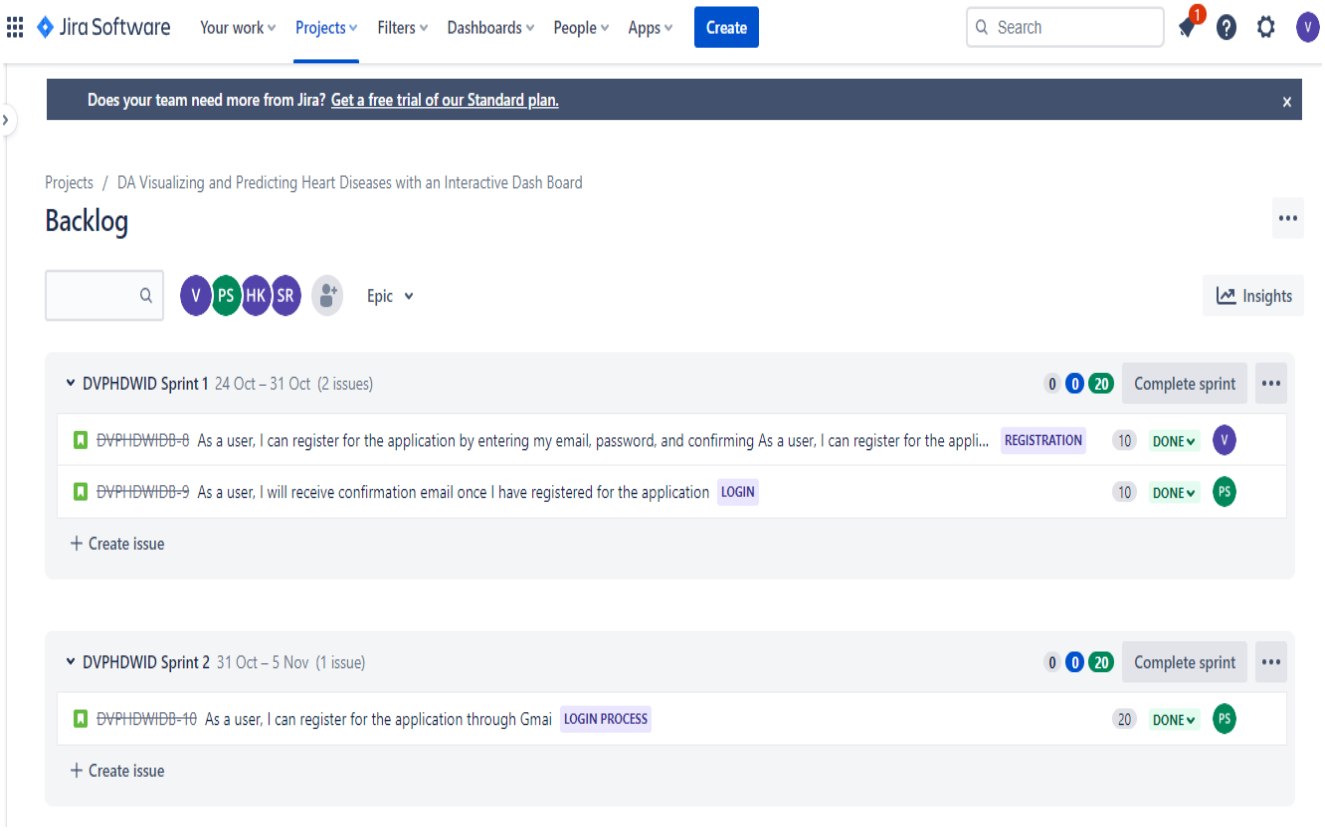
Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

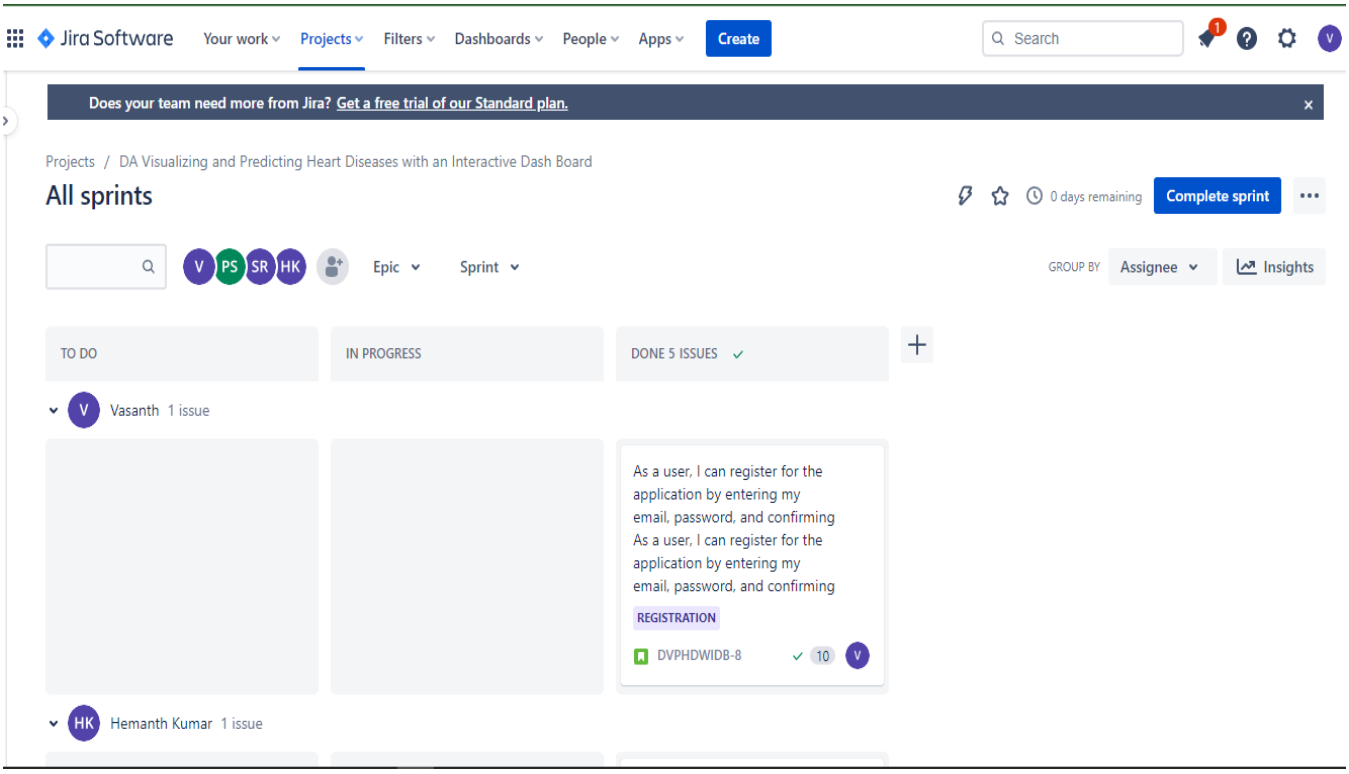
6.3 BURN DOWN CHART:



6.4 REPORTS FROM JIRA:




Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / DA Visualizing and Predicting Heart Diseases with an Interactive Dash Board
Backlog
Epic
Insights
DVPHDWID Sprint 1 24 Oct – 31 Oct (2 issues)
0 0 20
Complete sprint
DVPHDWIDB-8 As a user, I can register for the application by entering my email, password, and confirming As a user, I can register for the appli...
REGISTRATION
10
DONE
DVPHDWIDB-9 As a user, I will receive confirmation email once I have registered for the application
LOGIN
10
DONE
Create issue
DVPHDWID Sprint 2 31 Oct – 5 Nov (1 issue)
0 0 20
Complete sprint
DVPHDWIDB-10 As a user, I can register for the application through Gmai
LOGIN PROCESS
20
DONE
Create issue


Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / DA Visualizing and Predicting Heart Diseases with an Interactive Dash Board
All sprints
0 days remaining
Complete sprint
Epic
Sprint
GROUP BY
Assignee
Insights
TO DO
IN PROGRESS
DONE 5 ISSUES
Vasanth 1 issue
As a user, I can register for the application by entering my email, password, and confirming As a user, I can register for the application by entering my email, password, and confirming
REGISTRATION
DVPHDWIDB-8
10
Hemanth Kumar 1 issue

CODING & SOLUTIONING

HTML FILES:

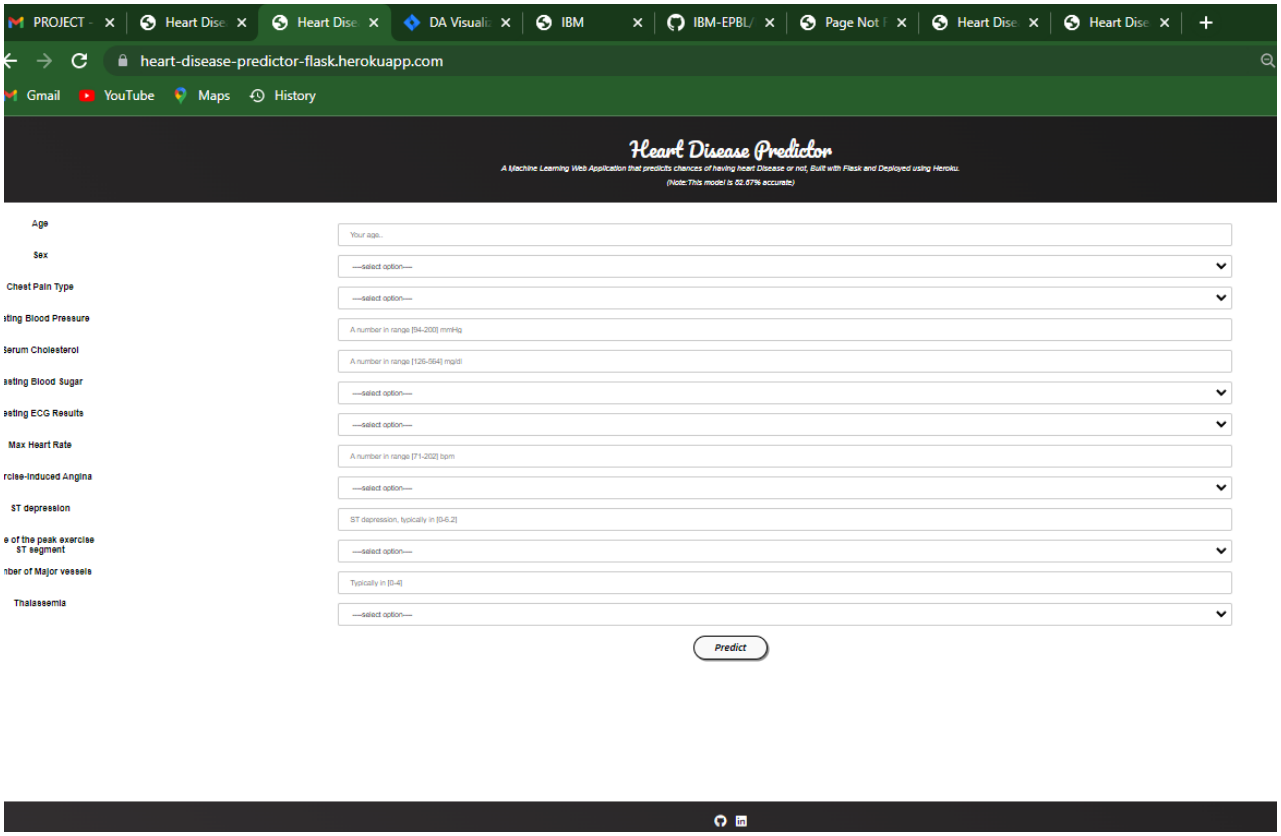
Main.html:

```
<!DOCTYPE html>
<html lang="en" dir="ltr">
<head>
<meta charset="utf-8">
<meta name="viewport" content="width=device-width, initial-scale=1.0">
<title>Heart Disease Predictor</title>
<link rel="stylesheet" type="text/css" href="{{ url_for('static', filename='style.css') }}">
<scriptsrc="https://kit.fontawesome.com/5f3f547070.js"
crossorigin="anonymous"></script>
<linkhref="https://fonts.googleapis.com/css2?family=Pacifico&display=swap"
rel="stylesheet">
</head>

<body>

<!-- Website Title -->

<div class="container">
<h2 class='container-heading'><span class="heading_font">Heart
Disease Predictor</span></h2>
<div class='description'>
<p>A Machine Learning Web Application that predicits chances of having heart
 Disease or not, Built with Flask and Deployed using Heroku.</p><br>
<p>(Note:This model is 82.67% accurate)</p>
</div>
</div>

<!-- Text Area -->
<div class="ml-container">
<form action="{{ url_for('predict')}}" method="POST">


<label for="age">Age</label>
<input type="text" id="age" name="age" placeholder="Your age.."><br>
```

```
<label for="sex">Sex</label>
<select id="sex" name="sex">
<option selected>----select option----</option>
<option value="1">Male</option>
<option value="0">Female</option>
</select><br>



<label for="cp">Chest Pain Type</label>
<select id="cp" name="cp">
<option selected>----select option----</option>
<option value="0">Typical Angina</option>
<option value="1">Atypical Angina</option>
<option value="2">Non-anginal Pain</option>
<option value="3">Asymtomatic</option>
</select><br>



<label for="trestbps">Resting Blood Pressure</label>
<input type="text" id="trestbps" name="trestbps" placeholder="A number in range
 [94-200] mmHg"><br>



<label for="chol">Serum Cholesterol</label>
<input type="text" id="chol" name="chol" placeholder="A number in range
 [126-564] mg/dl"><br>



<label for="fbs">Fasting Blood Sugar</label>
<select id="fbs" name="fbs">
<option selected>----select option----</option>
<option value="1">Greater than 120 mg/dl</option>
<option value="0">Less than 120 mg/dl</option>
</select><br>
```

```
<label for="restecg">Resting ECG Results</label>
<select id="restecg" name="restecg">
<option selected>----select option----</option>
<option value="0">Normal</option>
<option value="1">Having ST-T wave abnormality</option>
<option value="2">Probable or definite left ventricular
hypertrophy</option>
</select><br>



<label for="thalach">Max Heart Rate </label>
<input type="text" id="thalach" name="thalach" placeholder="A number in range
 [71-202] bpm"><br>



<label for="exang">Exercise-induced Angina</label>
<select id="exang" name="exang">
<option selected>----select option----</option>
<option value="1">Yes</option>
<option value="0">No</option>
</select><br>



<label for="oldpeak">ST depression</label>
<input type="text" id="oldpeak" name="oldpeak" placeholder="ST depression,
typically in [0-6.2]"><br>



<label for="slope">slope of the peak exercise ST segment</label>
<select id="slope" name="slope">
<option selected>----select option----</option>
<option value="0">Upsloping</option>
<option value="1">Flat</option>
<option value="2">Downsloping</option>
</select><br>
```

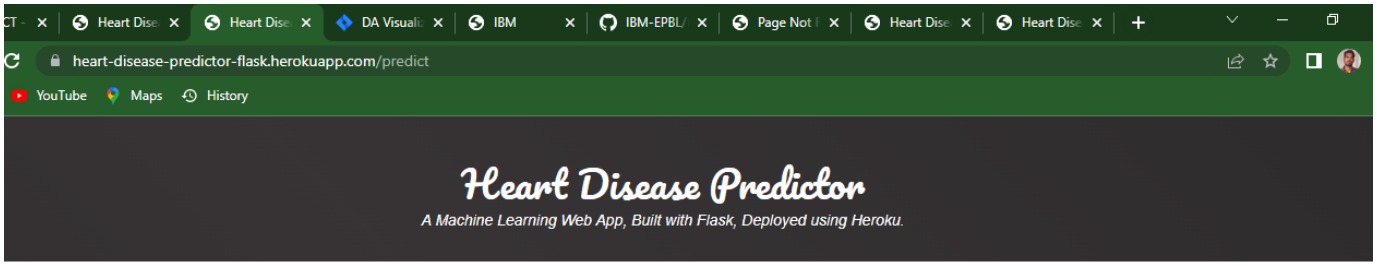
```
<label for="ca">Number of Major vessels</label>
<input type="text" id="ca" name="ca" placeholder="Typically in [0-4]"><br>

<label for="thal">Thalassemia</label>
<select id="thal" name="thal">
<option selected>----select option----</option>
<option value="0">Normal</option>
<option value="1">Fixed Defect</option>
<option value="2">Reversible Defect</option>
</select><br>

<input type="submit" class="my-cta-button" value="Predict">
</form>
</div>



<!-- Footer -->
<div class='footer'>
<div class="contact">
<a target="_blank" href="https://github.com/asthasharma98/Heart-Disease-
Prediction-Deployment"><i
class="fab fa-github fa-lg contact-icon"></i></a>
<a target="_blank" href="https://www.linkedin.com/in/astha-sharma-47266b11b/">
<i class="fab fa-linkedin fa-lg contact-icon"></i></a>
</div>
</div>

</body>

</html>
```



Result.html

```
<!DOCTYPE html>
<html lang="en" dir="ltr">
<head>
<meta charset="utf-8">
<meta name="viewport" content="width=device-width, initial-scale=1.0">
<title>Heart Disease Predictor</title>
<link rel="shortcut icon" href="{{ url_for('static', filename='diabetes-favicon.ico') }}">
<link rel="stylesheet" type="text/css" href="{{ url_for('static', filename='style.css') }}">
<scriptsrc="https://kit.fontawesome.com/5f3f547070.js"crossorigin="anonymous">
</script>

<linkhref=https://fonts.googleapis.com/css2?family=Pacifico&display=swap
rel="stylesheet">
</head>

<body>

<!-- Website Title -->
```

```
        <div class="container">
        <h2 class='container-heading'><spanclass="heading_font">Heart
Disease     Predictor</span></h2>
        <div class='description'>
                <p>A Machine Learning Web App, Built with Flask, Deployed
using Heroku.</p>
        </div>
    </div>

        <!-- Result -->
        <div class="results">
                {% if prediction==1 %}
        <h1>Prediction: <span class='danger'>Oops!
                 You have Chances of Heart Disease.</span></h1>

                {% elif prediction==0 %}
                        <h1>Prediction: <span class='safe'>Great! You DON'T
                chances have Heart Disease.</span></h1>

                {% endif %}
        </div>

        <!-- Footer -->
    <div class='footer'>
        <div class="contact">
          <atarget="_blank"href="https://github.com/IBM-EPBL/IBM-Project-36665-
1660297036 Prediction-Deployment">
<i class="fab fa-github fa-lg contact-icon"></i></a>
          <a target="_blank" href="https://www.linkedin.com">
  <i class="fab fa-linkedin fa-lg contact-icon"></i></a>
        </div>

    </div>

    </body>
</html>
```



Prediction: Oops! You have Chances of Heart Disease.

STYLESHEET:

Style.css

```
html{
        height: 100%;
        margin: 0;
    }

    body{
        font-family: Arial, Helvetica,sans-serif;
      text-align: center;
      margin: 0;
      padding: 0;
      width: 100%;
        height: 100%;
        display: flex;
        flex-direction: column;
    }
```

```
/* Website Title */

.container{
   padding: 30px;
        position:relative;
        background: linear-gradient(45deg, #161616, #383436, #161616);
        background-size: 500% 500%;
        animation: change-gradient 10s ease-in-out infinite;
}
@keyframes change-gradient {
        0%{
                background-position: 0 50%;
        }
        50%{
                background-position: 100% 50%;
        }
        100%{
                background-position: 0 50%;
        }
}

.container-heading{
   margin: 0;
}

.heading_font{
   color: #ffffff;
        font-family: 'Pacifico', cursive;
        font-size: 35px;
        font-weight: normal;
}

.description p{
        color: #ffffff;
   font-style: italic;
   font-size: 14px;
   margin: -5px 0 0;
}
```

```
/* Text Area */

.ml-container{
   margin: 30px 0;
        flex: 1 0 auto;

}

.form {
        text-align: center;
        width: 250px;
        height: 25px;
        margin-bottom: 5px;

}

input[type=text], select {
   width:60%;
   padding: 12px 20px;
   margin: 8px 0;
   display: inline-block;
   border: 1px solid #ccc;
   border-radius: 4px;
   box-sizing: border-box;
}

label {
        display: inline block;
        width: 200px;
        font-weight: bold;
        text-align: center;
        float: left;

 }
```

```
/* Predict Button */


.my-cta-button{
        background: #f9f9f9;
        border: 2px solid #000000;
   border-radius: 1000px;
   box-shadow: 3px 3px #8c8c8c;
        margin-top: 10px;
   padding: 10px 36px;
   color: #000000;
   display: inline-block;
        font: italic bold 20px/1 "Calibri", sans-serif;
   text-align: center;
}

.my-cta-button:hover{
   color: #141414;
   border: 2px solid #46424b;
}

.my-cta-button:active{
   box-shadow: 0 0;
}

/* Contact */
.contact-icon{
        color: #ffffff;
        padding: 7px;
}

.contact-icon:hover{
        color: #8c8c8c;
}
```

```
/* Footer */
.footer{
        flex-shrink: 0;
        position:relative;
        padding: 20px;
        background: linear-gradient(45deg, #161616, #383436, #161616);
        background-size: 500% 500%;
        animation: change-gradient 10s ease-in-out infinite;
}


/* Result */
.results{
   padding: 30px 0 0;
        flex: 1 0 auto;
}

.danger{
   color: #ff0000;
}

.safe{
   color: green;
}
```

PYTHON FILES:

app.py:

```
 # Importing essential libraries
from flask import Flask, render_template, request
import pickle
import numpy as np

# Load the Random Forest CLassifier model
filename = 'heart-disease-prediction-knn-model.pkl'
model = pickle.load(open(filename, 'rb'))

app = Flask(__name__)

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


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

      age = int(request.form['age'])
      sex = request.form.get('sex')
      cp = request.form.get('cp')
      trestbps = int(request.form['trestbps'])
      chol = int(request.form['chol'])
      fbs = request.form.get('fbs')
      restecg = int(request.form['restecg'])
      thalach = int(request.form['thalach'])
      exang = request.form.get('exang')
      oldpeak = float(request.form['oldpeak'])
      slope = request.form.get('slope')
      ca = int(request.form['ca'])
      thal = request.form.get('thal')
```

```
Data=np.array([[age,sex,cp,trestbps,chol,fbs,restecg,thalach,exang,oldpeak,slope,ca,thal]])
      my_prediction = model.predict(data)

      return render_template('result.html', prediction=my_prediction)



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

prediction.py:

```
# importing required libraries
import numpy as np
import pandas as pd
import pickle
from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
from sklearn.ensemble import RandomForestClassifier
from sklearn.svm import SVC
from sklearn.linear_model import LogisticRegression
from sklearn.neighbors import KNeighborsClassifier
from sklearn.tree import DecisionTreeClassifier


# loading and reading the dataset

heart = pd.read_csv("heart_cleveland_upload.csv")

# creating a copy of dataset so that will not affect our original dataset.
heart_df = heart.copy()

# Renaming some of the columns
heart_df = heart_df.rename(columns={'condition':'target'})
print(heart_df.head())

# model building
```

```
#fixing our data in x and y. Here y contains target data and X contains rest all the features.
x= heart_df.drop(columns= 'target')
y= heart_df.target

# splitting our dataset into training and testing for this we will use train_test_split library.
x_train, x_test, y_train, y_test= train_test_split(x, y, test_size= 0.25, random_state=42)

#feature scaling
scaler= StandardScaler()
x_train_scaler= scaler.fit_transform(x_train)
x_test_scaler= scaler.fit_transform(x_test)

# creating K-Nearest-Neighbor classifier
model=RandomForestClassifier(n_estimators=20)
model.fit(x_train_scaler, y_train)
y_pred= model.predict(x_test_scaler)
p = model.score(x_test_scaler,y_test)
print(p)

print('Classification Report\n', classification_report(y_test, y_pred))
print('Accuracy: {}%\n'.format(round((accuracy_score(y_test, y_pred)*100),2)))

cm = confusion_matrix(y_test, y_pred)
print(cm)

# Creating a pickle file for the classifier
filename = 'heart-disease-prediction-knn-model.pkl'
pickle.dump(model, open(filename, 'wb'))
```

TESTING

Project Development Phase (Model Performance Test)

8.1 MODEL PERFORMANCE TESTING:

S. No	Parameter	Screenshot / Values
1.	Dashboard design	Visualization-7
2.	Data Responsiveness	Yes, the website is responsive completely, by resizing the browser window size as per the test scenario.
3.	Amount Data to Rendered (DB2 Metrics)	Totally there are **298** records in the dataset.
4.	Utilization of Data Filters	Data Filter used in Visualizing and Predicting Heart Disease with an Interactive Dash Board
5.	Effective User Story	• To work on the given dataset • To Understand the Dataset • Load the dataset to Cloud platform then Build the required Visualizations • With the help of Heart Disease dataset, create various graphs & Charts to highlight the insights in the dataset • Build a Visualizations to showcase the Heart Disease Prediction

Acceptance Testing

UAT Execution & Report Submission

PURPOSE OF DOCUMENT:

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

DEFECT ANALYSIS:

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

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

8.2 TEST CASE ANALYSIS:

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

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	7	0	0	7
Client Application	51	0	0	51
Security	2	0	0	2
Outsource Shipping	3	0	0	3

RESULTS

9.1 PERFORMANCE METRICS

1. Hours worked: 50 hours
2. Stick to Timelines: 100%
3. Stay within budget: 100%
4. Consistency of the product: 85%
5. Efficiency of the product: 85%

ADVANTAGES & DISADVANTAGES:

ADVANTAGES:

- Smooth User Interface
- Accuracy is achieved quickly

DISADVANTAGES:

Random forest can be used for both classification and regression tasks, butit is no moresuitable for Regression tasks

CONCLUSION:

This overview of the project conveys the idea that numerous methods have been investigated for diagnosing cardiovascular disease. Big data, machine learning, and data mining can be used to great success to analyses the prediction model with the highest degreeof accuracy. The primary goal of this project is to diagnose cardiovascular disease or heart disease utilizing a variety of techniques and procedures to obtain a prognosis.

FUTURE SCOPE

A future update shall comprise of section for viewing renowned cardiologists and scan centers in their city. The obtained output can be further processed and sent to smart devices to provide necessary assistance. Constant monitoring can provide necessary data to recommend to consult a doctor in case of an emergency.

APPENDIX

Source Code: https://github.com/IBM-EPBL/IBM-Project-36665-1660297036/tree/main/PROJECT%20DEVELOPMENT

GitHub Link: https://github.com/IBM-EPBL/IBM-Project-36665-1660297036

Project Demo link:
https://drive.google.com/file/d/1nwbsvKZFCqtxM5n5thCcDBwyrGPhGAjX/view?usp=sharing