VISUALIZING AND PREDICTING HEART DISEASES WITH AN INTERACTIVE DASH BOARD

NALAIYA THIRAN PROJECT REPORT-2022

Team Details	
Team ID	PNT2022TMID34621
Team Leader	Bibishek G S Steephen
Team Member 1	Bibin S
Team Member 2	Ajay A J
Team Member 3	Gino M
Team Member 4	Ajay E
Faculty Mentor	Ajin M

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.

The application is fed with various details and the heart disease associated with those details. The application allows user to share their heart related issues. It then processes user specific details to check for various illness that could be associated with it. Here we use some intelligent data mining techniques to guess the most accurate illness that could be associated with patient's details. Based on result, the person can contact doctor accordingly for further treatment. The system allows user to view doctor's details too. The system can be used for free heart disease consulting online.

1.2 Purpose

The main goal of this project is to provide a tool for doctors to detect heart disease as early stage. This in turn will help to provide effective treatment to patients and avoid severe consequences. 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 analyzing 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 analyzing data.

2.2 References

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

classifier performances and evaluation metrics. In the result section, the visualized data shows that the prediction is accurate.

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

2.3 Problem Statement Definition

Who does the problem affect?

People with unhealthy lifestyles, stress, depression, age above 40 and when their ancestors got heart disease (since heart disease is hereditary).

When does the issue occur?

The issue occurs for people with unhealthy lifestyles and age above 40.

Where is the issue occurring?

The issue is originating from an unhealthy lifestyle. It mostly occurs in the blood valves of the heart.

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

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

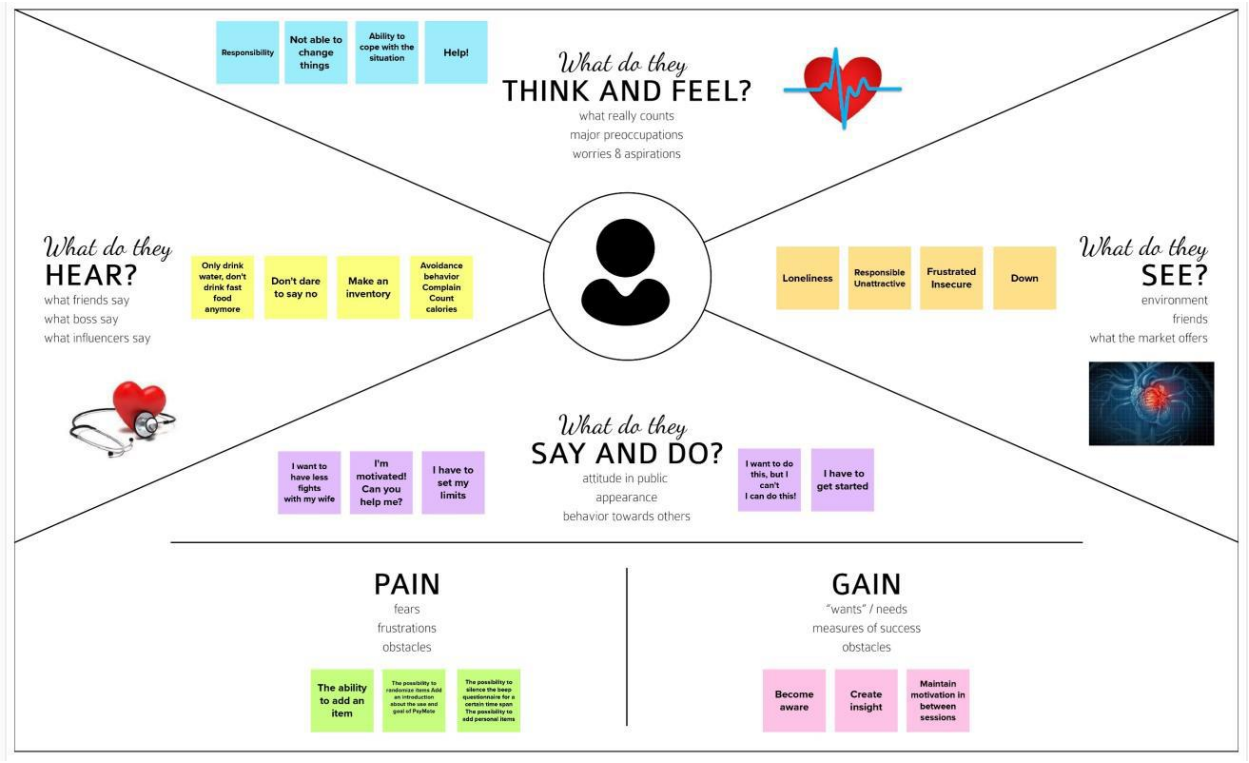
Why is it important to fix the problem?

We should predict the problem before giving treatment to the patients. As the problem is predicted early, we can solve it easily and early.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

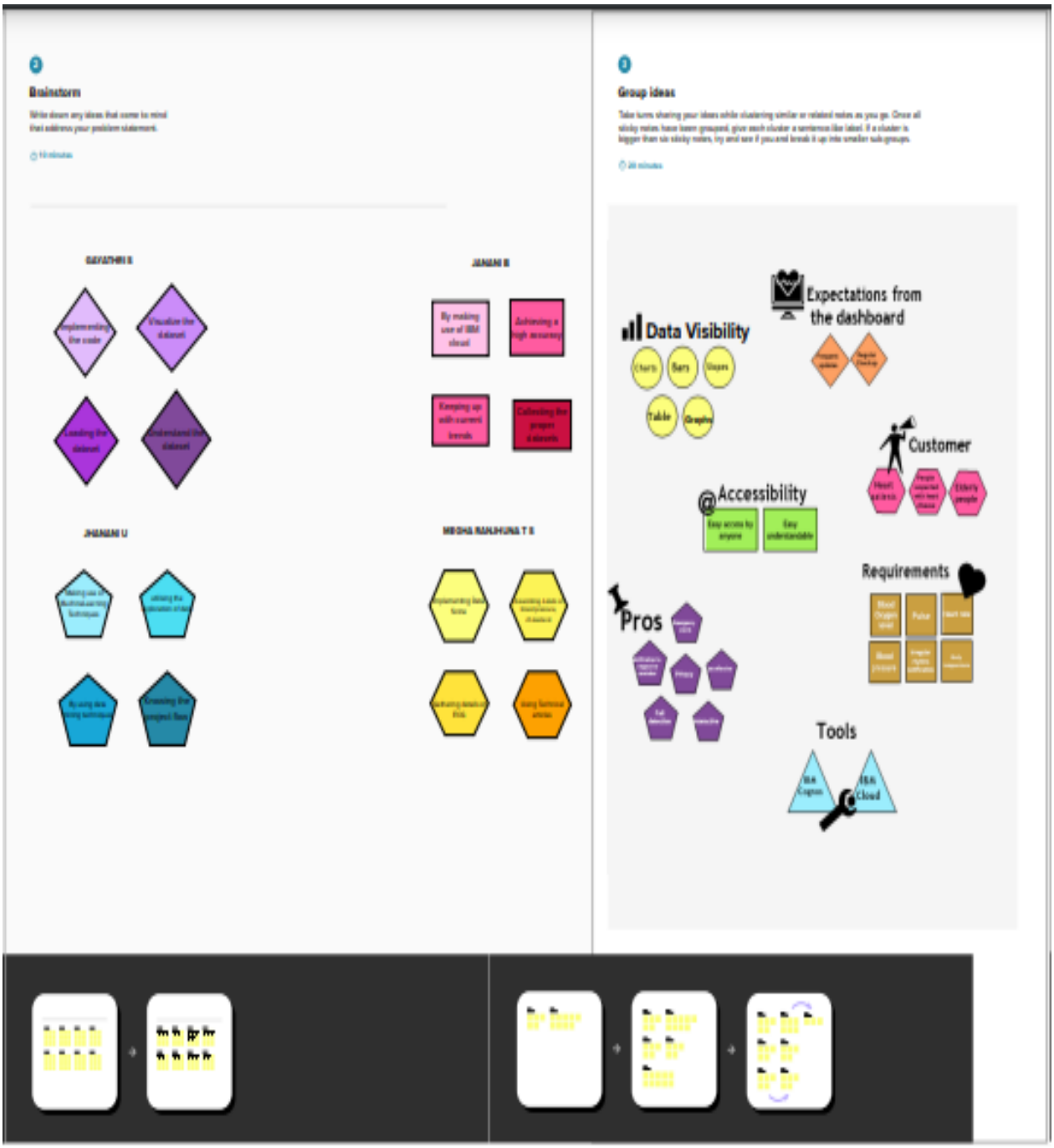
Empathy map



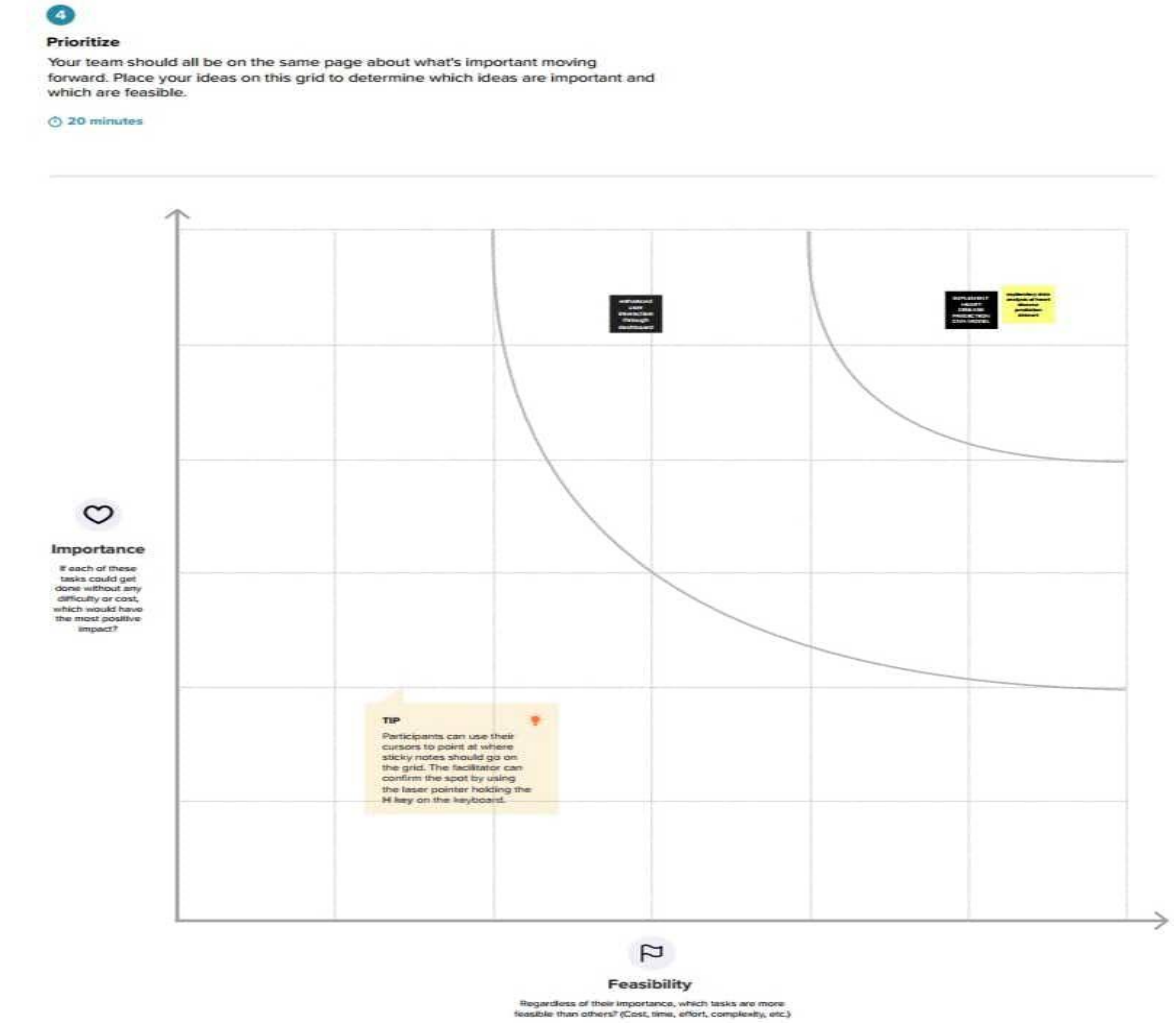
3.2 Ideation & Brainstorming

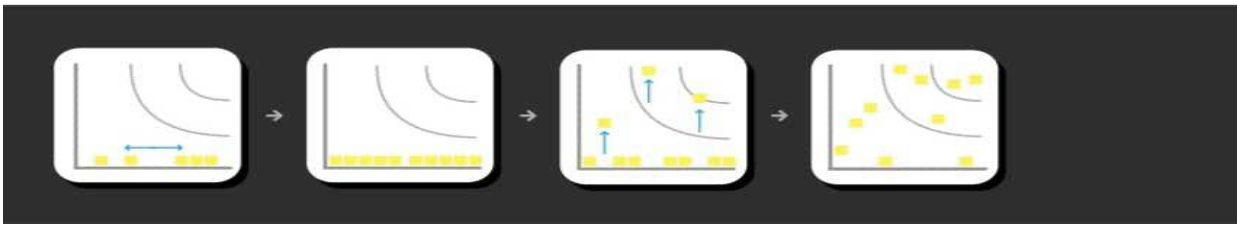
Step-1: Team Gathering, Collaboration and Select the Problem Statement

Step-2: Brainstorm, Idea Listing and Grouping



Step-3: Idea Prioritization



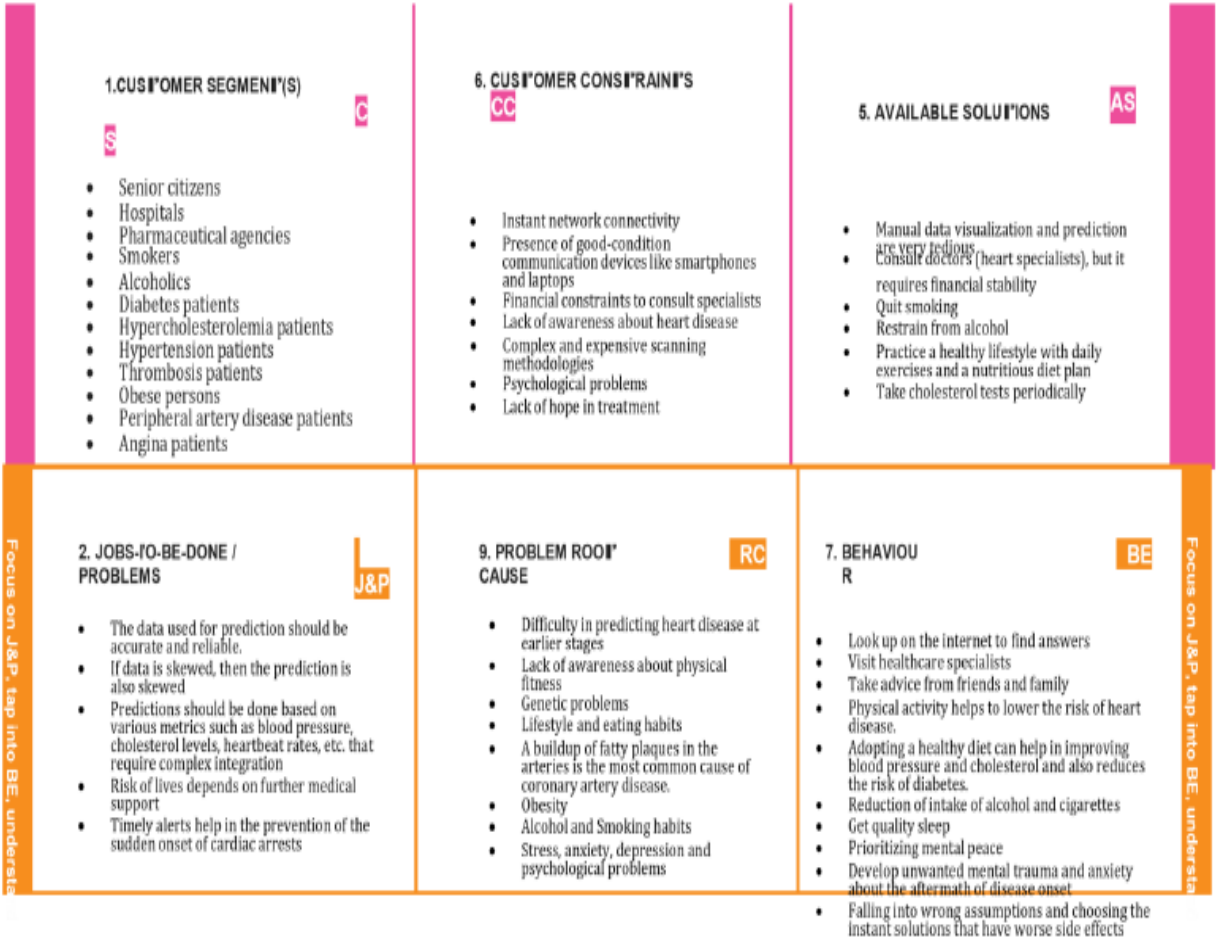


3.3 Proposed Solution

.

- The prime novelty of the solution is the fusion of highly efficient algorithms, that eliminates the disadvantage of every algorithm when employed individually and also provides a higher level of accuracy in the prediction. Another innovation is employed in the dashboard by providing diet and fitness related suggestions to the user based on his/her medical reports and history. In addition to it, the patient is given a list of hospitals closer to the patient's locality and severity of the disease.
- It helps with disease prediction at an early stage and alerts the user about his/her current health status. Heart disease can be cured by a mix of medication, lifestyle modifications, and occasionally, surgery. The system helps the user as well as the doctor to make better decisions. Complex questions related to heart diseases can be answered by extracting hidden knowledge, i.e., patterns and relationships from the heart disease database.
- This interactive dashboard for heart disease prediction can be installed in hospitals and healthcare facilities. Predicted outcomes can be utilized to avoid expensive surgeries.
- It can be used in educational institutions, industries and all types of workplaces to monitor the employees' health conditions and thereby helping them lead a healthier life.
- The proposed solution works efficiently in both smaller and larger datasets.
- This predictive model can be used to detect diseases in other internal organs too.

3.4 Problem Solution Fit



4. REQUIREMENT ANALYSIS

4.1 Functional Requirement

User Registration:

Enables user to make registration for the application through Gmail.

User Confirmation:

Once after registration, the user will get confirmation via., Email.

Visualizing Data:

User can visualize the trends on the heart disease through Dashboard created using IBM Cognos Analytics.

Generation Report:

User can view his/her health report and can make decisions accordingly.

4.2 Non-Functional Requirements

Usability:

Users can access the application with a minimal spec Devices and no additional dependencies are required to access the application.

Security:

The cloud vendor provides security to the deployed application.

Reliability:

The User's credentials are protected by the secured database provided by the cloud vendor. So, the customers can trust the application.

Performance:

Since the application is deployed in the Docker container, Docker provides a smooth performance to the clients/customers.

Availability:

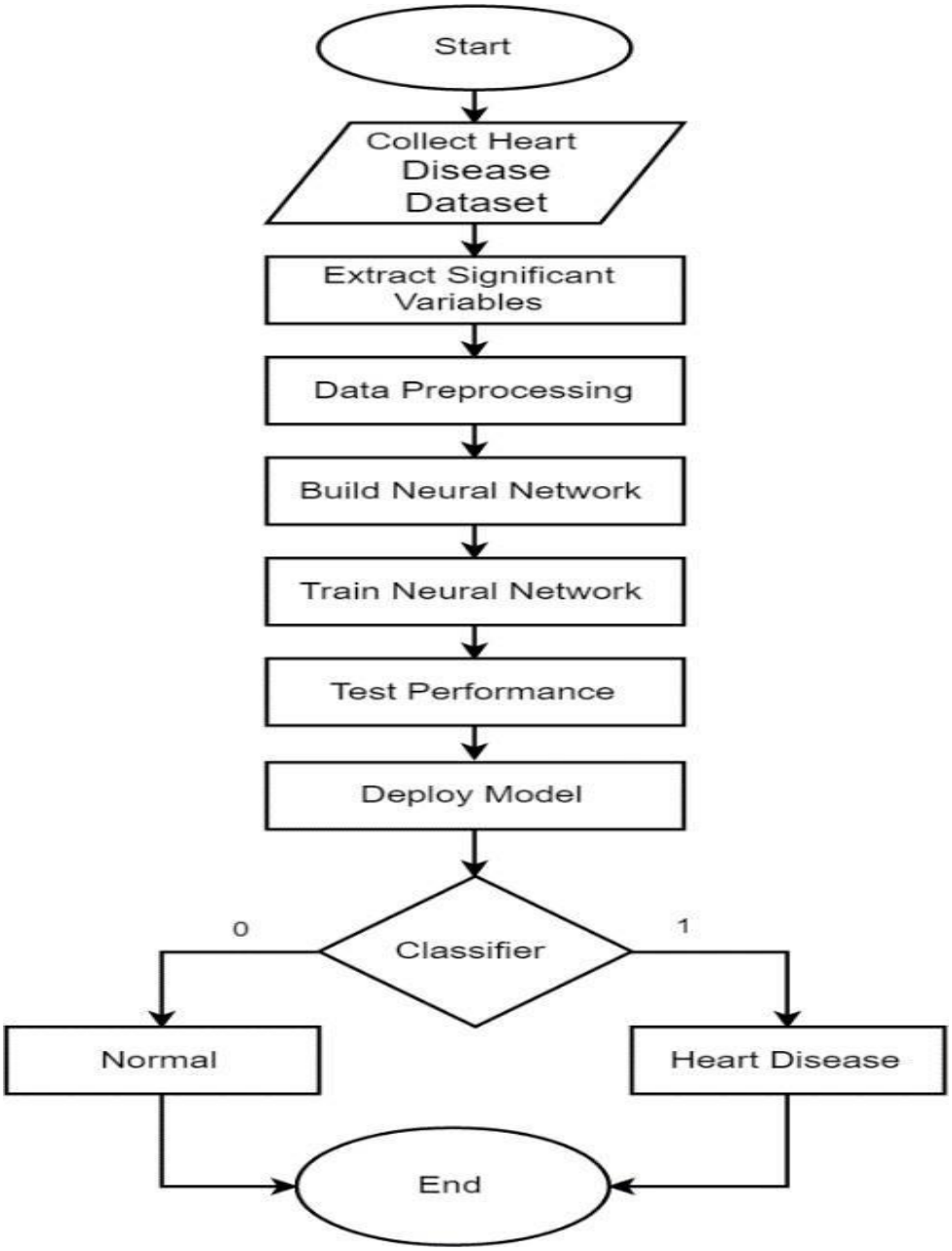
Since the application is deployed in cloud. The cloud vendor provides the availability of the application.

Scalability:

If the number of customers are increased, the application can be scaled through the cloud vendor.

5. PROJECT DESIGN

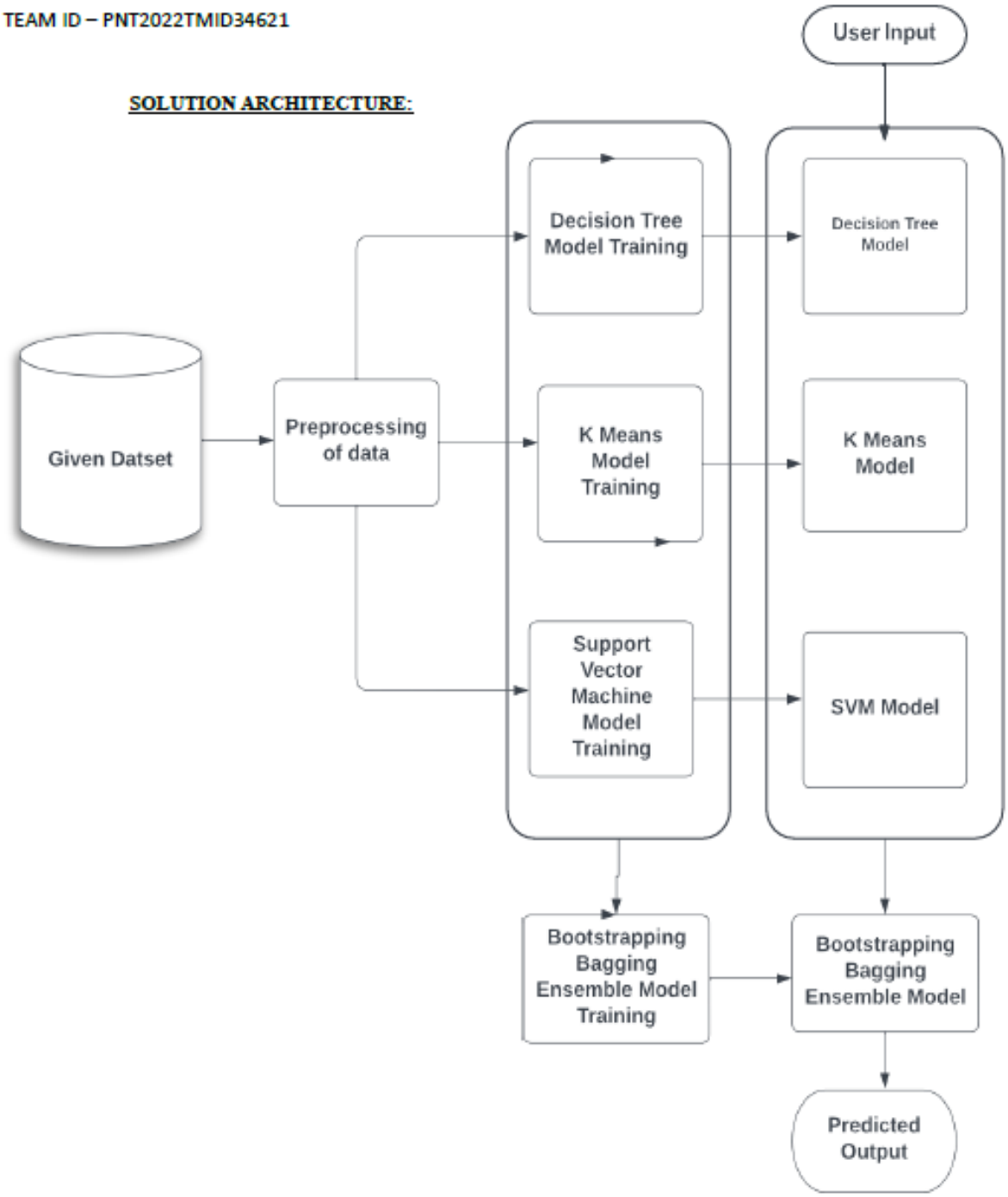
5.1 Data Flow Diagrams



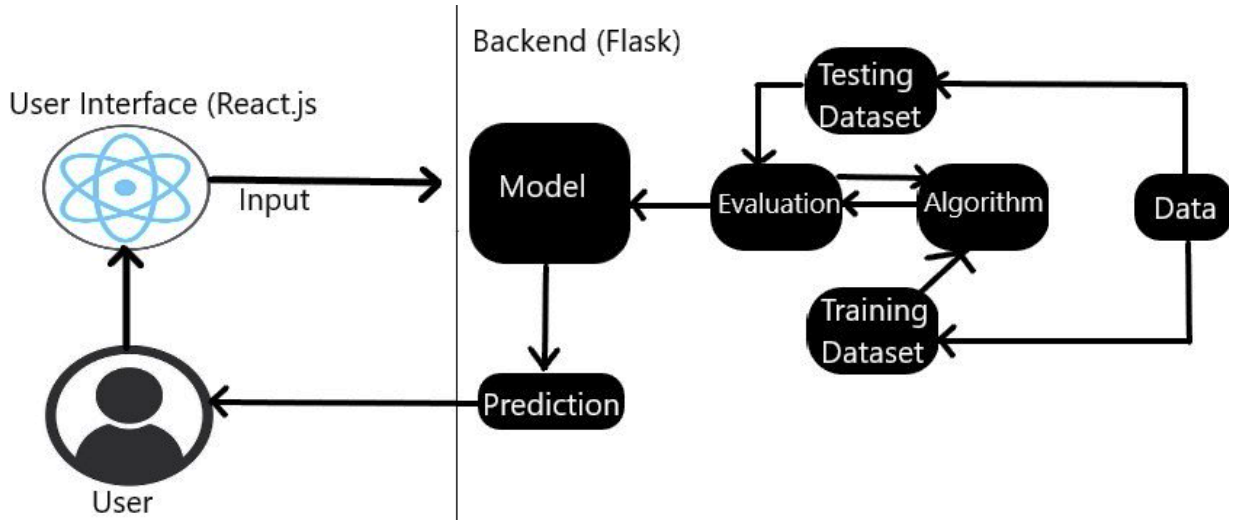
5.2 Solution & Technical Architecture

TEAM ID – PNT2022TMID34621

SOLUTION ARCHITECTURE:



Technical Architecture



6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

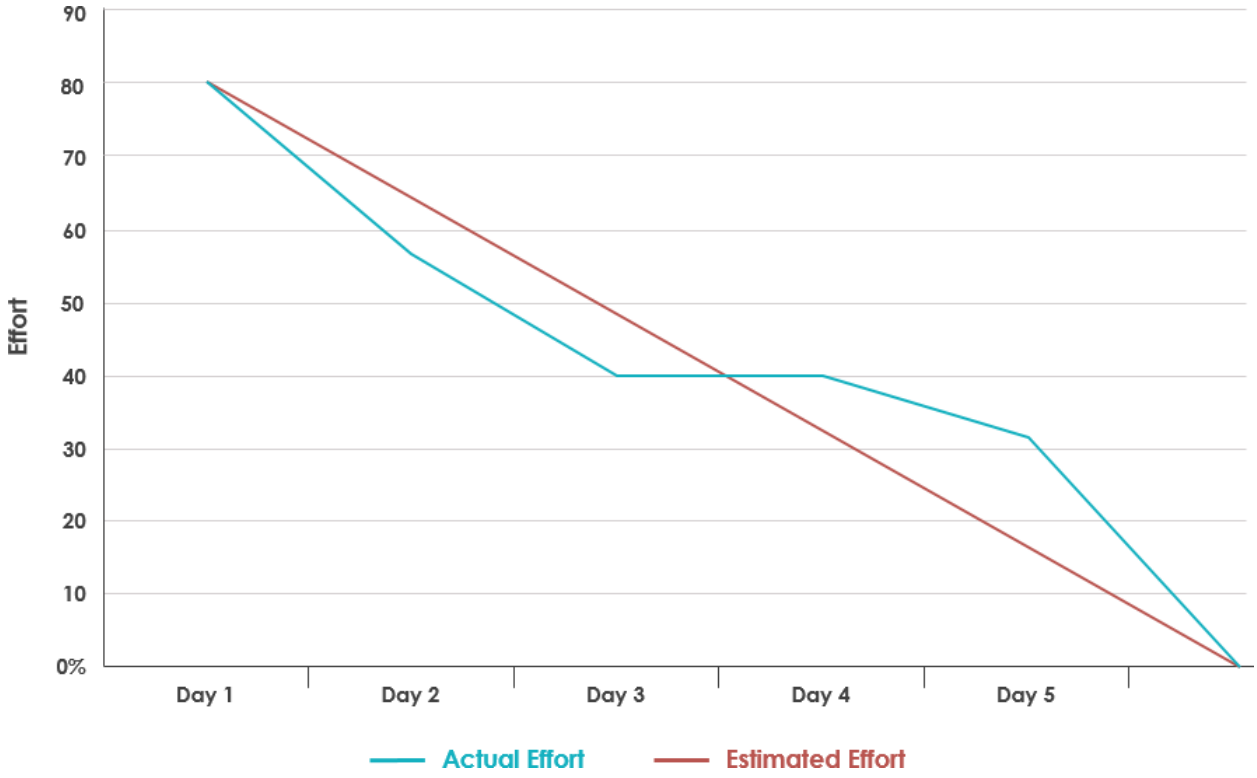
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Story points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the Dashboard by entering Email, password, and confirm password.	I can access my account / dashboard.	10	High	Bibin, Gino, Ajay.A.J, Ajay.E, Bibishek.G.S.Steephen
		USN-2	As a user, I can register for the application through Gmail.	I can register & access the dashboard with gmail login.	10	High	Bibin, Gino, Ajay.A.J, Ajay.E, Bibishek.G.S.Steephen
Sprint-2	Login	USN-3	After Registration Login page will appear, the user can login using the login credentials.	I can register & access the dashboard with gmail login.	20	High	Bibin, Gino, Ajay.A.J, Ajay.E, Bibishek.G.S.Steephen
Sprint-3	Cognos Dashboard	USN-4	The user is allowed to view or	I can view the profile.	10	Medium	Bibin, Gino, Ajay.A.J,

			update his/her profile.				Ajay.E, Bibishek.G.S.Steephen
Sprint-4	Classified result	USN-6	Home - Analyze your Heart.	I can predict the heart condition.	5	High	Bibin, Gino, Ajay.A.J, Ajay.E, Bibishek.G.S.Steephen
		USN-7	The user will have to fill in the 13 required fields for the system to predict a heart disease.	As a user, I can enter the datas in the specified fields.	10	High	Bibin, Gino, Ajay.A.J, Ajay.E, Bibishek.G.S.Steephen
		USN-8	The report is generated based on the condition.	The user is able to view	5	Medium	Bibin, Gino, Ajay.A.J, Ajay.E, Bibishek.G.S.Steephen

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	28 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	02 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	09 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	16 Nov 2022

Burndown Chart

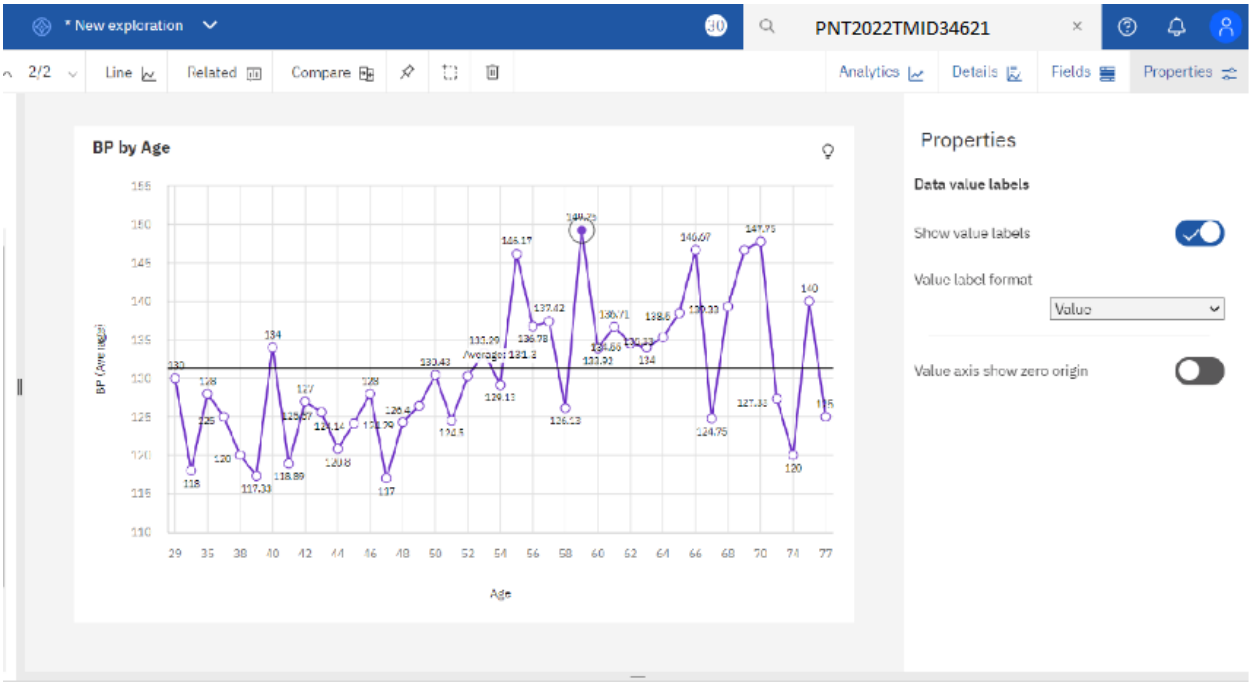


7. CODING & SOLUTIONING

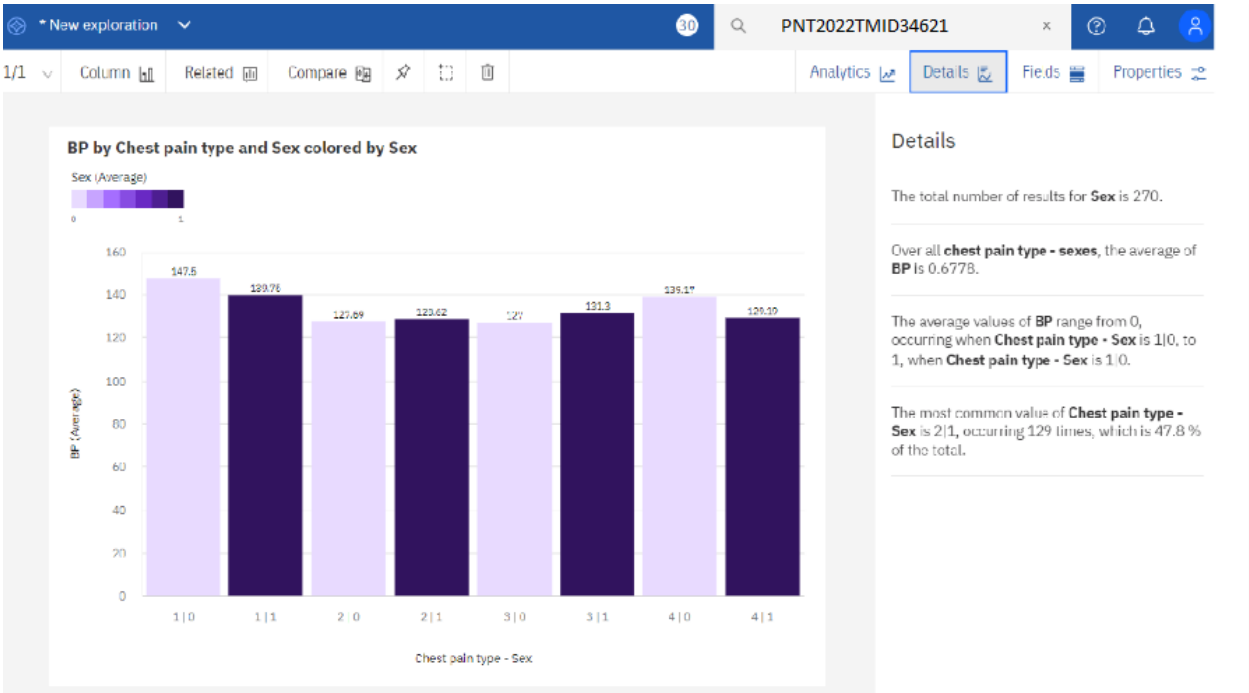
7.1 Machine Learning

7.2 Dashboard

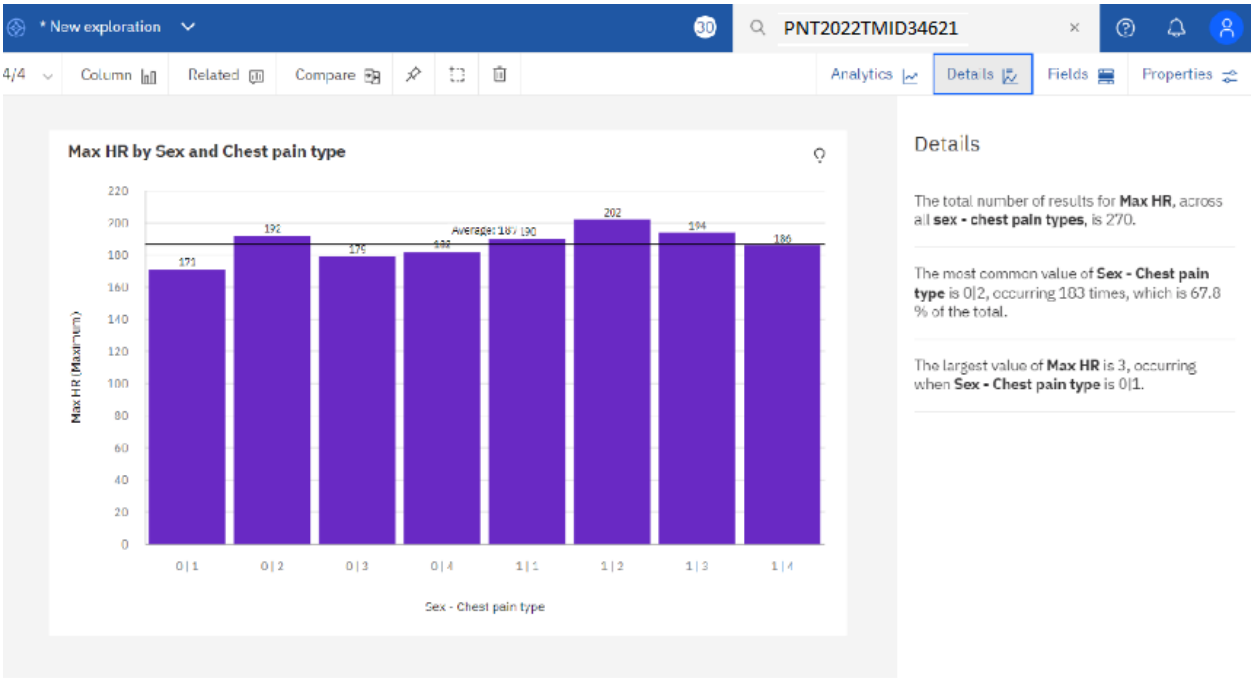
Average BP during chest pain



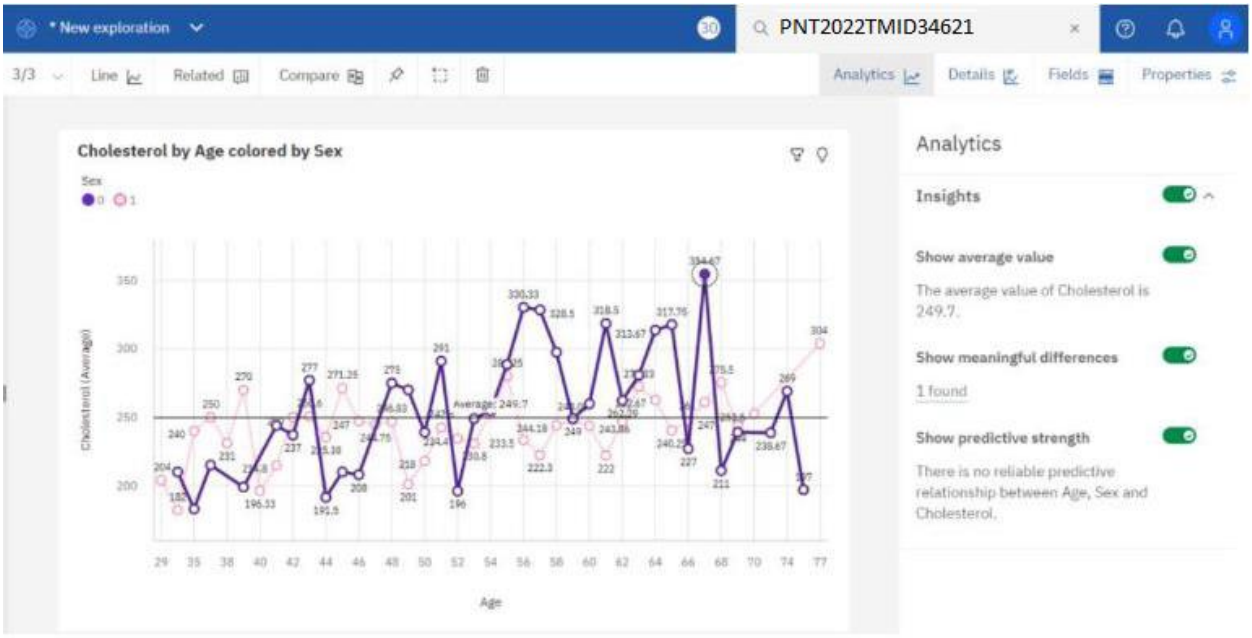
Exploration of BP vs Chest Pain Type and Gender:



Exploration of Max Heart Rate During the Chest Pain:

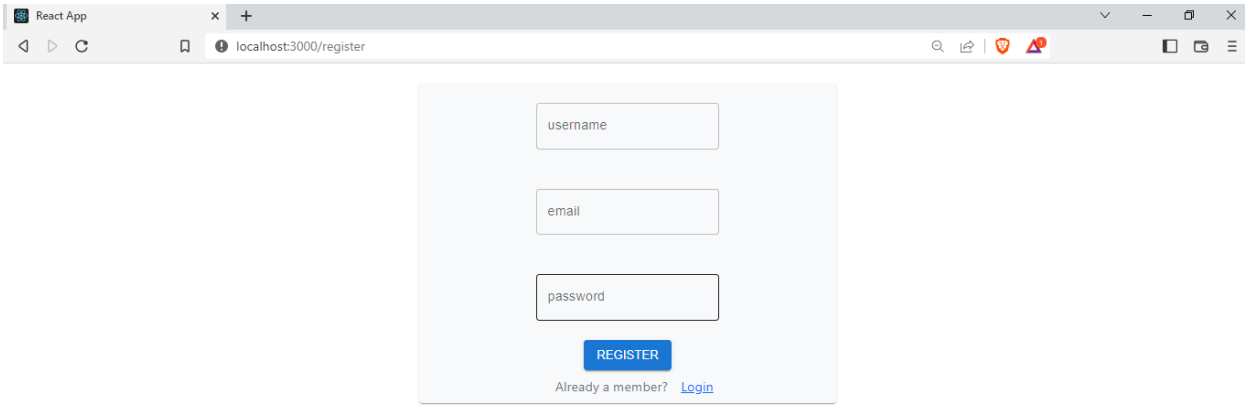


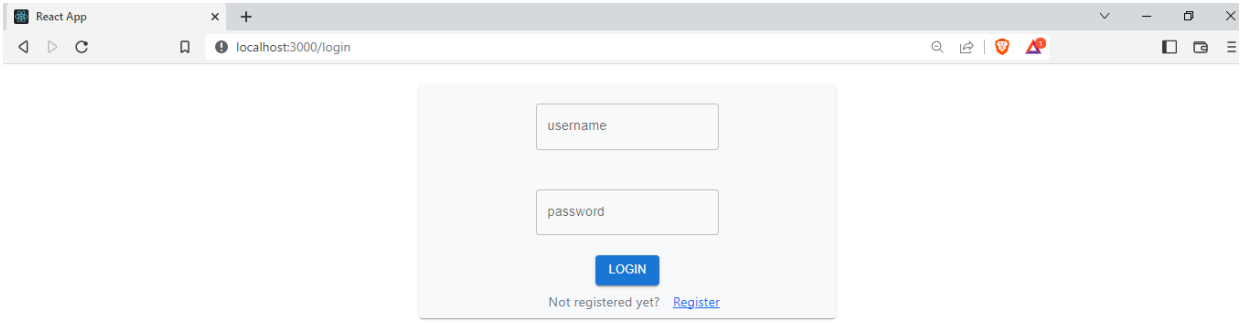
Exploration Of Cholesterol by age and Gender:

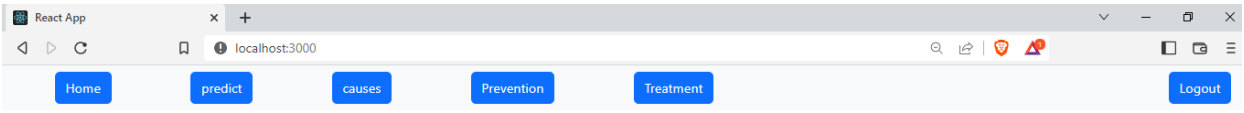


Dashboard Showing Different Types of Visuals:

8. TESTING

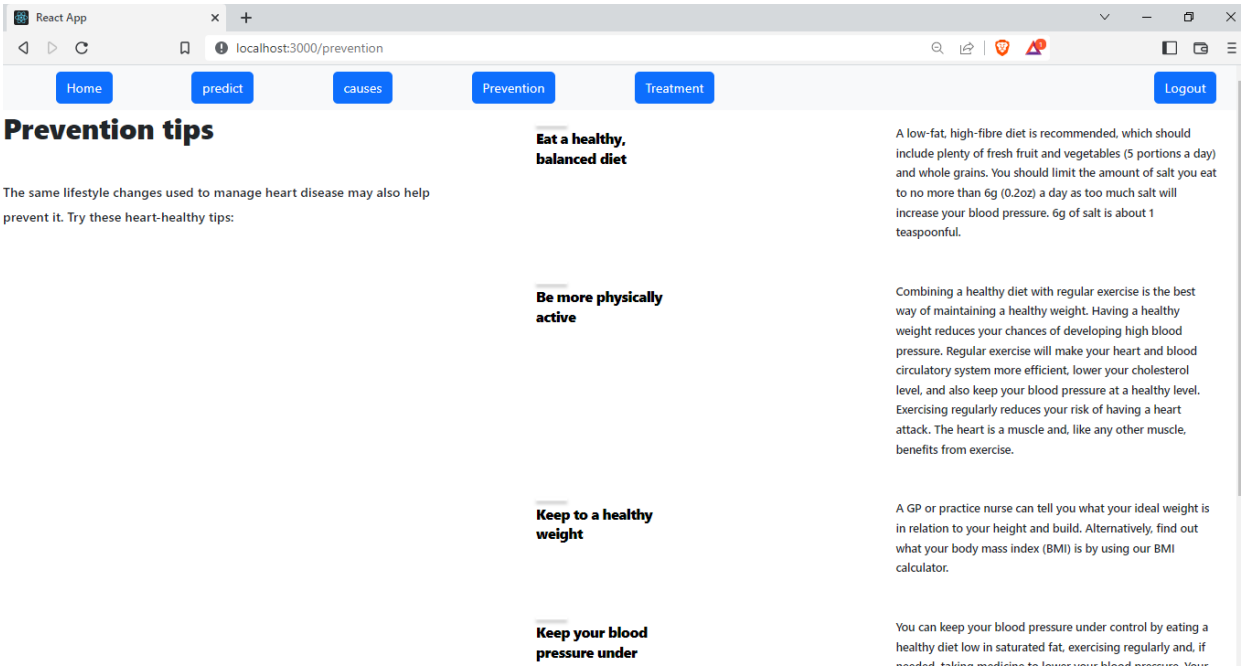

React App
localhost:3000/register
username
email
password
REGISTER
Already a member? Login

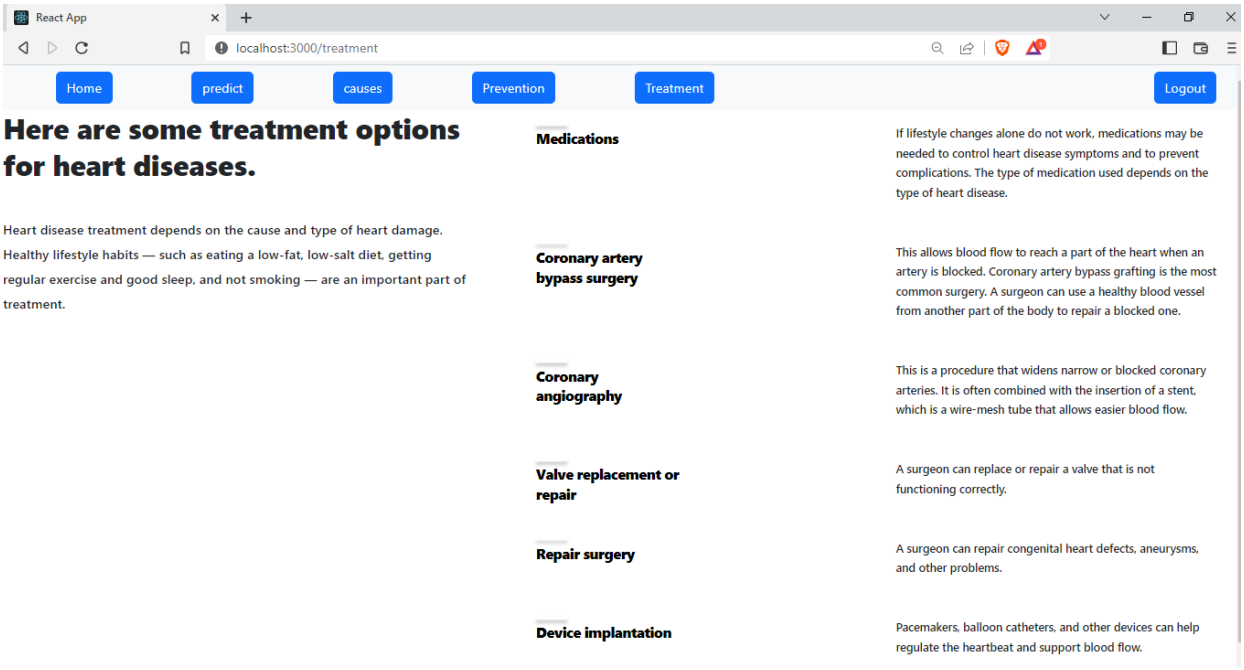

React App
localhost:3000/login
username
password
LOGIN
Not registered yet? Register

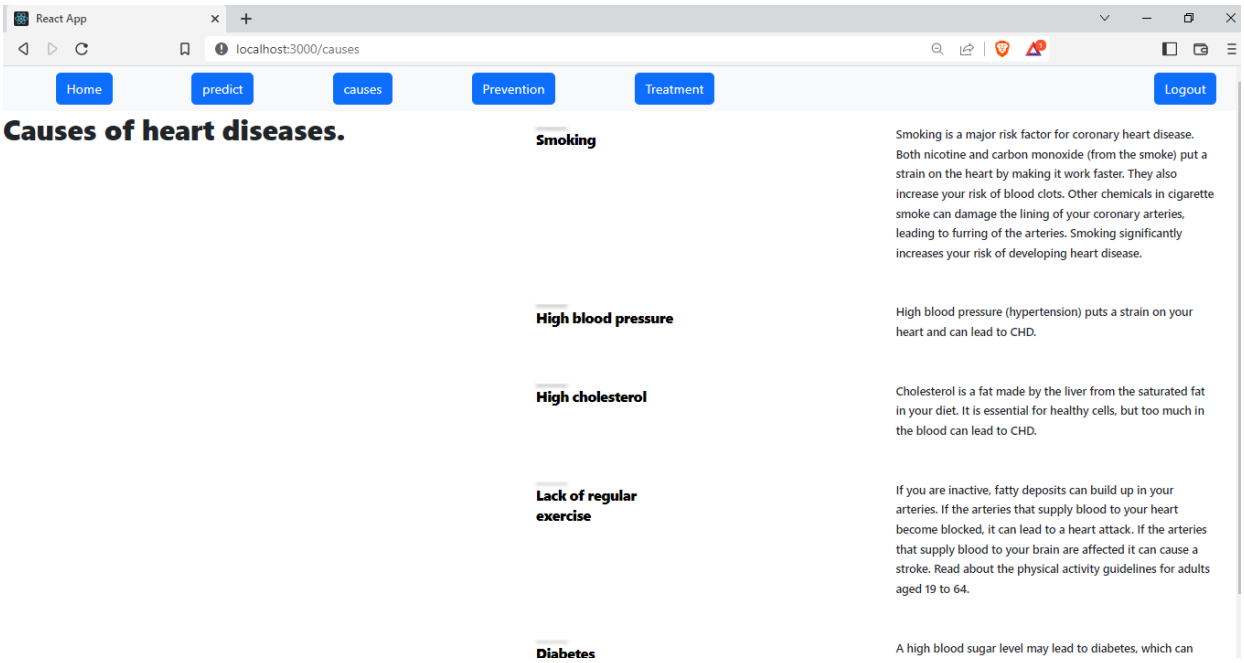


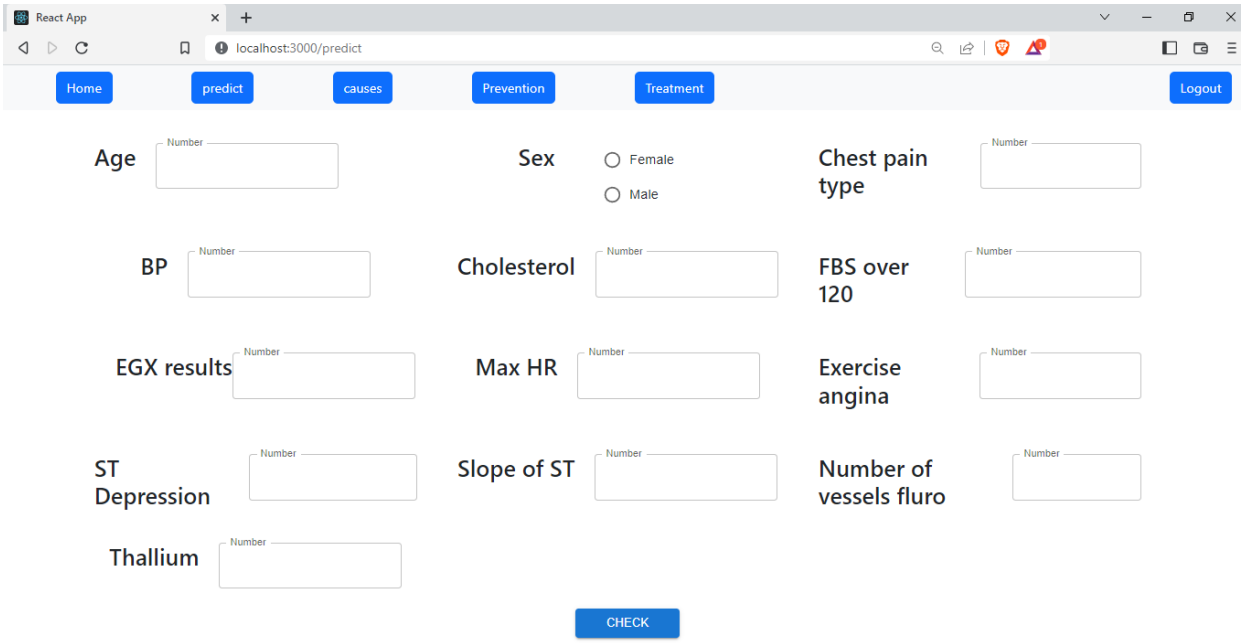
Heart Disease Prediction Tool

Heart Disease Prediction Tool provides you the option of diagnosing whether you have heart disease or not without visiting a hospital.



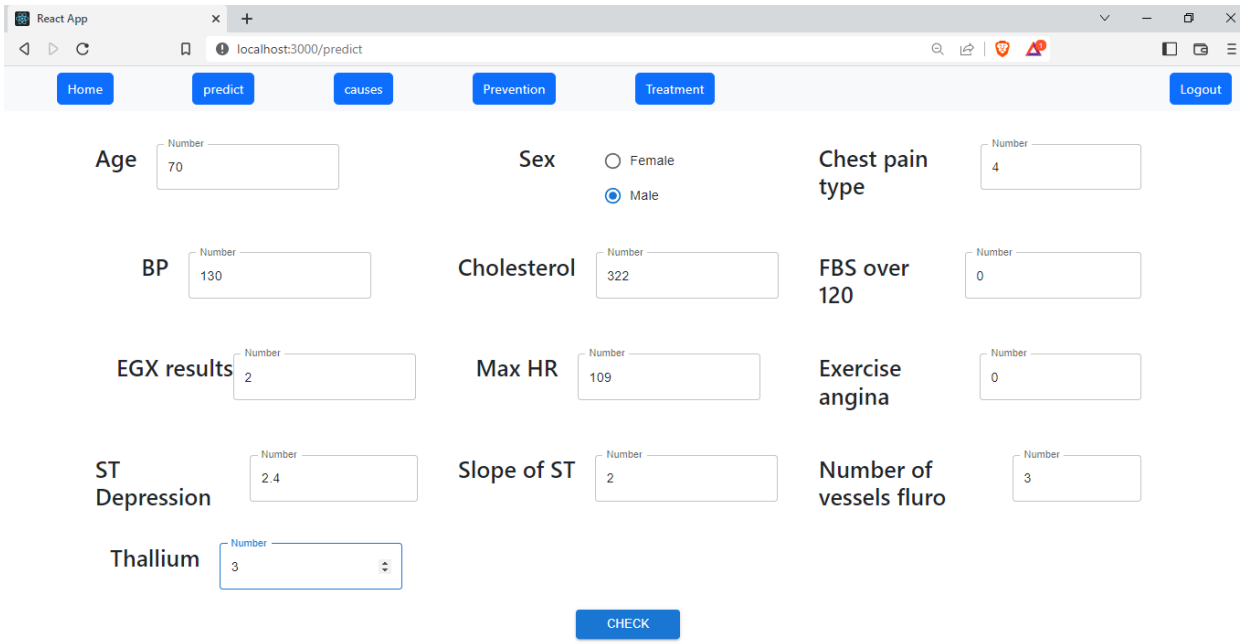

React App
localhost:3000/treatment
Home
predict
causes
Prevention
Treatment
Logout
Here are some treatment options for heart diseases.
Heart disease treatment depends on the cause and type of heart damage. Healthy lifestyle habits — such as eating a low-fat, low-salt diet, getting regular exercise and good sleep, and not smoking — are an important part of treatment.
Medications
If lifestyle changes alone do not work, medications may be needed to control heart disease symptoms and to prevent complications. The type of medication used depends on the type of heart disease.
Coronary artery bypass surgery
This allows blood flow to reach a part of the heart when an artery is blocked. Coronary artery bypass grafting is the most common surgery. A surgeon can use a healthy blood vessel from another part of the body to repair a blocked one.
Coronary angiography
This is a procedure that widens narrow or blocked coronary arteries. It is often combined with the insertion of a stent, which is a wire-mesh tube that allows easier blood flow.
Valve replacement or repair
A surgeon can replace or repair a valve that is not functioning correctly.
Repair surgery
A surgeon can repair congenital heart defects, aneurysms, and other problems.
Device implantation
Pacemakers, balloon catheters, and other devices can help regulate the heartbeat and support blood flow.

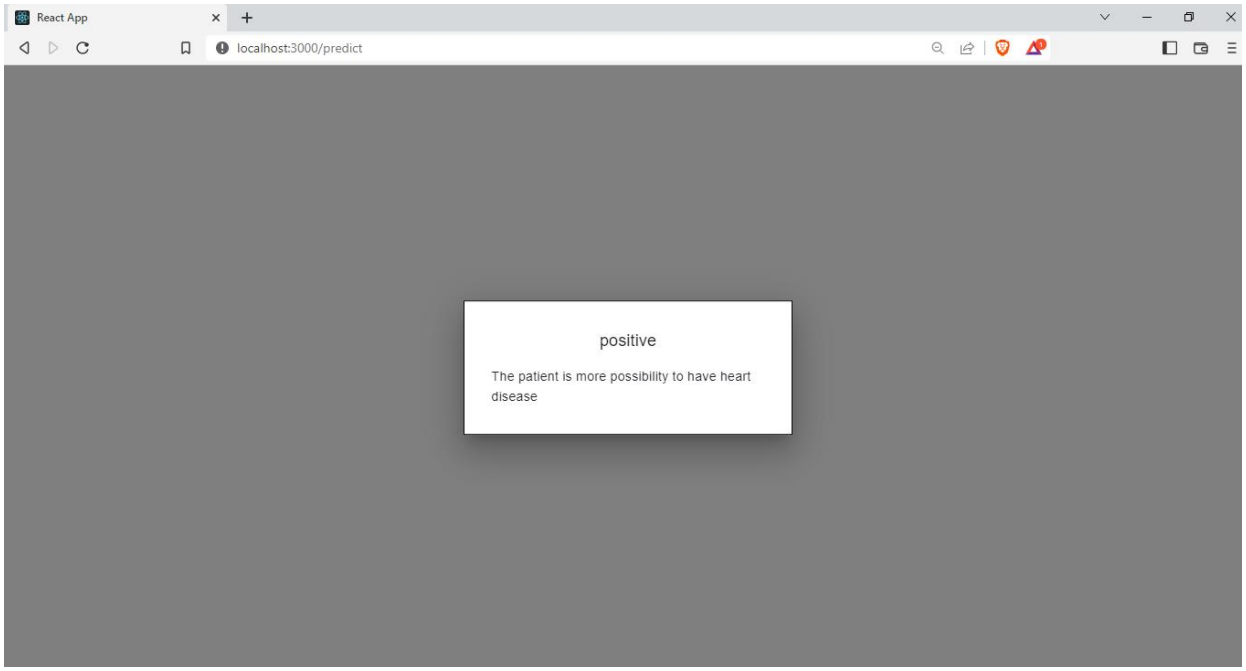

React App
localhost:3000/causes
Home
predict
causes
Prevention
Treatment
Logout
Causes of heart diseases.
Smoking
Smoking is a major risk factor for coronary heart disease. Both nicotine and carbon monoxide (from the smoke) put a strain on the heart by making it work faster. They also increase your risk of blood clots. Other chemicals in cigarette smoke can damage the lining of your coronary arteries, leading to furring of the arteries. Smoking significantly increases your risk of developing heart disease.
High blood pressure
High blood pressure (hypertension) puts a strain on your heart and can lead to CHD.
High cholesterol
Cholesterol is a fat made by the liver from the saturated fat in your diet. It is essential for healthy cells, but too much in the blood can lead to CHD.
Lack of regular exercise
If you are inactive, fatty deposits can build up in your arteries. If the arteries that supply blood to your heart become blocked, it can lead to a heart attack. If the arteries that supply blood to your brain are affected it can cause a stroke. Read about the physical activity guidelines for adults aged 19 to 64.
Diabetes
A high blood sugar level may lead to diabetes, which can



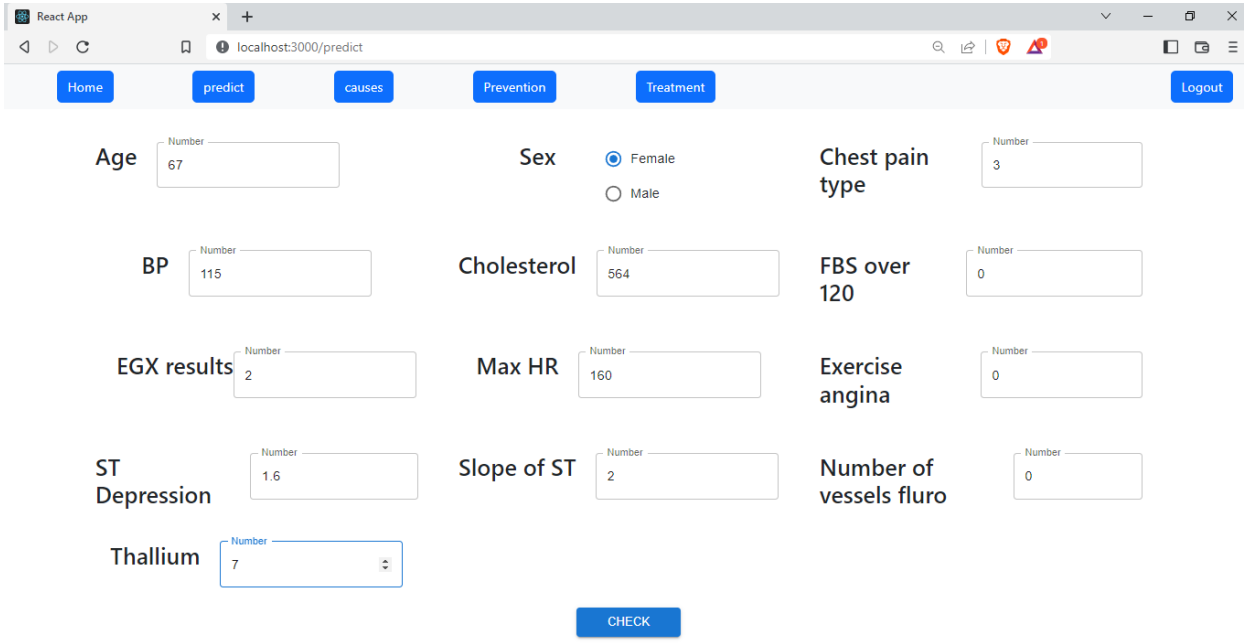
8.1 User Acceptance Testing

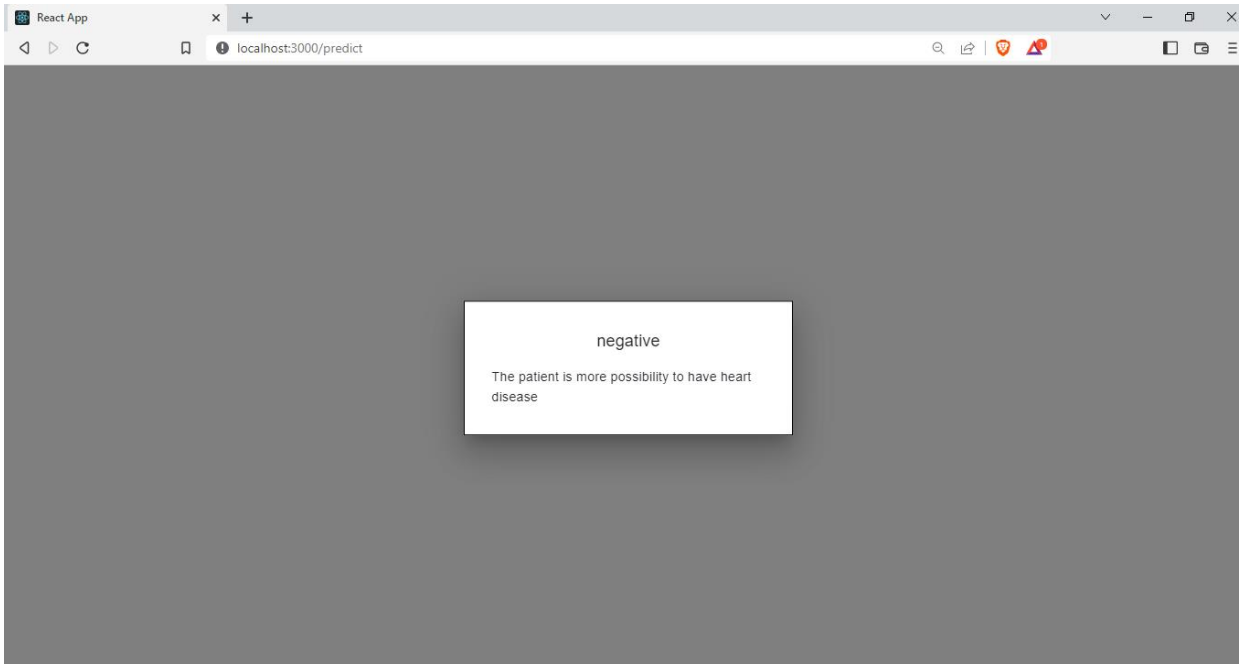
Testing a case where user has heart disease





Testing a case where user does not have heart disease




React App
localhost:3000/predict
negative
The patient is more possibility to have heart disease

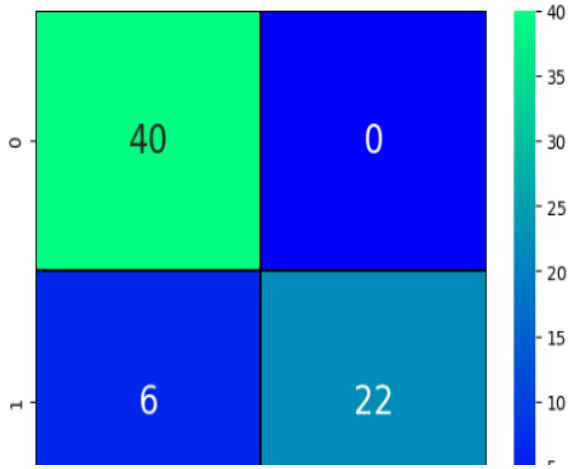
9. RESULTS

9.1 Performance Metrics

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

```
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

10.1 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
- Secure
- Dashboard provides insightful information.

10.2 Disadvantages

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

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.

13. APPENDIX

Source Code:

https://github.com/IBM-EPBL/IBM-Project-49273-1660817397/tree/main/Final%20Deliverables

GitHub & Project Demo Link:

https://drive.google.com/file/d/1HWYqEY6cHOqc46PWtDTOKezXJy01Nf4b/view?usp=sharing