Visualizing and Predicting Heart Diseases with an Interactive Dashboard

NALAIYA THIRAN PROJECT REPORT 2022

Submitted by

Team Leader: Bibin Francis V J (71381904010)

Team Member 1: Abirami S (71381904002)

Team Member 2: Kayalvizhe J (71381904030)

Team Member 3: Kishankanth T G (71381904031)

Team ID: PNT2022TMID07860

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 Kmeans 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, Kmeans 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.

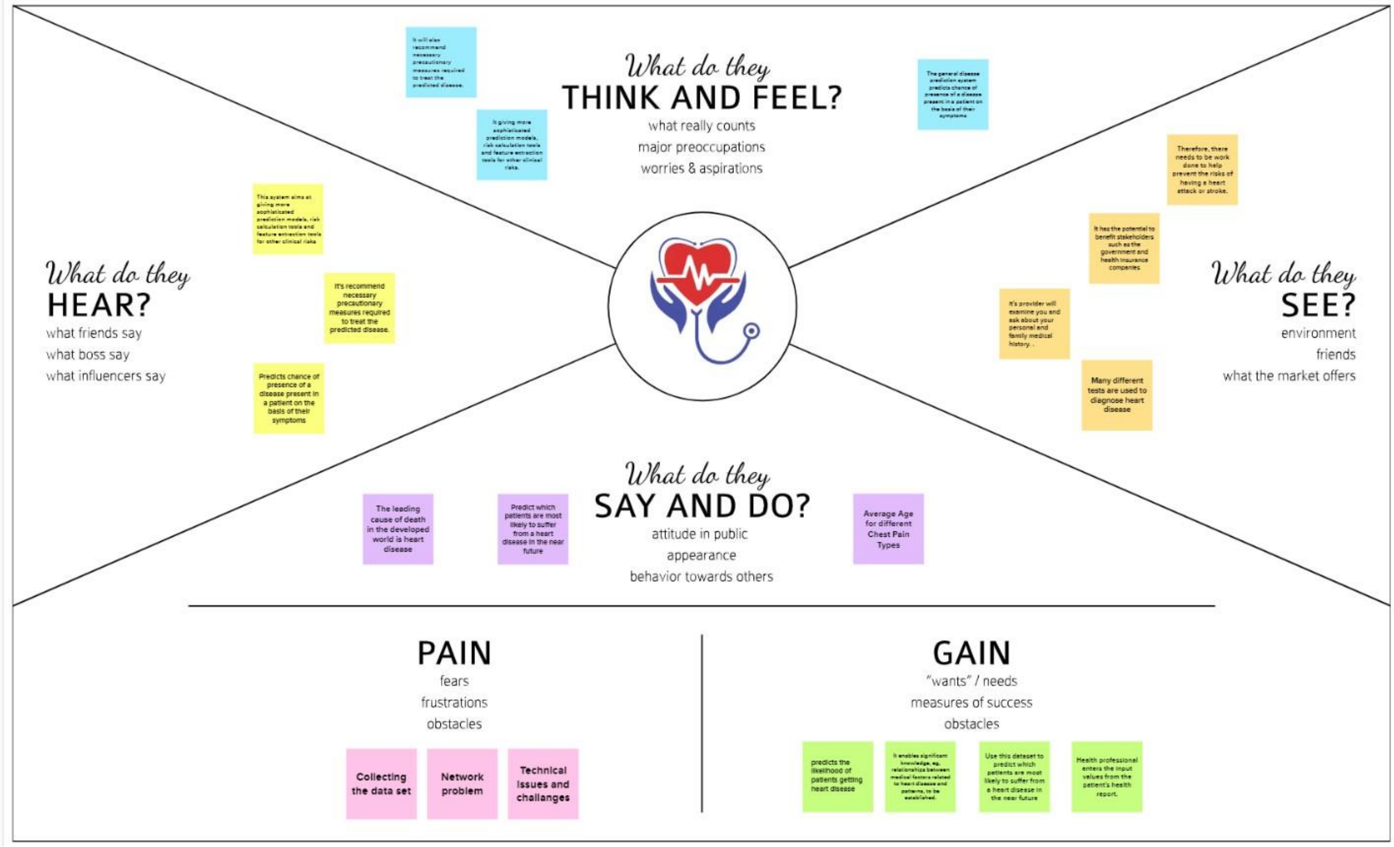
2.3 Problem Statement Definition

Who does the problem affect?	The majority of people who die of coronary heart disease are 65 or older. While heart attacks can strike people of both genders in old age, women are at greater risk of dying.
What are the boundaries of the problem?	Several health conditions, your lifestyle, and your age and family history can increase your risk for heart disease.
What is the issue?	If the person is affected by heart disease, then it produces the side effects like Chest pain, chest tightness, chest pressure, chest discomfort (angina), Shortness of breath, pain in the neck, jaw, throat, upper belly area or back.

When does the issue occur?	Heart disease and the other conditions that lead to it can happen at any age. High rates of obesity and high blood pressure among younger people are putting them at risk for heart disease earlier in their life.
Where is the issue coming?	Coronary artery disease happens when coronary arteries struggle to supply the heart with enough blood, oxygen and nutrients. Cholesterol deposits, or plaques, are almost always to blame. These buildups narrow your arteries, decreasing blood flow to your heart. This can cause chest pain, shortness of breath or even a heart attack.
Why is it important to fix the problem?	Predict if the patient suffers from heart disease. The health professional enters the input values from the patient's health report. The data is fed into model which predicts the probability of having heart disease.
Which solution can be used to address the issue?	A machine learning powered web application model with the strong building of algorithm that helps to identify and predicts the disease with the identification of symptoms. It processes the breathing signals using a neural network that infers whether the person has heart disease, and if they are identified then it assesses the severity of their disease in accordance with the Movement Disorder Society Unified Heart Disease using ML algorithms.
Which methodologies are used to solve the issue?	Supervised and Un-supervised machine learning, Data mining, Computer vision with OpenCV, Python web application interface - Flask, Jupyter Notebook, IBM Cloud.

3. Ideation and Proposed Solution

3.1 Empathy Map Canvas



What do they THINK AND FEEL?
what really counts
major preoccupations
worries & aspirations

It will also recommend necessary precautionary measures required to treat the predicted disease.

It giving more sophisticated prediction models, risk calculation tools and feature extraction tools for other clinical risks.

The general disease prediction system predicts chance of presence of a disease present in a patient on the basis of their symptoms

What do they HEAR?
what friends say
what boss say
what influencers say

This system aims at giving more sophisticated prediction models, risk calculation tools and feature extraction tools for other clinical risks

It's recommend necessary precautionary measures required to treat the predicted disease.

Predicts chance of presence of a disease present in a patient on the basis of their symptoms

What do they SEE?
environment
friends
what the market offers

Therefore, there needs to be work done to help prevent the risks of having a heart attack or stroke.

It has the potential to benefit stakeholders such as the government and health insurance companies

It's provider will examine you and ask about your personel and family medical history. .

Many different tests are used to diagnose heart disease

What do they SAY AND DO?
attitude in public
appearance
behavior towards others

The leading cause of death in the developed world is heart disease

Predict which patients are most likely to suffer from a heart disease in the near future

Average Age for different Chest Pain Types

PAIN
fears
frustrations
obstacles

Collecting the data set

Network problem

Technical issues and challanges

GAIN
"wants" / needs
measures of success
obstacles

predicts the likelihood of patients getting heart disease

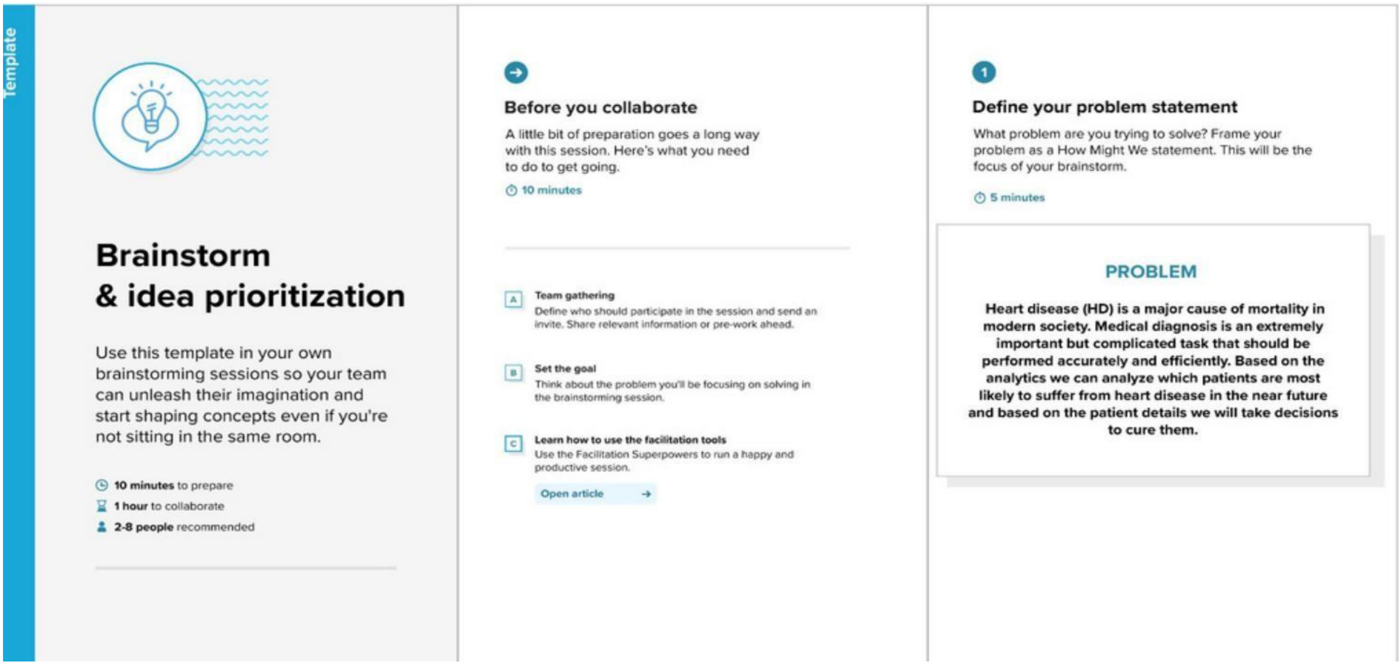
It enables significant knowledge, eg, relationships between medical factors related to heart disease and patterns, to be established.

Use this dataset to predict which patients are most likely to suffer from a heart disease in the near future

Health professional enters the input values from the patient's health report.

3.2 Ideation and Brainstorming

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



Template

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

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**

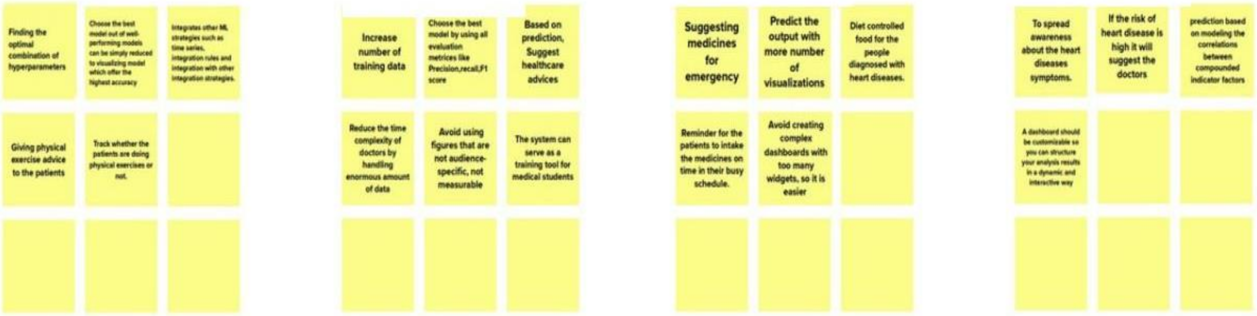
What problem are you trying to solve? Frame your problem as a How Might We statement. This will be the focus of your brainstorm.

5 minutes

PROBLEM

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. 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 take decisions to cure them.

Step-2: Brainstorm, Idea Listing and Grouping



Finding the optimal combination of hyperparameters

Choose the best model out of well-performing models can be simply reduced to visualizing model which offer the highest accuracy

Integrates other ML strategies such as time series, integration rules and integration with other integration strategies.

Giving physical exercise advice to the patients

Track whether the patients are doing physical exercises or not.

Increase number of training data

Choose the best model by using all evaluation metrices like Precision,recall,F1 score

Based on prediction, Suggest healthcare advices

Reduce the time complexity of doctors by handling enormous amount of data

Avoid using figures that are not audience-specific, not measurable

The system can serve as a training tool for medical students

Suggesting medicines for emergency

Predict the output with more number of visualizations

Diet controlled food for the people diagnosed with heart diseases.

Reminder for the patients to intake the medicines on time in their busy schedule.

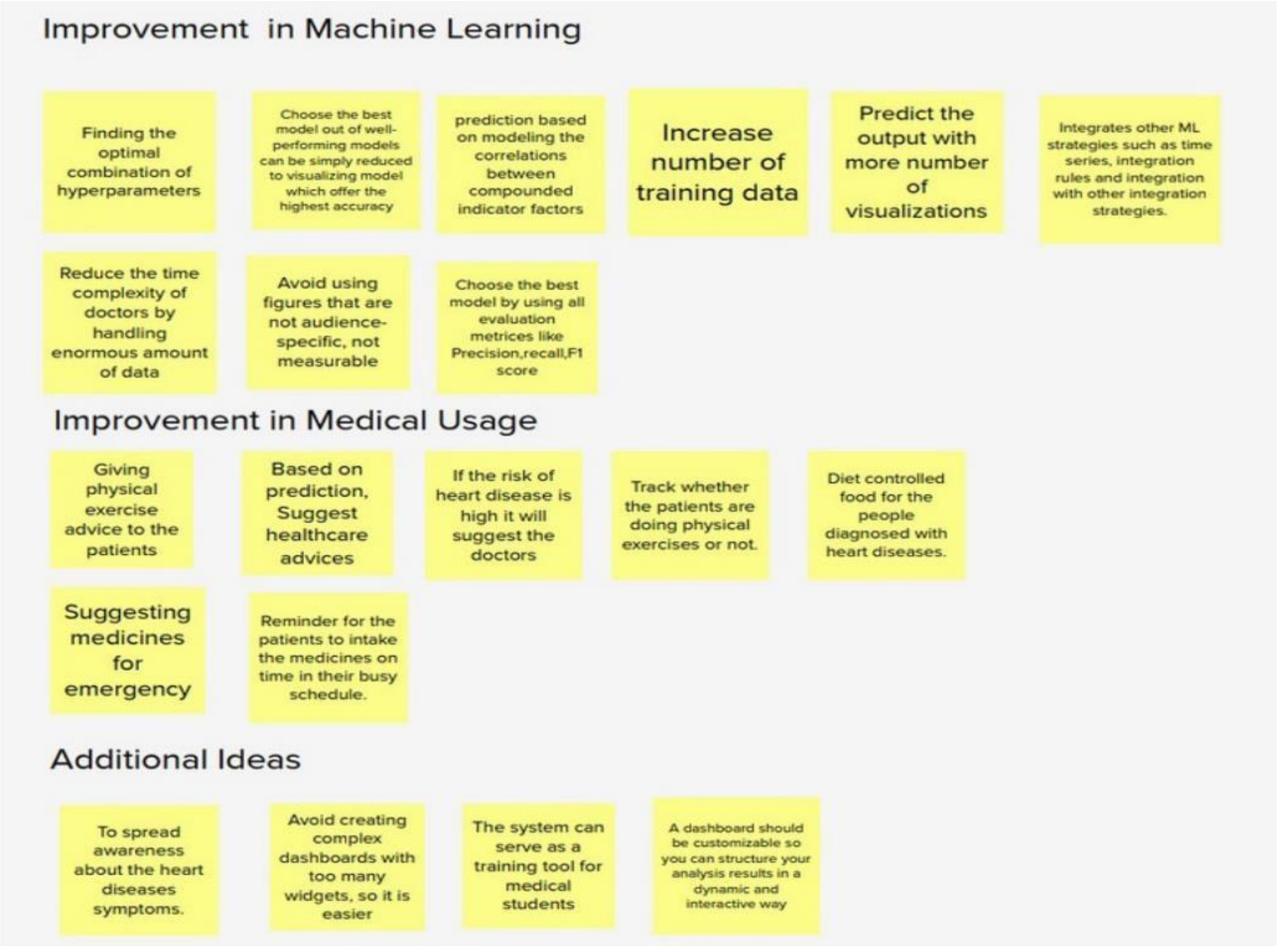
Avoid creating complex dashboards with too many widgets, so it is easier

To spread awareness about the heart diseases symptoms.

If the risk of heart disease is high it will suggest the doctors

prediction based on modeling the correlations between compounded indicator factors

A dashboard should be customizable so you can structure your analysis results in a dynamic and interactive way



Improvement in Machine Learning

Finding the optimal combination of hyperparameters

Choose the best model out of well-performing models can be simply reduced to visualizing model which offer the highest accuracy

prediction based on modeling the correlations between compounded indicator factors

Increase number of training data

Predict the output with more number of visualizations

Integrates other ML strategies such as time series, integration rules and integration with other integration strategies.

Reduce the time complexity of doctors by handling enormous amount of data

Avoid using figures that are not audience-specific, not measurable

Choose the best model by using all evaluation metrices like Precision,recall,F1 score

Improvement in Medical Usage

Giving physical exercise advice to the patients

Based on prediction, Suggest healthcare advices

If the risk of heart disease is high it will suggest the doctors

Track whether the patients are doing physical exercises or not.

Diet controlled food for the people diagnosed with heart diseases.

Suggesting medicines for emergency

Reminder for the patients to intake the medicines on time in their busy schedule.

Additional Ideas

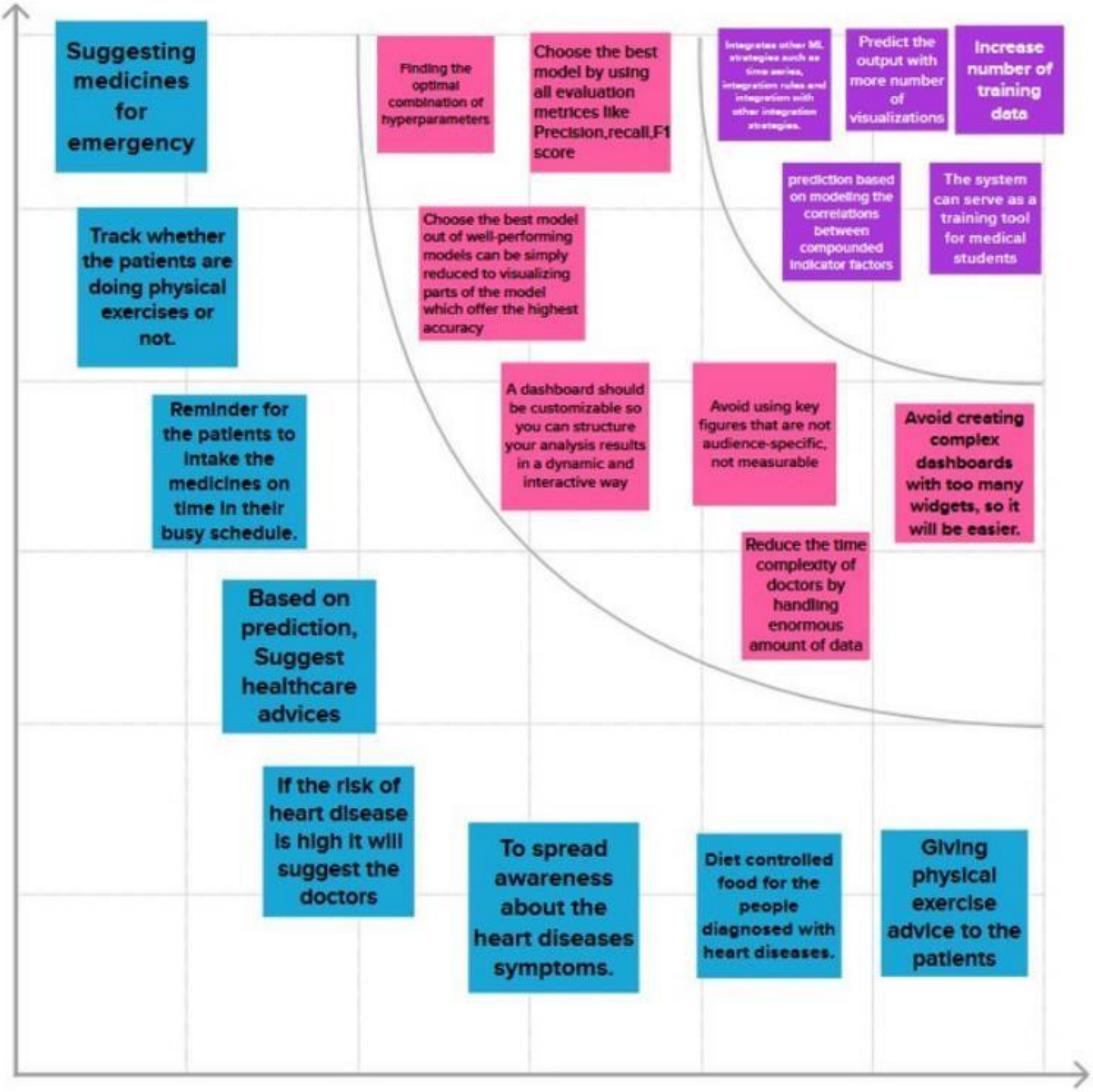
To spread awareness about the heart diseases symptoms.

Avoid creating complex dashboards with too many widgets, so it is easier

The system can serve as a training tool for medical students

A dashboard should be customizable so you can structure your analysis results in a dynamic and interactive way

Step-3: Idea Prioritization

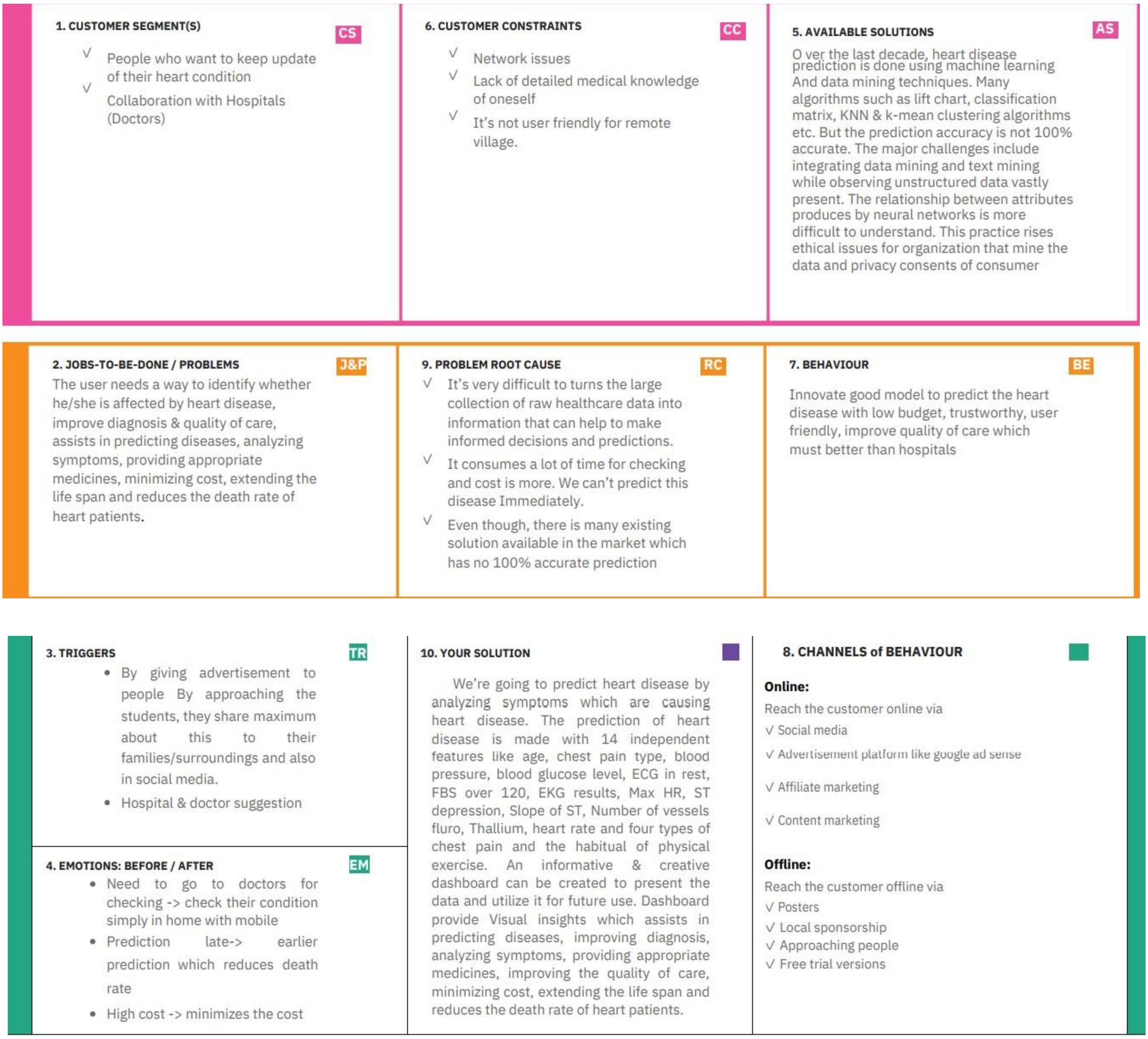

Suggesting medicines for emergency
Finding the optimal combination of hyperparameters
Choose the best model by using all evaluation metrices like Precision,recall,F1 score
Integrates other ML strategies such as time series, integration rules and integration with other integration strategies.
Predict the output with more number of visualizations
Increase number of training data
prediction based on modeling the correlations between compounded Indicator factors
The system can serve as a training tool for medical students
Track whether the patients are doing physical exercises or not.
Choose the best model out of well-performing models can be simply reduced to visualizing parts of the model which offer the highest accuracy
Reminder for the patients to intake the medicines on time in their busy schedule.
A dashboard should be customizable so you can structure your analysis results in a dynamic and interactive way
Avoid using key figures that are not audience-specific, not measurable
Avoid creating complex dashboards with too many widgets, so it will be easier.
Reduce the time complexity of doctors by handling enormous amount of data
Based on prediction, Suggest healthcare advices
If the risk of heart disease is high it will suggest the doctors
To spread awareness about the heart diseases symptoms.
Diet controlled food for the people diagnosed with heart diseases.
Giving physical exercise advice to the patients

3.3 Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	The user needs a way to identify whether he/she is affected by Heart disease, improve diagnosis & quality of care, assists in predicting diseases, analysing symptoms, providing appropriate medicines, minimizing cost, extending the life span and reduces the death rate of heart patients.
2.	Idea / Solution description	By predicting and visualizing the fundamentals properties that are related to heart disease and visualizing them in a

		dashboard.
3.	Novelty / Uniqueness	The use of analytics in healthcare improves care by facilitating preventive care and visually represented data provide various insights easily. Prediction is non invasive. So it is cost efficient. Earlier prediction is very helpful in reducing mortality rate.
4.	Social Impact / Customer Satisfaction	It will reduce the mortality rate due to heart disease. Heart prediction can be done easier and earlier by visual analytics. As it is cost efficient, it is preferred by most of the customers. Most importantly, it is very helpful for doctors to give treatments according to the patients conditions and it's preferred by the doctors as it saves time.
5.	Business Model (Revenue Model)	There are 2 ways to generate revenue from this project by creating a product model. By introducing an app for predicting heart disease or it can be integrated with smart watches for producing more efficient models.
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



1. CUSTOMER SEGMENT(S) CS

√ People who want to keep update of their heart condition

√ Collaboration with Hospitals (Doctors)

6. CUSTOMER CONSTRAINTS CC

√ Network issues

√ Lack of detailed medical knowledge of oneself

√ It's not user friendly for remote village.

5. AVAILABLE SOLUTIONS AS

O ver the last decade, heart disease prediction is done using machine learning And data mining techniques. Many algorithms such as lift chart, classification matrix, KNN & k-mean clustering algorithms etc. But the prediction accuracy is not 100% accurate. The major challenges include integrating data mining and text mining while observing unstructured data vastly present. The relationship between attributes produces by neural networks is more difficult to understand. This practice rises ethical issues for organization that mine the data and privacy consents of consumer

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

The user needs a way to identify whether he/she is affected by heart disease, improve diagnosis & quality of care, assists in predicting diseases, analyzing symptoms, providing appropriate medicines, minimizing cost, extending the life span and reduces the death rate of heart patients.

9. PROBLEM ROOT CAUSE RC

√ It's very difficult to turns the large collection of raw healthcare data into information that can help to make informed decisions and predictions.

√ It consumes a lot of time for checking and cost is more. We can't predict this disease Immediately.

√ Even though, there is many existing solution available in the market which has no 100% accurate prediction

7. BEHAVIOUR BE

Innovate good model to predict the heart disease with low budget, trustworthy, user friendly, improve quality of care which must better than hospitals

3. TRIGGERS TR

- By giving advertisement to people By approaching the students, they share maximum about this to their families/surroundings and also in social media.
- Hospital & doctor suggestion

4. EMOTIONS: BEFORE / AFTER EM

- Need to go to doctors for checking -> check their condition simply in home with mobile
- Prediction late-> earlier prediction which reduces death rate
- High cost -> minimizes the cost

10. YOUR SOLUTION

We're going to predict heart disease by analyzing symptoms which are causing heart disease. The prediction of heart disease is made with 14 independent features like age, chest pain type, blood pressure, blood glucose level, ECG in rest, FBS over 120, EKG results, Max HR, ST depression, Slope of ST, Number of vessels fluro, Thallium, heart rate and four types of chest pain and the habitual of physical exercise. An informative & creative dashboard can be created to present the data and utilize it for future use. Dashboard provide Visual insights which assists in predicting diseases, improving diagnosis, analyzing symptoms, providing appropriate medicines, improving the quality of care, minimizing cost, extending the life span and reduces the death rate of heart patients.

8. CHANNELS of BEHAVIOUR

Online:

Reach the customer online via

√ Social media

√ Advertisement platform like google ad sense

√ Affiliate marketing

√ Content marketing

Offline:

Reach the customer offline via

√ Posters

√ Local sponsorship

√ Approaching people

√ Free trial versions

4.Requirement Analysis

4.1 Functional Requirement

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Enables user to make registration for the application through 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

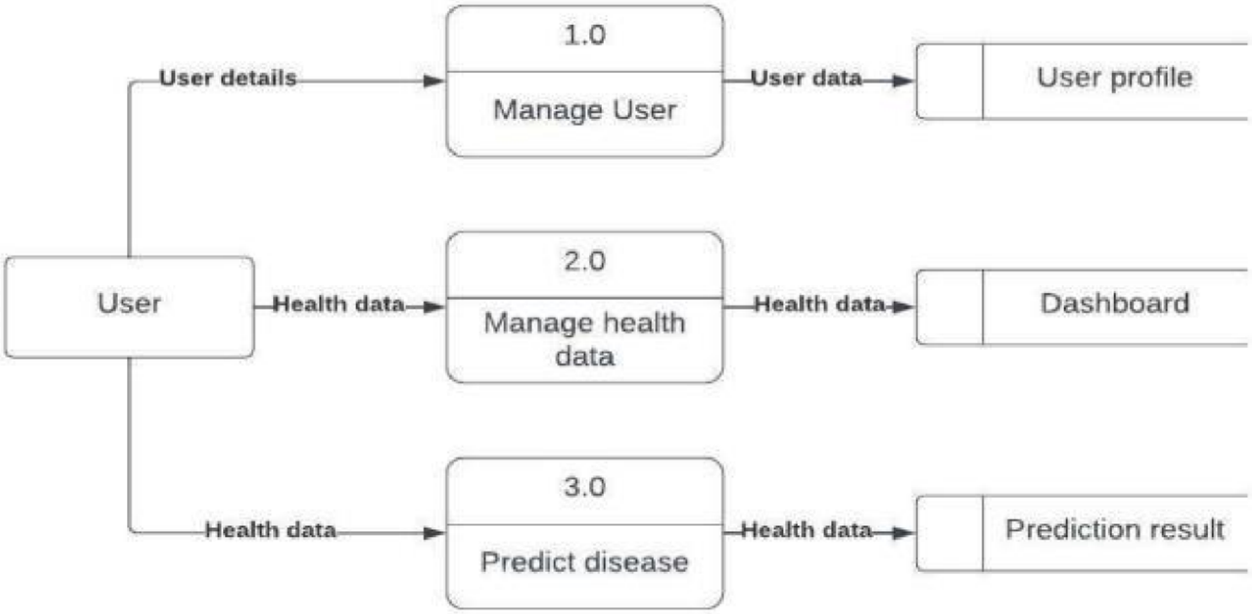
4.2 Non-Functional Requirement

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

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The application will have a simple and userfriendly 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. 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
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 Diagram



5.2 Solution and Technical Architecture

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2

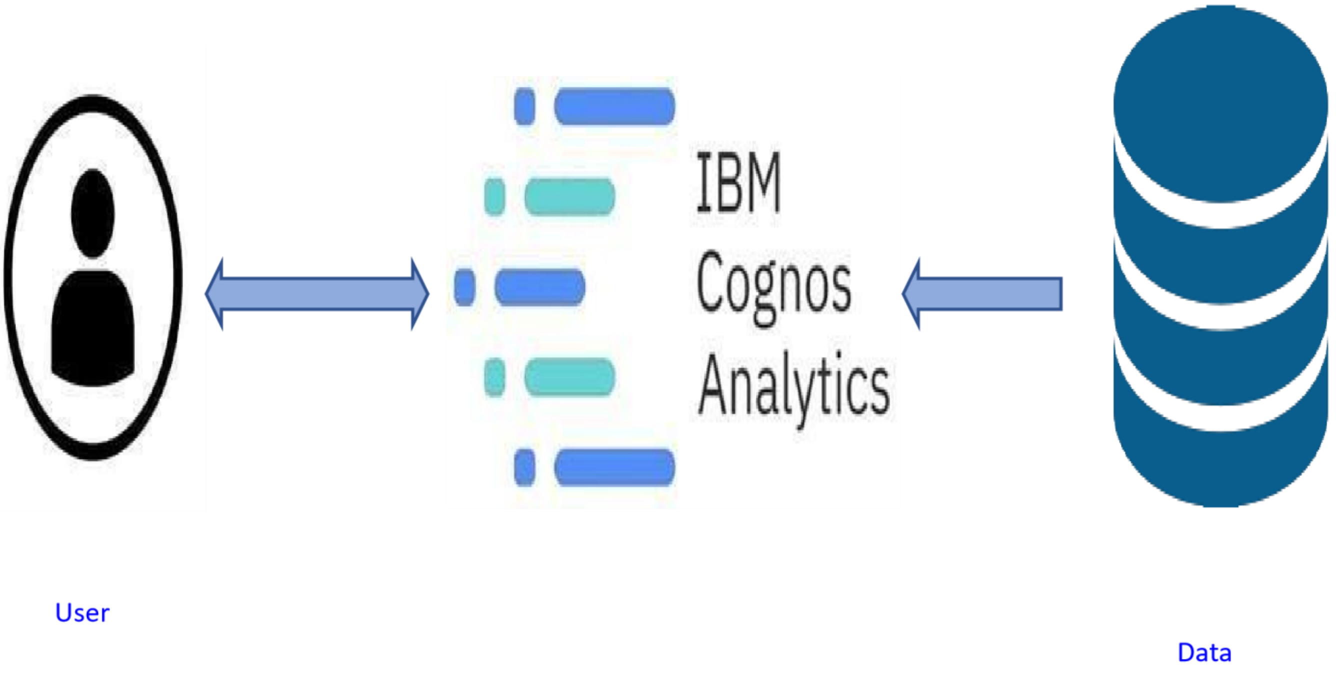


Table-1 : Components & Technologies:

S.NO.	Component	Description	Technology
1.	User Interface	How user interacts with application e.g. Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript / Angular JS / React JS etc.
2.	Application Logic-1	Logic for a process in the application	Python
3.	Application Logic-2	Logic for a process in the application	IBM Watson, IBM Cognos Analytics
4.	Application Logic-3	Logic for a process in the application	IBM Cognos Analytics
5.	Database	Data Type, Configurations etc.	MySQL, NoSQL, Cloudant DB ,OLAP
6.	Cloud Database	Database Service on Cloud	IBM DB2, IBM Cloudant etc.
7.	File Storage	File storage requirements	IBM Block Storage or Other Storage Service or Local Filesystem
8.	External API-1	Purpose of External API used in the application	IBM Watson API
9.	Machine Learning Model	Purpose of Machine Learning Model	Regression Model, Classification Model, Clustering Model, Object Recognition Model, etc.,

10.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :	Local, Cloud Foundry, Kubernetes, etc.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Power BI is truly an interactive tool that gets connected with online platforms to fetch the data for you. With the connectors and pre-installed dashboards, Power BI can analyse the data and present visually creative reports by connecting with Google Analytics, Salesforce, and other important software.	Microsoft Power BI, IBM Cognos
2.	Security Implementations	Authenticated users Hosted on Cloud-based servers, it offers strong, multilayer security to all data exchanged, also remains protected from Cyber attacks	IBM Cloud
3.	Scalable Architecture	Support feature increase in throughput and able to handle data of any patient at any given point of time without affecting the stability.	Cognos BI
4.	Availability	Ensure that data is available to the end users, Reliable access to data.	AWS, Cloud platforms, Microsoft Power BI
5.	Performance	The process of quickly examining extremely large data sets to find insights. This is done by using the parallel processing of high performance computing to run powerful analytic software.	IBM Cognos Analytics

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 when logged 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

Administrator	User Profile	USN-8	As an admin, he/she can update the health details of users.	I can view my updated health details.	High	Sprint-4
		USN-9	As an admin, he/she can add or delete users.	I can access my account / Dashboard when logged in	High	Sprint-4
		USN-10	As an admin, he/she can manage the user details.	I can view the organized data of myself.	High	Sprint-4

6. Project Planning and Scheduling

6.1 Script Planning and Execution

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	1	High	Bibin Francis V J, Abirami S, Kayalvizhe J, Kishankanth T G
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	1	High	Bibin Francis V J, Abirami S, Kayalvizhe J, Kishankanth T G
Sprint-1	Login	USN-3	As a user, I can log into the application by entering email & password	1	High	Bibin Francis V J, Abirami S, Kayalvizhe J, Kishankanth T G
Sprint-2	Dashboard	USN-4	User can view his/her complete medical analysis and accuracy of disease prediction	2	High	Bibin Francis V J, Abirami S, Kayalvizhe J, Kishankanth T G
Sprint-2		USN-5	User can view the accuracy of occurrence of heart disease	2	High	Bibin Francis V J, Abirami S, Kayalvizhe J, Kishankanth T G

Sprint-3	Helpdesk	USN-6	As a customer care executive, he/she can view the customer queries.	2	Medium	Bibin Francis V J, Abirami S, Kayalvizhe J, Kishankanth T G
Sprint-3		USN-7	As a customer care executive, he/she can answer the customer queries.		High	Bibin Francis V J, Abirami S, Kayalvizhe J, Kishankanth T G
Sprint-4	User Profile	USN-8	As an admin, he/she can update the health details of users.		High	Bibin Francis V J, Abirami S, Kayalvizhe J, Kishankanth T G
Sprint-4		USN-9	As an admin, he/she can add or delete users.		High	Bibin Francis V J, Abirami S, Kayalvizhe J, Kishankanth T G
Sprint-4		USN-10	As an admin, he/she can manage the user details.		High	Bibin Francis V J, Abirami S, Kayalvizhe J, Kishankanth T G

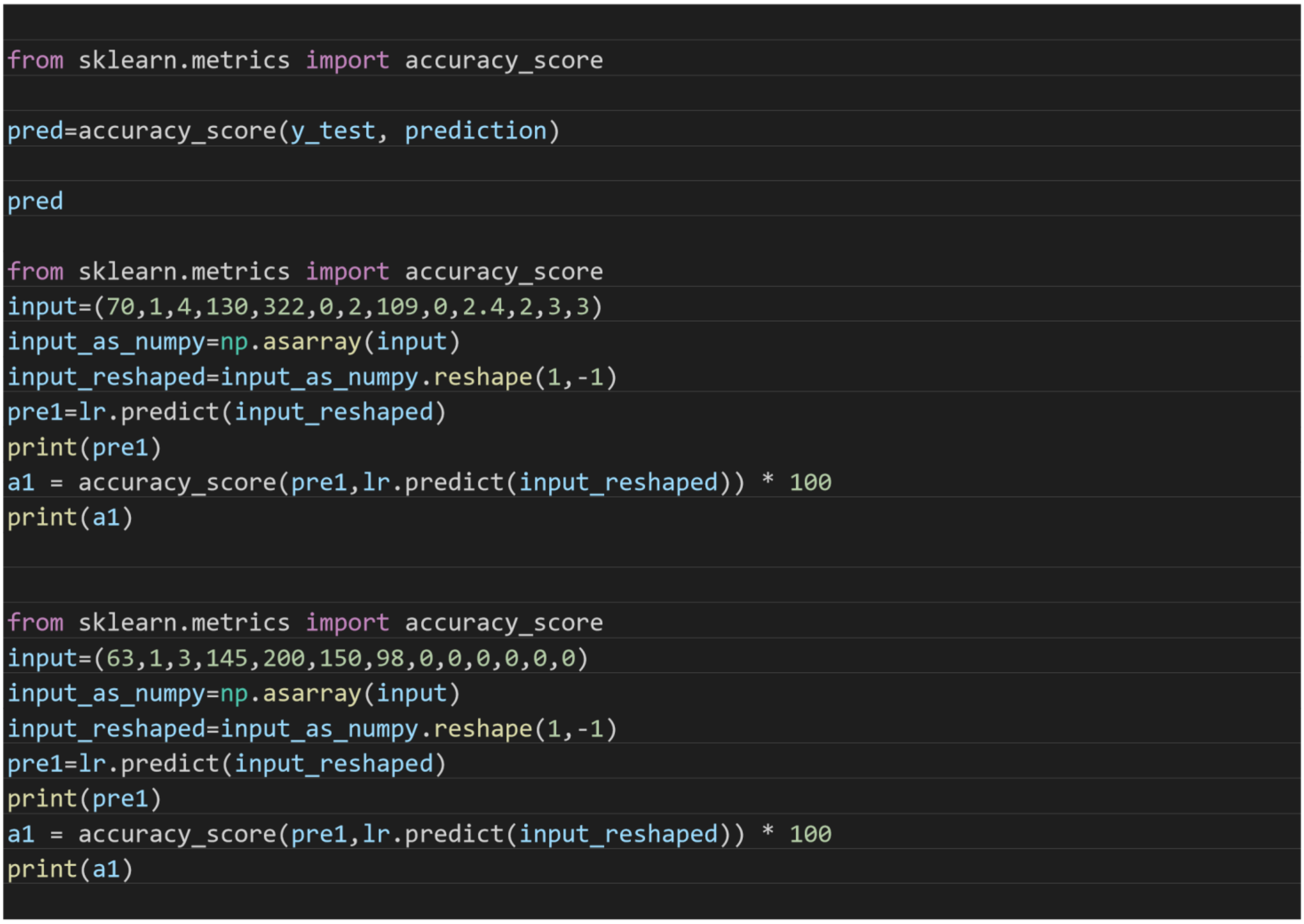
6.2 Sprint Delivery Schedule

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

7 Coding And Solutioning

7.1 Machine Learning

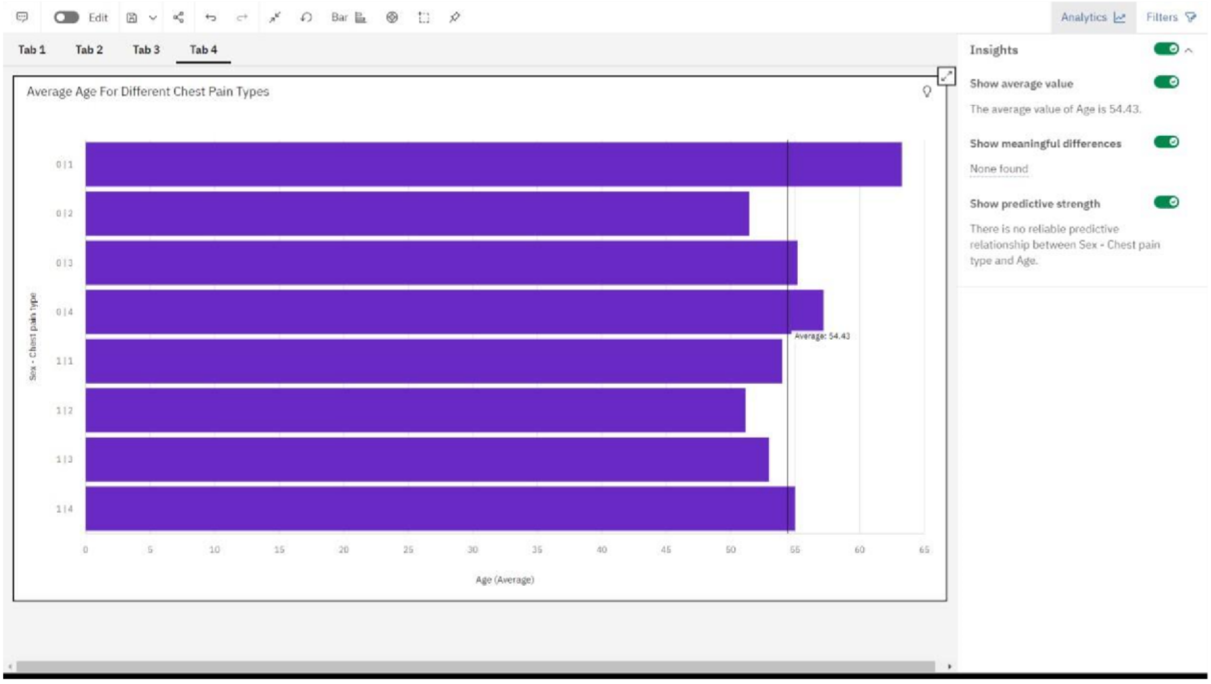
```
import numpy as np import pandas as pd import matplotlib.pyplot as plt import seaborn
as sns import warnings from sklearn.model_selection import KFold, StratifiedKFold,
cross_val_score,train_test_split from sklearn import linear_model, tree, ensemble from
sklearn.metrics import confusion_matrix,plot_confusion_matrix,classification_report
from sklearn.model_selection import GridSearchCV from sklearn.linear_model import
LogisticRegression

from sklearn.preprocessing import OneHotEncoder
 from sklearn import
svm
 dataframe=pd.read_csv("/Heart_Disease_Prediction.csv")
dataframe.head(10)

x = dataframe.iloc[ : , :-1].values y
= dataframe.iloc[ : , -1].values

X_train, X_test,y_train, y_test=train_test_split(x,y,test_size=0.25,random_state=40)

lr=LogisticRegression(max_iter=1000)

lr.fit(X_train,y_train)

prediction=lr.predict(X_test)
```



```
from sklearn.metrics import accuracy_score

pred=accuracy_score(y_test, prediction)

pred

from sklearn.metrics import accuracy_score
input=(70,1,4,130,322,0,2,109,0,2.4,2,3,3)
input_as_numpy=np.asarray(input)
input_reshaped=input_as_numpy.reshape(1,-1)
pre1=lr.predict(input_reshaped)
print(pre1)
a1 = accuracy_score(pre1,lr.predict(input_reshaped)) * 100
print(a1)


from sklearn.metrics import accuracy_score
input=(63,1,3,145,200,150,98,0,0,0,0,0,0)
input_as_numpy=np.asarray(input)
input_reshaped=input_as_numpy.reshape(1,-1)
pre1=lr.predict(input_reshaped)
print(pre1)
a1 = accuracy_score(pre1,lr.predict(input_reshaped)) * 100
print(a1)
```

7.2 Dashboard

Average age for different chest pain types

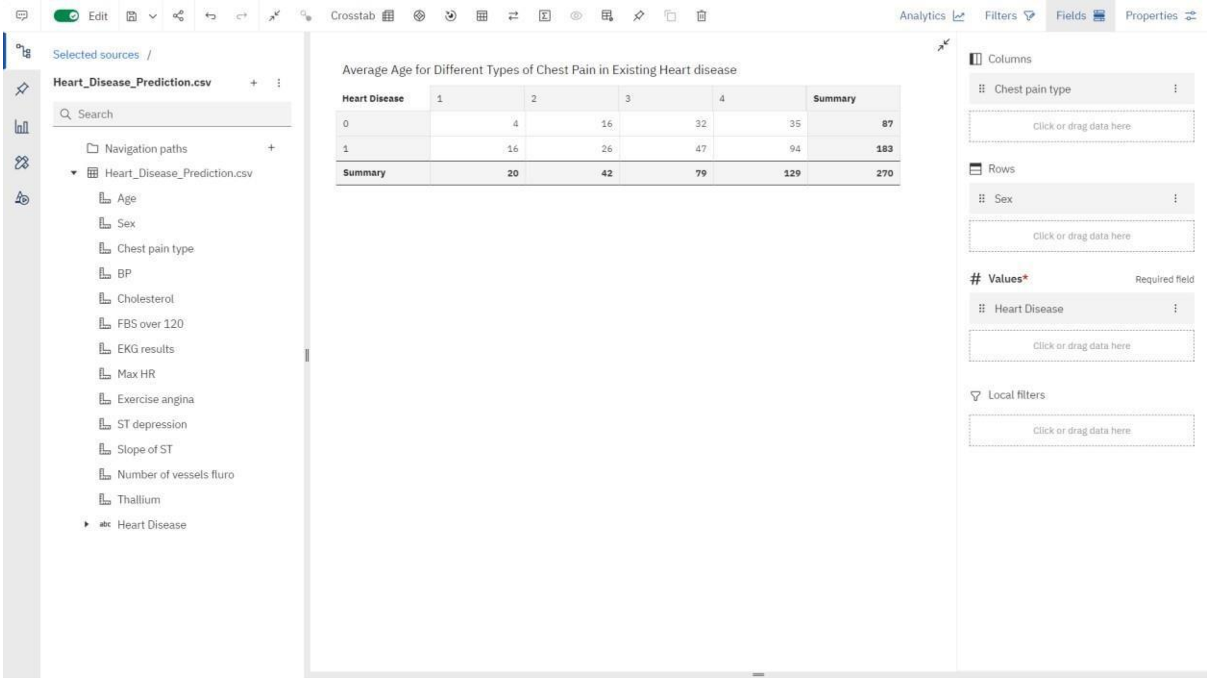


The average value of Age is 54.43.

None found

There is no reliable predictive relationship between Sex - Chest pain type and Age.

Average age for different types of chest pain in existing heart disease



Average Age for Different Types of Chest Pain in Existing Heart disease

Heart Disease	1	2	3	4	Summary
0	4	16	32	35	**87**
1	16	26	47	94	**183**
Summary	**20**	**42**	**79**	**129**	**270**

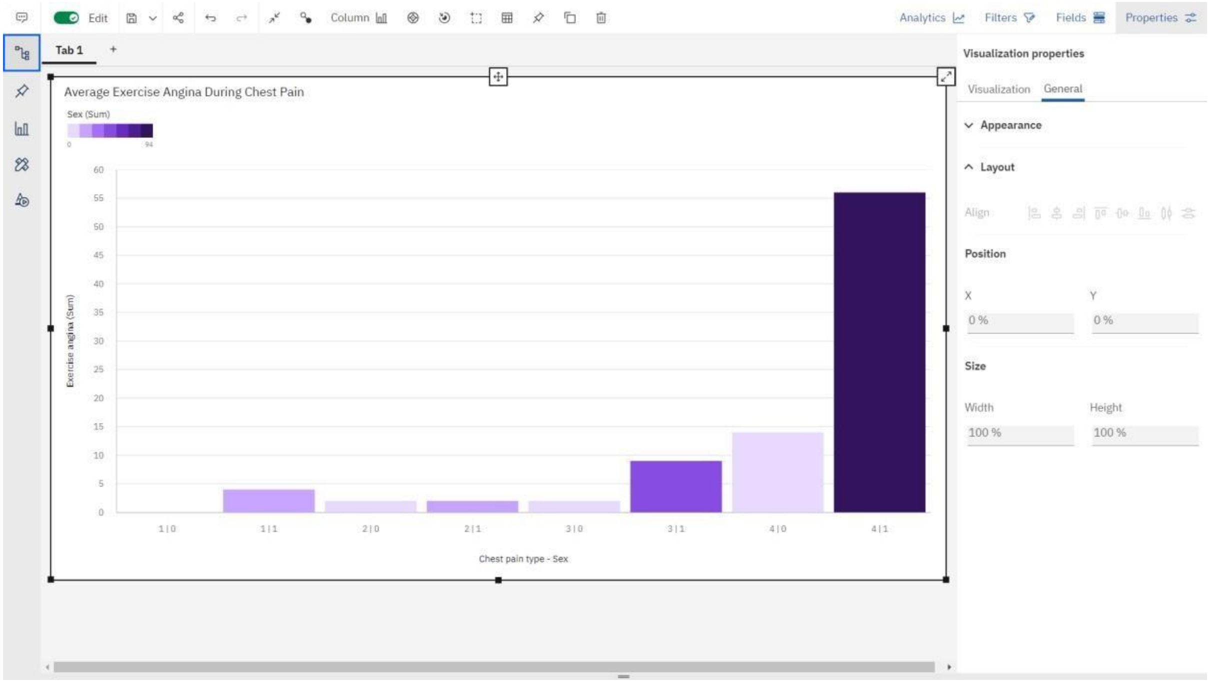
Columns: Chest pain type

Rows: Sex

Values*: Heart Disease

Local filters

Average exercise angina during chest pain



Position

X: 0 %

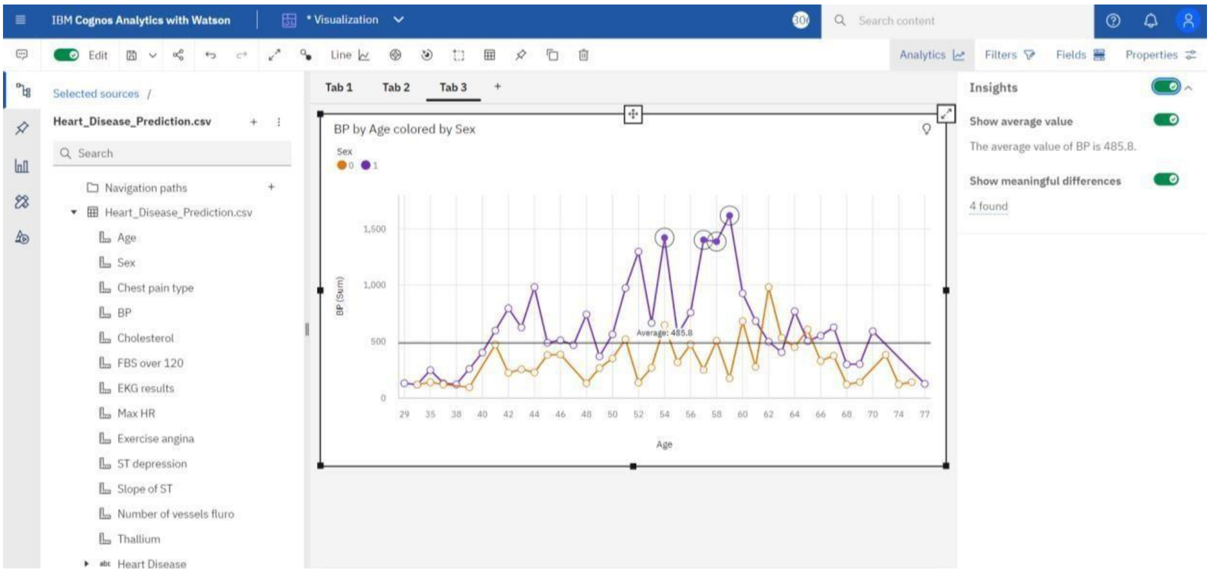
Y: 0 %

Size

Width: 100 %

Height: 100 %

BP variation with respect to age



Insights

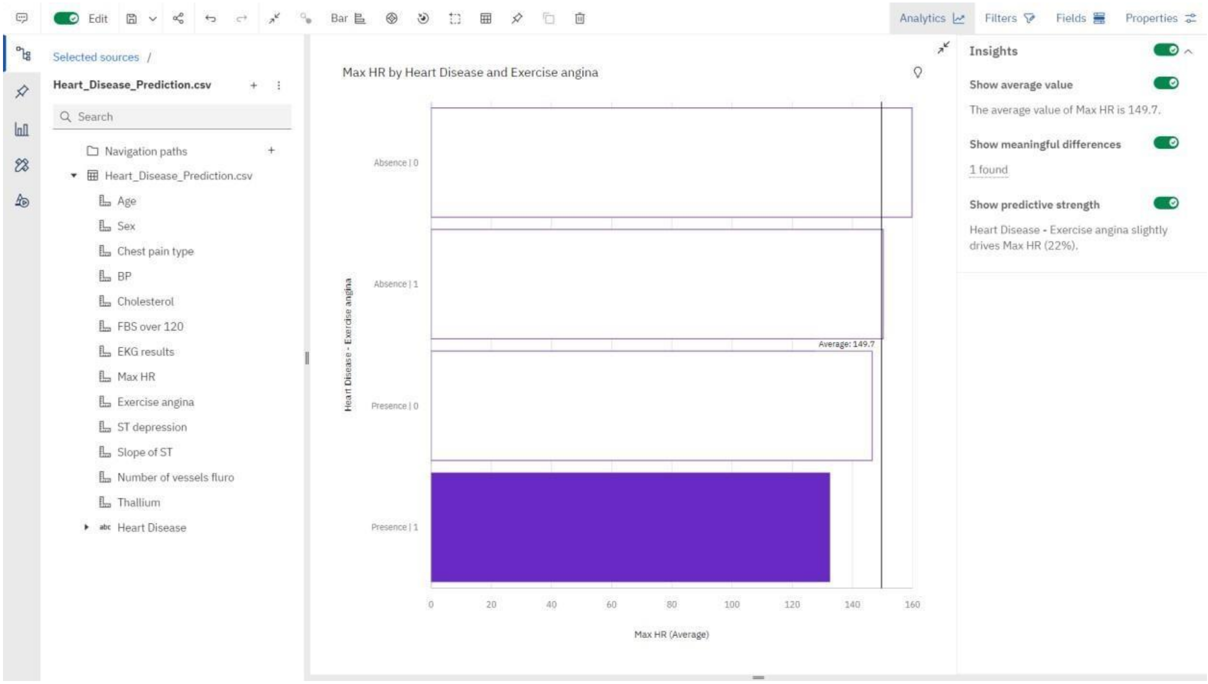
Show average value

The average value of BP is 485.8.

Show meaningful differences

4 found

Effect of existing heart disease on average of exercise angina



Insights

Show average value

The average value of Max HR is 149.7.

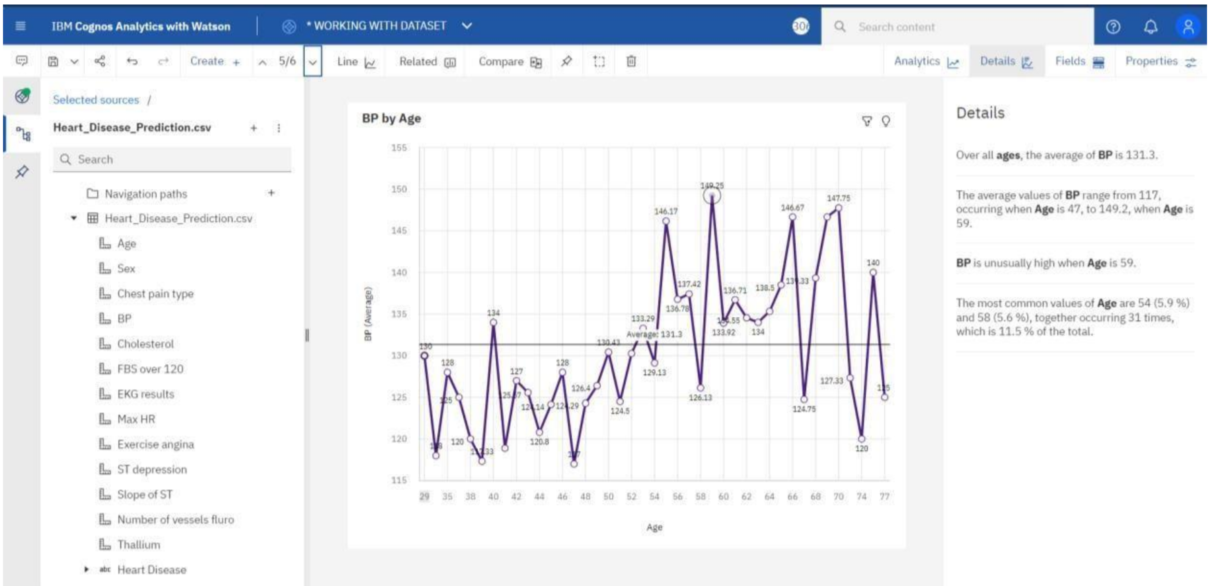
Show meaningful differences

1 found

Show predictive strength

Heart Disease - Exercise angina slightly drives Max HR (22%).

Blood pressure variation by age



Details

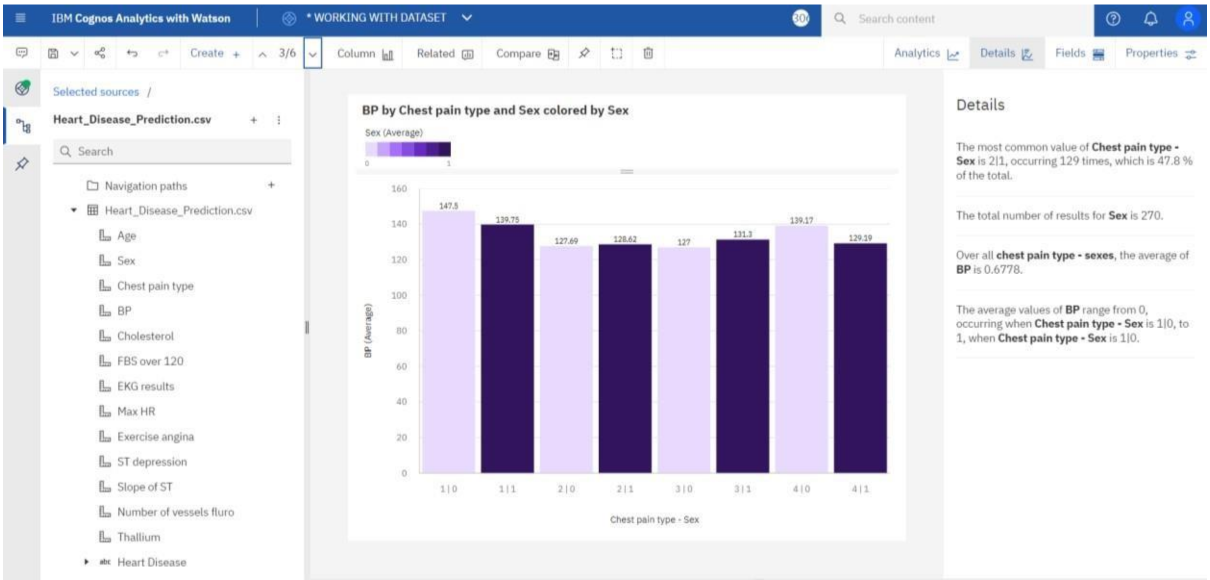
Over all **ages**, the average of **BP** is 131.3.

The average values of **BP** range from 117, occurring when **Age** is 47, to 149.2, when **Age** is 59.

BP is unusually high when **Age** is 59.

The most common values of **Age** are 54 (5.9 %) and 58 (5.6 %), together occurring 31 times, which is 11.5 % of the total.

Average BP during the chest pain



Details

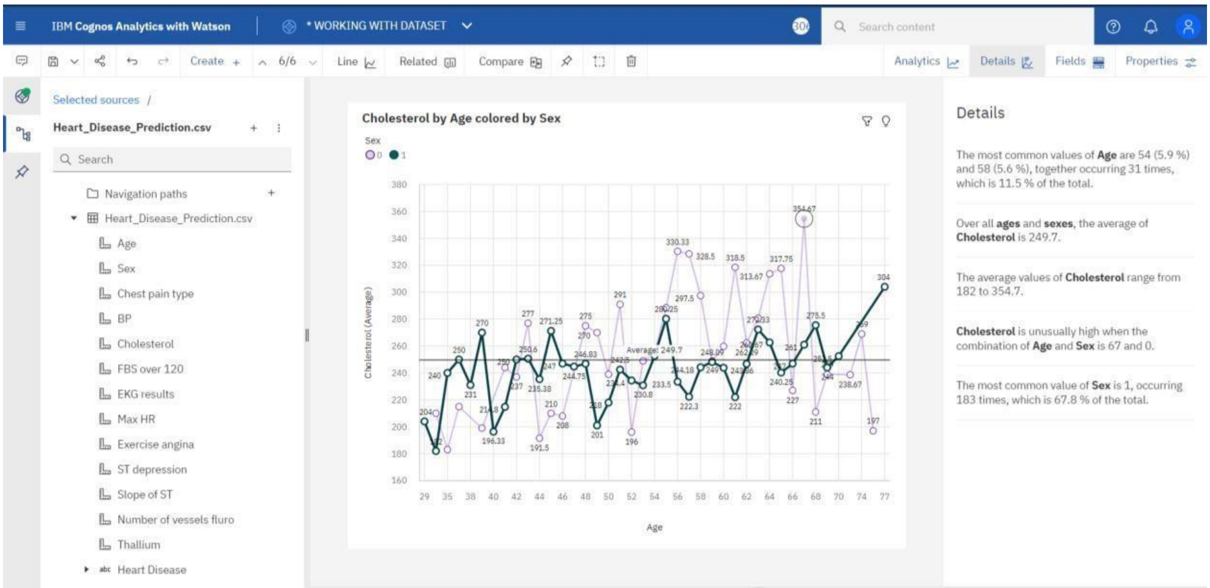
The most common value of **Chest pain type - Sex** is 2|1, occurring 129 times, which is 47.8 % of the total.

The total number of results for **Sex** is 270.

Over all **chest pain type - sexes**, the average of **BP** is 0.6778.

The average values of **BP** range from 0, occurring when **Chest pain type - Sex** is 1|0, to 1, when **Chest pain type - Sex** is 1|0.

Exploration Of cholesterol by age and gender



Details

The most common values of **Age** are 54 (5.9 %) and 58 (5.6 %), together occurring 31 times, which is 11.5 % of the total.

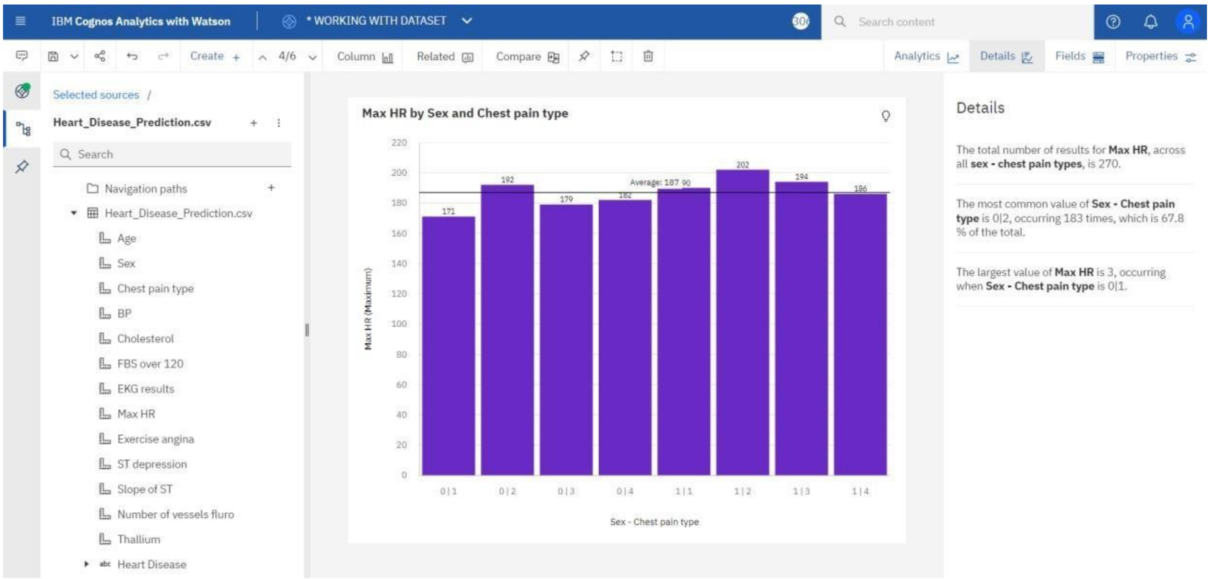
Over all **ages** and **sexes**, the average of **Cholesterol** is 249.7.

The average values of **Cholesterol** range from 182 to 354.7.

Cholesterol is unusually high when the combination of **Age** and **Sex** is 67 and 0.

The most common value of **Sex** is 1, occurring 183 times, which is 67.8 % of the total.

Average Max Heart Rate achieved during chest pain



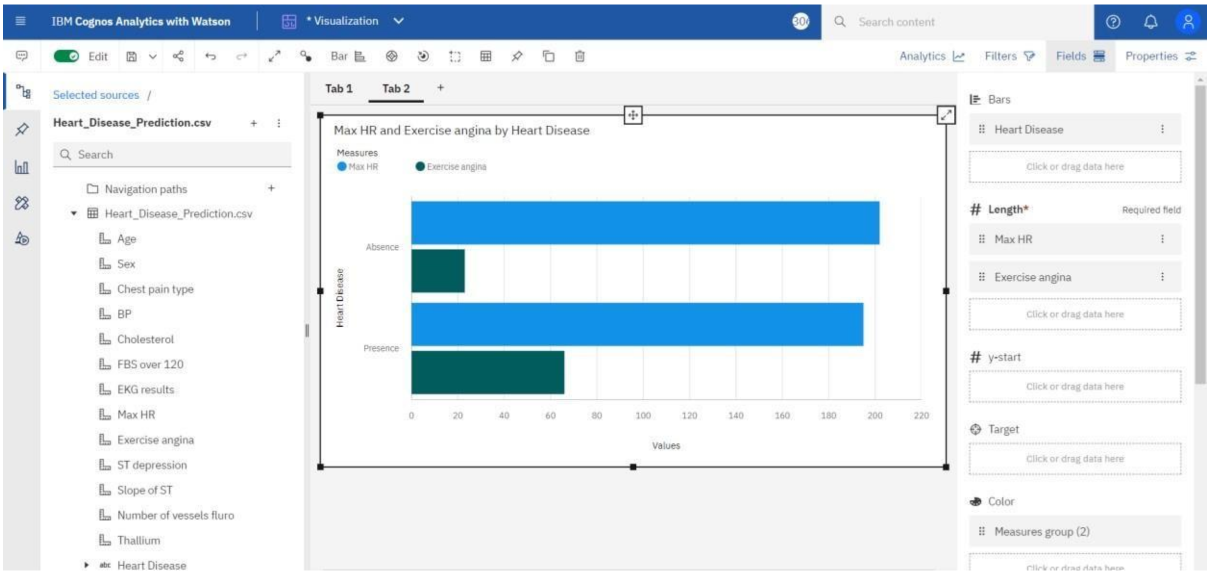
Details

The total number of results for **Max HR**, across all **sex - chest pain types**, is 270.

The most common value of **Sex - Chest pain type** is 0|2, occurring 183 times, which is 67.8 % of the total.

The largest value of **Max HR** is 3, occurring when **Sex - Chest pain type** is 0|1.

Maximum heart rate in existing heart disease by exercise angina



Max HR and Exercise angina by Heart Disease

Bars: Heart Disease

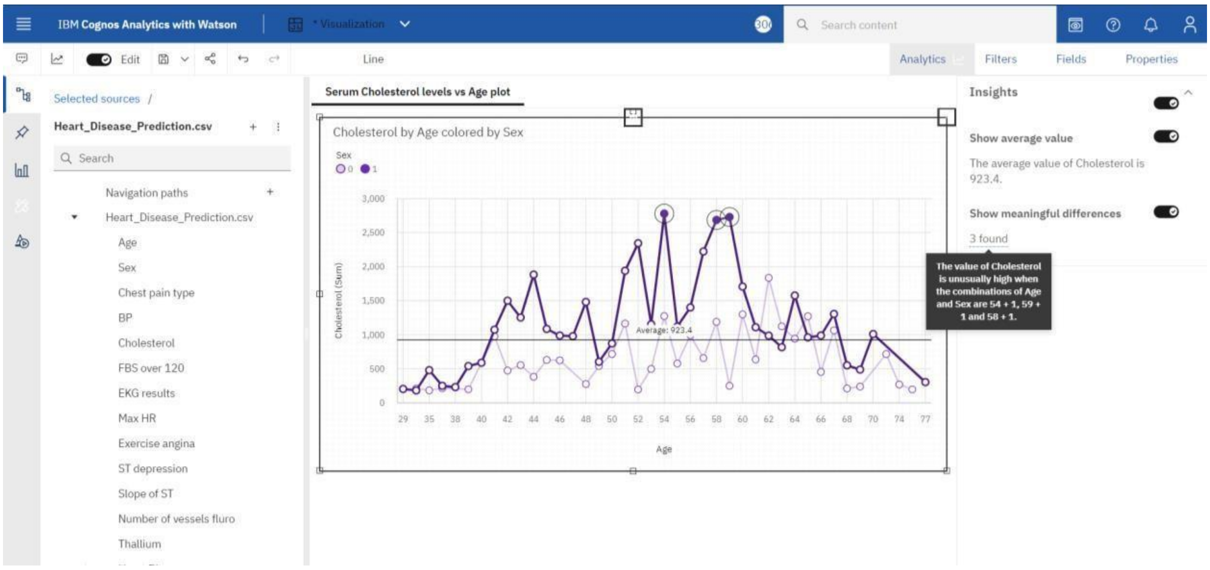
Length*: Max HR; Exercise angina

y-start

Target

Color: Measures group (2)

Serum cholesterol levels vs age plot



Cholesterol by Age colored by Sex

Insights

Show average value

The average value of Cholesterol is 923.4.

Show meaningful differences

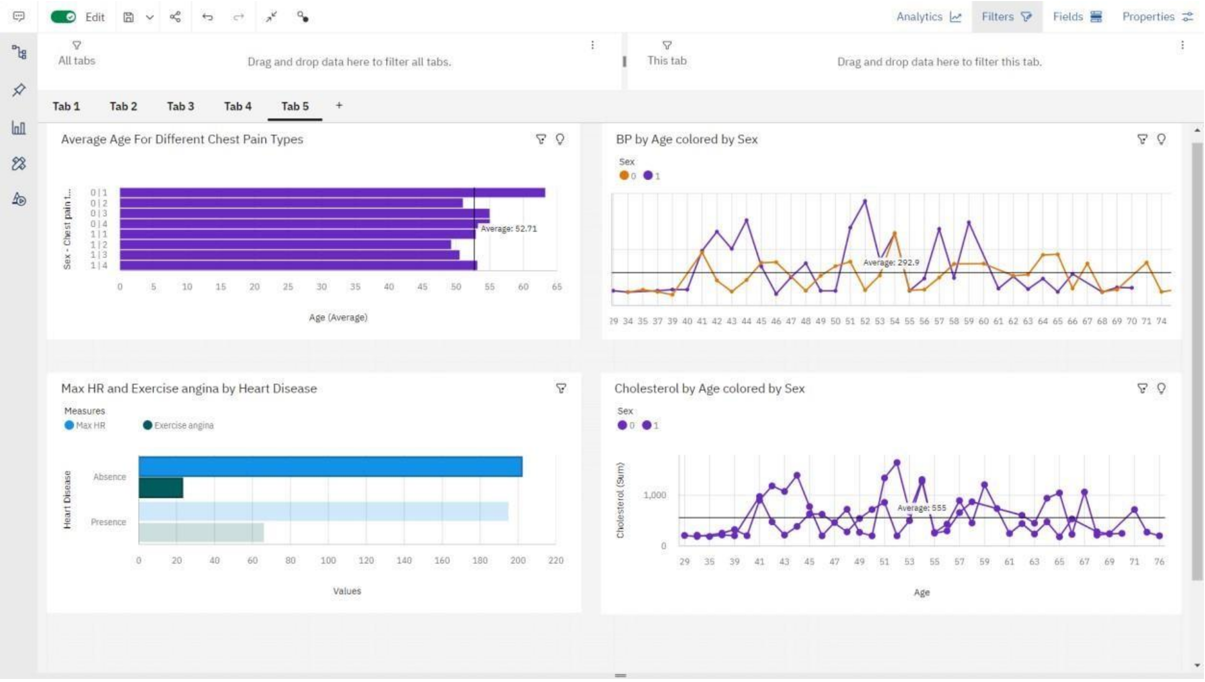
3 found

The value of Cholesterol is unusually high when the combinations of Age and Sex are 54 + 1, 59 + 1 and 58 + 1.

Dashboard showing different types of visuals

Dashboard presenting the following visuals,

1. Average Age for Different Chest Pain Types
2. Exploration Of BP By Age
3. Max Heart Rate vs exercise Angina in case of Existing Heart Disease.
4. Serum Cholesterol levels vs Age in both the Gender people.



7.3 Web page

Dashboard.html

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Heart Disease Prediction</title>
    <link href="static/style_main.css" rel="stylesheet" type="text/css" />
</head>
<body>
    <div>
    <center><iframe
```

```
src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&amp;pathRef=.my_folders
%2FVisualization&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_na
vbar=false&amp;shareMode=embedded&amp;action=view&amp;mode=dashboard&amp;su
bView=model0000018441e02e59_00000000" width="900" height="650" frameborder="0"
gesture="media" allow="encrypted-media" allowfullscreen=""></iframe></center>     </div>
</body>
</html>
```

Result.html

```
<!DOCTYPE html>
<html>
  <head>
    <meta charset="utf-8">
    <meta name="viewport" content="width=device-width">
    <title>Result</title>
    <link href="static/style.css" rel="stylesheet" type="text/css" />
</head>
  <body>
    <a class="btn" href="/back">Back</a>
    <h1 class=heading>{{result}}</h1>
    <label class=side>{{tip}}</label>
    <form action="/dashboard" enctype="multipart/form-data">
        <input type="submit" class="btn2 btn-primary btn-block" value="View Dashboard">

</form>
  </body>
</html>
```

Style.css body

```
{     margin:
0px;
    background-image: url("background.jpg");
background-size: cover;     padding: 0px;
background-position: center;
    background-repeat: no-repeat;
      }
.heading{
      position: relative;
text-align: center;
top:50px;
left:30px;       color:
white;
  }
```

```
.side{
     position: relative;
     text-align: center;
top:110px;
left:450px;      color:
white;
  }
.btn{      width: 100px;
height: 44px;      position:
relative;      top:20px;
left:1200px;      display:
flex;      justify-content:
center;      align-items:
center;
     color: white;
     background-color: #1db096;      border-
radius: 20px;
     box-shadow: 5px 10px 30px rgba(24,139,119,0.2);
  }
.btn:hover{      background-
color: #23cdaf;      transition:
all ease 0.2s;
     box-shadow: 0 12px 16px 0 rgba(0,0,0,0.24), 0 17px 50px 0 rgba(0,0,0,0.19); }
.btn2{
          width: 190px;
height: 44px;           position:
relative;           top:440px;
left:585px;           display:
flex;           justify-content:
center;           align-items:
center;           color: white;
          background-color: #1db096;
border-radius: 10px;
          box-shadow: 5px 10px 30px rgba(24,139,119,0.2);
        }
     .btn2:hover{
background-color: #23cdaf;
transition: all ease 0.2s;
          box-shadow: 0 12px 16px 0 rgba(0,0,0,0.24), 0 17px 50px 0 rgba(0,0,0,0.19); }
```

Style_main.css

```
body {
margin: 0px;
   background-image: url("background.jpg");
   background-size: cover;
padding: 0px;     background-
position: center;
   background-repeat: no-repeat;
     }

App.py import
numpy as np import
pickle import
sklearn
from flask import Flask, render_template, request, redirect, url_for, flash

app = Flask(__name__)
model = pickle.load(open('heart.pkl', 'rb'))
@app.route('/') def index():     return
render_template('index.html')

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

@app.route('/back') def back():     return
render_template('index.html')
@app.route('/predict', methods =['GET',
'POST']) def predict():   if request.method ==
'POST':
   features = [float(i) for i in request.form.values()]
array_features = [np.array(features)]     prediction
= model.predict(array_features)     output =
prediction     print(output)     result=""     tip1=""
tip2="     tip3=""     if output == 1:
      result = "You are not likely to have heart disease!"
tip1=" Have a great day!"
      tip2=" "         tip3="
"
final=tip1+tip2+tip3
else:

      result="Hey there! You are likely to have heart disease "
tip1="Eat a heart-healthy diet, "
      tip2="Maintain a healthy weight and "         tip3="Get
good quality sleep"         final=tip1+tip2+tip3     return
render_template("result.html",result=result,tip=final) if
__name__ == '__main__':
 app.run(debug=True)
```

index.html

```
 <html>
<head>
<!-- Bootstrap CSS --> <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>
<title>Heart Disease Test</title>
</head>
<body >
<!-- Java Script -->
<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>
<!-- Navbar-->
<nav class="navbar navbar-dark" style="background-color: rgb(13, 102, 87);">
<span class="navbar-brand mb-0 h1">Heart Disease Prediction</span> </nav>
<div class="container">
<br>
```

```
<!--Form-->
<form action = "/predict" method ="POST" >
<fieldset>
<legend style="color: rgb(41, 15, 134);"><b>Predict your heart disease possibility
here!</b></legend><br>
<div class="card card-body" style="background-color: pink;">
<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" 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> -- Select an Option -- </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> -- Select an Option -- </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">Resting Blood Pressure in mm Hg</label>
<input type="number" class="form-control" id="trestbps" name="trestbps"  required> </div>
<div class="col-sm">
<label for="chol">Serum Cholestoral in mg/dl</label>
<input type="number" class="form-control" id="chol" name="chol" required> </div>
<div class="col-sm">
<label for="fbs">Fasting Blood Sugar > 120 mg/dl</label>
<select class="form-control" id="fbs" name="fbs" required>
<option disabled selected value> -- Select an Option -- </option>
<option value = "0">False</option>
```

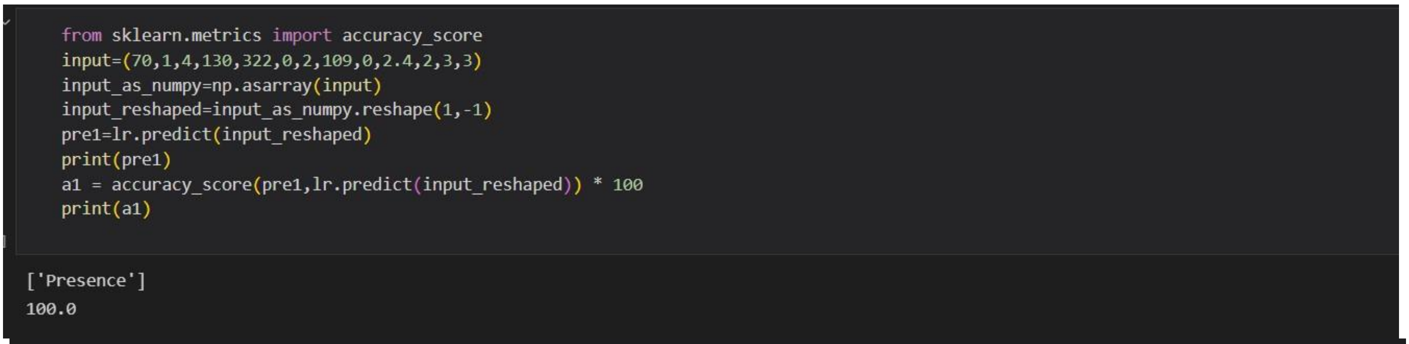
```
<option value = "1">True</option>
</select>
</div>
<br>
<div class="form-group row">
<div class="col-sm">
<label for="restecg">Resting ECG Results </label>
<select class="form-control" id="restecg" name="restecg" required>
<option disabled selected value> -- Select an Option -- </option>
<option value = "0">Normal  </option>
<option value = "1">Having ST-T wave abnormality  </option>
<option value = "2">Probable or definite left ventricular hypertrophy</option> </select>
</div>
<div class="col-sm">
<label for="thalach">Maximum Heart Rate</label>
<input type="number" class="form-control" id="thalach" name="thalach" required> </div>
<div class="col-sm">
<label for="exang">Exercise Induced Angina </label>
<select class="form-control" id="exang" name="exang" required>
<option disabled selected value> -- Select an Option -- </option>
<option value = "0">No</option>
<option value = "1">Yes</option>
</select>
</div>
<div class="col-sm">
<label for="oldpeak">ST Depression Induced</label>
<input type="number" step="any" class="form-control" id="oldpeak" name="oldpeak"
required> </div>
</div>
<br>
<div class="form-group  row">
<div class="col-sm">
<label for="slope">Slope of the Peak Exercise ST Segment </label>
<select class="form-control" id="slope" name="slope" required> <option
disabled selected value> -- Select an Option -- </option>
<option value = "1">Upsloping</option>
<option value = "2">Flat</option>
<option value = "3">Downsloping</option>
</select>
</div>
<div class="col-sm">
<label for="ca">Number of Vessels Colored by Flourosopy</label>
```

```
<select class="form-control" id="ca" name = "ca" required>
<option disabled selected value> -- Select an Option -- </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">Thalassemia</label>
<select class="form-control" id="thal" name = "thal" required>
<option disabled selected value> -- Select an Option -- </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-primary" type="submit" value="Predict">
</div>
<!--Prediction Result-->
<div id ="result">
<strong style="color:red">{{result}}</strong>
</div>
</div>
</fieldset>
</form>
</div>
</body>
</html>
```

8. Testing

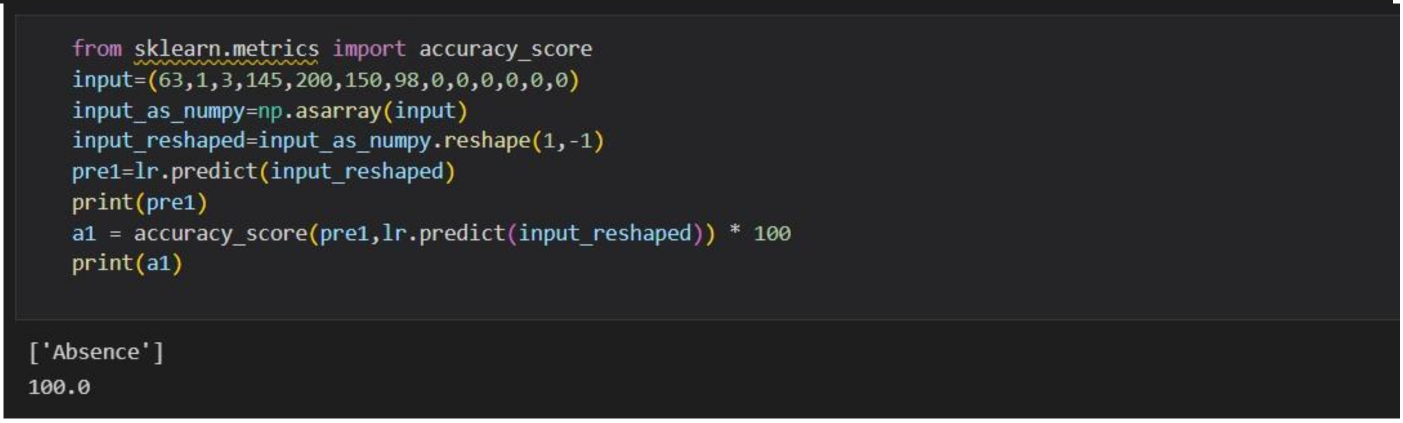
8.1 Test Cases

Testing the data model for various input values.



```
from sklearn.metrics import accuracy_score
input=(70,1,4,130,322,0,2,109,0,2.4,2,3,3)
input_as_numpy=np.asarray(input)
input_reshaped=input_as_numpy.reshape(1,-1)
pre1=lr.predict(input_reshaped)
print(pre1)
a1 = accuracy_score(pre1,lr.predict(input_reshaped)) * 100
print(a1)
```

```
['Presence']
100.0
```

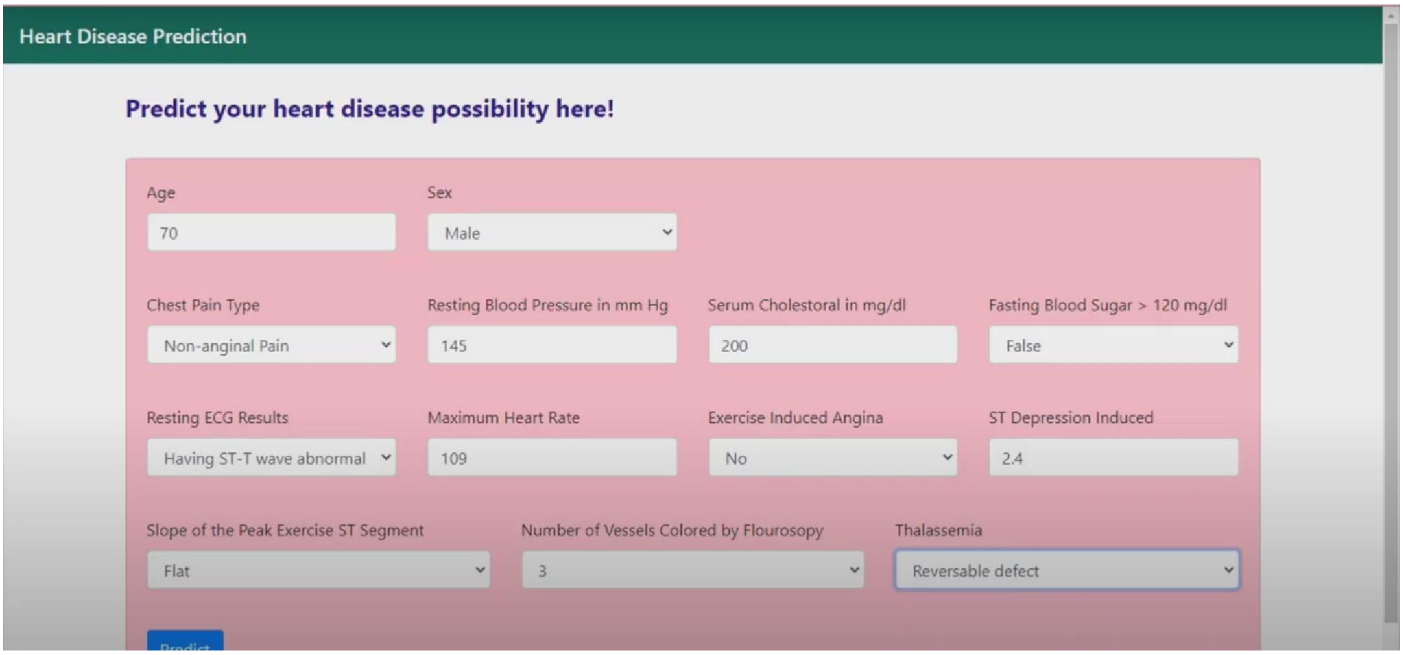


```
from sklearn.metrics import accuracy_score
input=(63,1,3,145,200,150,98,0,0,0,0,0,0)
input_as_numpy=np.asarray(input)
input_reshaped=input_as_numpy.reshape(1,-1)
pre1=lr.predict(input_reshaped)
print(pre1)
a1 = accuracy_score(pre1,lr.predict(input_reshaped)) * 100
print(a1)
```

```
['Absence']
100.0
```

8.2 User acceptance Testing

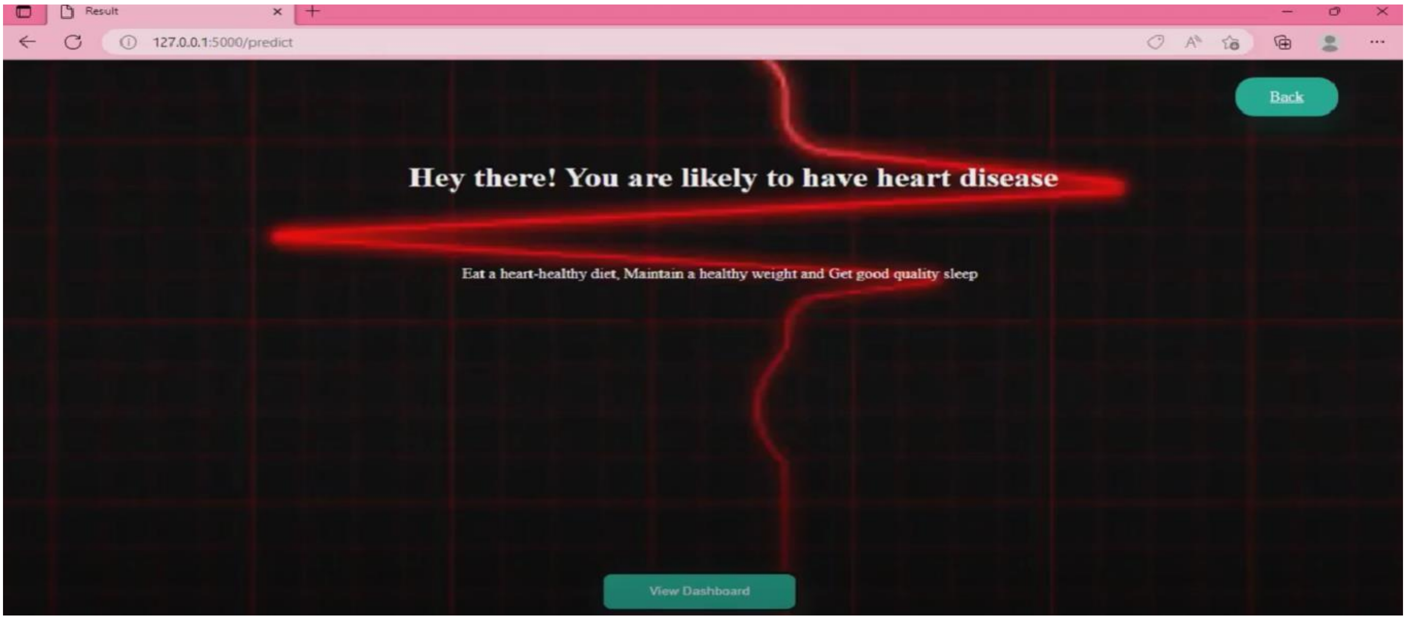
Testing a case where user has heart disease



Heart Disease Prediction

Predict your heart disease possibility here!

Age: 70	Sex: Male		
Chest Pain Type: Non-anginal Pain	Resting Blood Pressure in mm Hg: 145	Serum Cholestoral in mg/dl: 200	Fasting Blood Sugar > 120 mg/dl: False
Resting ECG Results: Having ST-T wave abnormal	Maximum Heart Rate: 109	Exercise Induced Angina: No	ST Depression Induced: 2.4
Slope of the Peak Exercise ST Segment: Flat	Number of Vessels Colored by Flourosopy: 3	Thalassemia: Reversable defect	



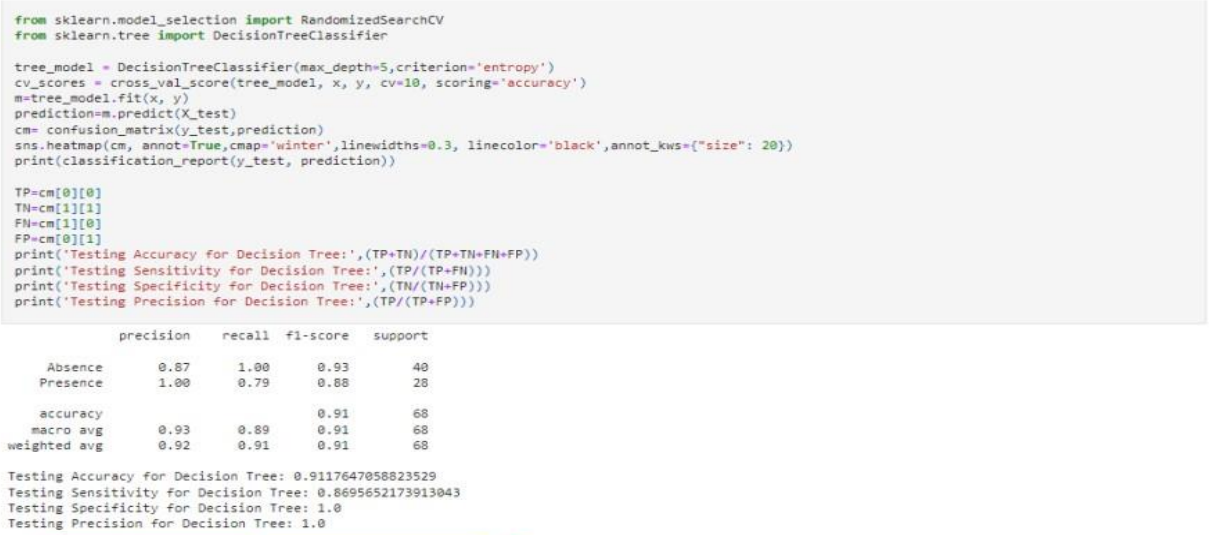
Hey there! You are likely to have heart disease

Eat a heart-healthy diet, Maintain a healthy weight and Get good quality sleep

9. Result

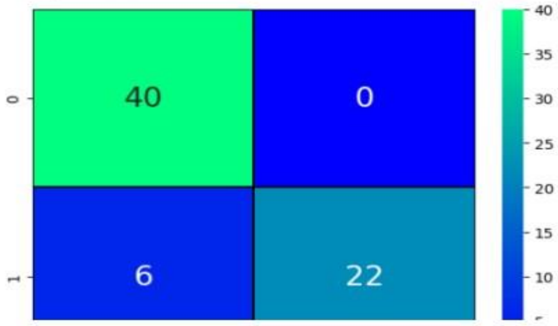
9.1 Performance Metrics

The confusion matrix below shows the performance metrics of the machine learning model.



```python
from sklearn.model_selection import RandomizedSearchCV
from sklearn.tree import DecisionTreeClassifier

tree_model = DecisionTreeClassifier(max_depth=5,criterion='entropy')
cv_scores = cross_val_score(tree_model, x, y, cv=10, scoring='accuracy')
m=tree_model.fit(x, y)
prediction=m.predict(X_test)
cm= confusion_matrix(y_test,prediction)
sns.heatmap(cm, annot=True,cmap='winter',linewidths=0.3, linecolor='black',annot_kws={"size": 20})
print(classification_report(y_test, prediction))

TP=cm[0][0]
TN=cm[1][1]
FN=cm[1][0]
FP=cm[0][1]
print('Testing Accuracy for Decision Tree:',(TP+TN)/(TP+TN+FN+FP))
print('Testing Sensitivity for Decision Tree:',(TP/(TP+FN)))
print('Testing Specificity for Decision Tree:',(TN/(TN+FP)))
print('Testing Precision for Decision Tree:',(TP/(TP+FP)))
```

```
              precision    recall  f1-score   support

     Absence       0.87      1.00      0.93        40
    Presence       1.00      0.79      0.88        28

    accuracy                           0.91        68
   macro avg       0.93      0.89      0.91        68
weighted avg       0.92      0.91      0.91        68

Testing Accuracy for Decision Tree: 0.9117647058823529
Testing Sensitivity for Decision Tree: 0.8695652173913043
Testing Specificity for Decision Tree: 1.0
Testing Precision for Decision Tree: 1.0
```



10. Advantages Disadvantages

Advantages:

- This is one of the fastest ways to determine if a person is likely to suffer from a heart disease or not.
- Useful for medical practitioners to easily classify their patients.
- User Friendly
- Easy to understand
- Dashboard provides insightful information.

Disadvantages:

- Needs work
- Users need to know all the fields
- Does Not take null value as input

11. Conclusion

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 stages of heart disease.

12. 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. To use this website we need to take a lot of tests beforehand. So it would be better if we require less attributes and still give an effective result