

NEWS TRACKER APPLICATION

USING CLOUD

A Project report submitted in partial fulfillment of 7th semester in degree Of

BACHELOR OF ENGINEERING

IN

ELECTRONICS AND COMMUNICATION ENGINEERING

Submitted By

Team Members:

Team ID: PNT2022TMID33466

**DEPARTMENT OF ELECTRONIC AND COMMUNICATION
ENGINEERING**

V.S.B ENGINEERING COLLEGE, KARUR

NOV-2022

V.S.BENGINEERINGCOLLEGE,KARUR

(ApprovedbyAICTE&AffiliatedbyAnnaUniversity,Chennai)

BONAFIDECERTIFICATE

Certified that this project report “Visualizing and Predicting

Heart Diseases with an

InteractiveDashboard”istheBonafiderecordwork done by Ms

ABIRAMI T(9225191061003), Ms AFYA FARGATH

A(922519106005),Ms JENIFER G(922519106055),andMs

JOTHIKA M.G(922519106057)forIBM-

NALAIYATHIRAN in VII semester of B.E., degree course

inElectronicsandCommunicationEngineeringbranchduringthe

**Head of the
Department Mrs. Dr
P.S Gomathi**

ACKNOWLEDGEMENT

First and foremost, we express my thanks to our parents for providing us a very nice environment for doing this mini project. We wish to express our sincere thanks to our founder and Chairman **Shri.V.S.BALSAMY** for his endeavour in educating us in this premier institution.

We wish to express our appreciation and gratefulness to our principal, **Dr.V.NIRMAL KANNAN** and vice principal **Mr.T.S.KIRUBASANKAR** for their encouragement and sincere guidance.

We are grateful to our head of the department **Mrs. Dr P.S GOMATHI** and our IBM Nalaiyathiran project coordinator **Mrs. Dr P.S GOMATHI** Department of Electronics and Communication Engineering for their valuable support.

We express our indebtedness to the supervisor of our IBM Nalaiyathiran project, **Mrs. Dr P.S GOMATHI** Professor, Department of Electronics and Communication Engineering, for guidance throughout the course of our project. Our sin

ce thanks

Content

1. INTRODUCTION

Project Overview

Purpose

2. LITERATURE SURVEY

Existing problem

References

Problem Statement Definition

3. IDEATION & PROPOSED SOLUTION

Empathy Map Canvas

Ideation & Brainstorming

Proposed Solution

Problem Solution fit

4. REQUIREMENT ANALYSIS

Functional requirement

Non-Functional requirements

5. PROJECT DESIGN

Data Flow Diagrams

Solution & Technical Architecture

User Stories

6. PROJECT PLANNING & SCHEDULING / development phase

Sprint Planning & Estimation

Sprint Delivery Schedule

Reports from JIRA

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

Feature 1

Feature 2

8. TESTING

Test Cases

User Acceptance Testing

9. RESULTS

Performance Metrics

10. ADVANTAGES & DISADVANTAGES

11. CONCLUSION

12. FUTURE SCOPE

13. APPENDIX

Source Code

GitHub & Project Demo Link

TEAMID:PNT2022TMID033843

VISUALIZING AND PREDICTING HEART DISEASES WITH AN INTERACTIVE DASHBOARD

1. Introduction

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.

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

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 modeling. 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

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.

Who does the problem affect?

People with unhealthy lifestyles, stress, depression, age above 40 and when their ancestors go the 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?

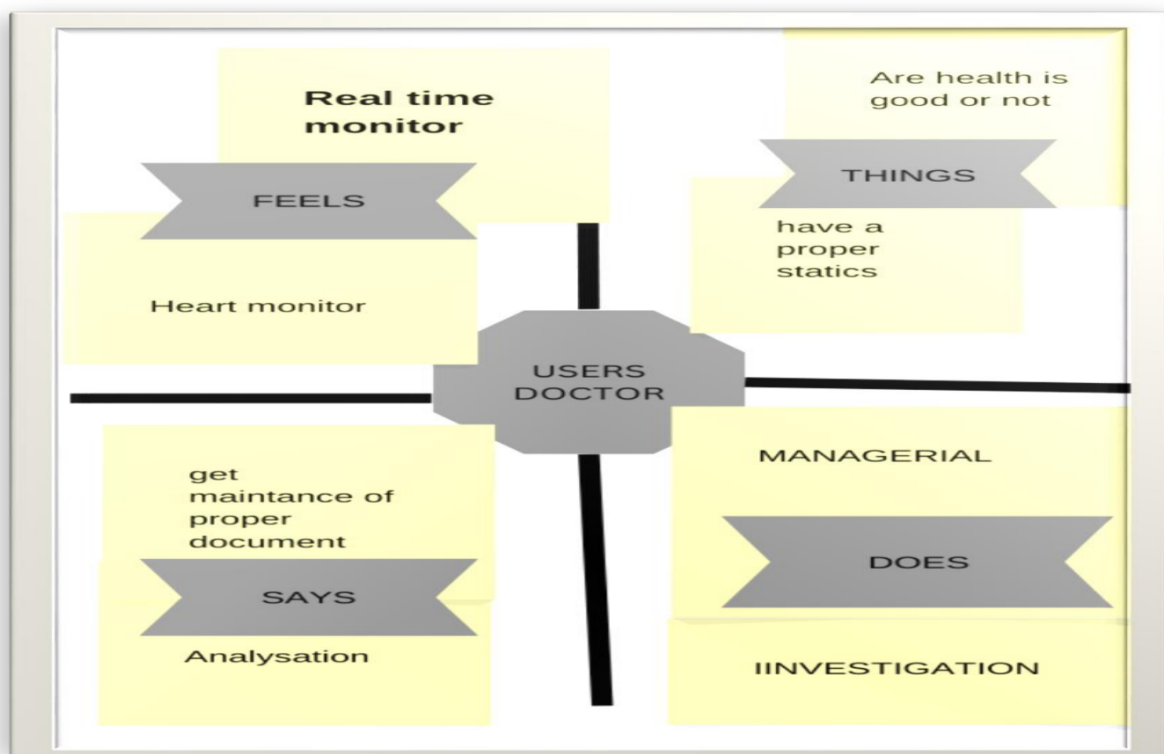
If we don't solve the problem, many people will die at a young age. The death rate due to heart disease will increase rapidly.

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 and Proposed Solution

Empathy Map Canvas



Ideation and Brainstorming

AfiyafargathA's Idea

Predicting heart disease with a visualizing like creating a common pie chart and dashboards with the help of collecting data in which age all should be aware of the heart disease. creating this and telling people to be precaution in that time and lead a healthy life.

Jothika MG's Idea

We can have regular checkup of every individuals monthly wise and get collected in village to city orderwise, this might help to predict the heart disease of individual source.

Jenifer's Idea

We can create our own chart on the basis of our health and get used to it and make sure of predicting the sugar level and other criterias by collecting their data which also use ful for us predict the future heart diseases easily and we can be prepared and get a precaution for this

BEST IDEA

Team member Afiya fargath idea is best because the modulation of our body changes and creating data with age wise, what are all the mandatory things that one should do and do not everything should be tabled and visualize to all human beings. The pie chart will show the major effects of heart disease, and also show the advantage and level of danger with age and cholesterol level.

By this idea we can easily predict the heart disease and we can make sure that the person stays healthy or not. Heart disease is the major part of one's body to stay healthy. This chart will demonstrate all the levels in the body which should maintain in range. The data will display all the parameters in the body which cause a person heart attack.

The mapping also ensure that the project will be done in the accordance with the data creates visualizing and easily predictable and helps to see the causes of the other body parameters so that one can make themselves save by viewing this survey.

Heart Attack

Symptoms



chest pain



**shortness
of breath**



**pain, stiffness,
or numbness in
the upper body**



cold sweat



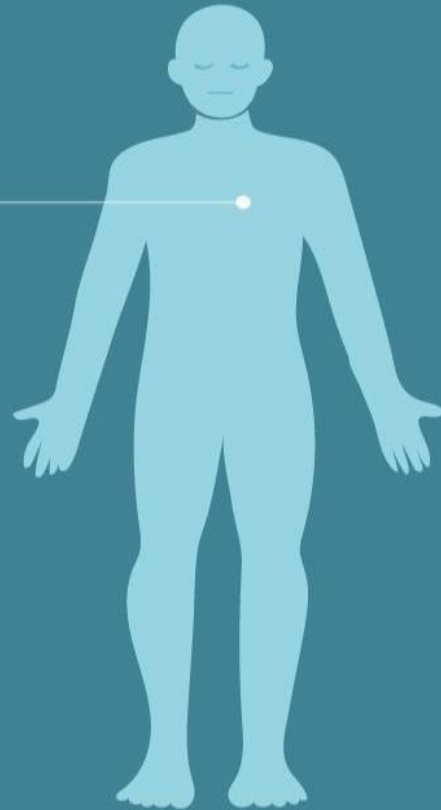
**nausea and/
or vomiting**



**light-headedness,
dizziness, or
fainting**



fatigue



**heartbeat
changes**



**stomach
discomfort**



**a blue tint to
the lips, hands,
or feet**

Proposed Solution

S.No.	Parameter	Description
1.	ProblemStatement(Problem to be solved)	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 heart improved. The predicted results can be used to prevent and thus reduce cost for surgical treatment and other expenses. Data Analytics will be very useful for this to predict the heart by visualization.
2	Idea/ Solution description	We have an idea to predict heart disease by means of collecting the data according to the human's age. The data collected must include the parameters such as blood pressure, sugar, cholesterol and habitual. Giving the validate correct medicines and involving diet for those individuals make them get rid of heart disease. It can be predicted easily.

3.	Novelty/Uniqueness	Reduced heart rate variability has also been observed in depressed patients when... with heart disease has been shown in numerous studies to be related to decreased heart rate.
4.	Social Impact/Customer Satisfaction	Involve the heart or blood vessels. Cardiovascular disease comprises coronary artery diseases (CAD) such as angina and myocardial infarction (heart attack), stroke, hypertensive heart disease.

5.	Business Model (Revenue Model)	... Business aspirations dovetail in that the cardiac problems ... She is convinced that a predictive mathematical model can be ... Furthermore, the revenue model based on fixed installation...
----	--------------------------------	---

6.	ScalabilityoftheSolution	<p>A scalable framework that useshealthcare data to predict heartdisease based on certainattributes. Our main contributionin this work is to predict thediagnosisofheartdiseasewithas mall number of attributes. Ourprediction solution uses randomforest on Apache Spark, whichgives massive opportunity forhealth care analysts to deploythis solution on ever changing,scalable big data landscape forinsightfuldecisionmaking.Usingt his approach, we show that upto98%accuracyisachieved. We also present a comparisonagainst Naïve-Bayes classifier,where we show the randomforestapproachoutperforms theformerbyasignificantmargin.</p>
----	--------------------------	--

ProblemSolution Fit

<u>1.Customersegments:-</u>	<u>6.Customerconstrains:-</u>	<u>5.Availablesolutions</u>
Manyurbanareapeoplesareunabletopredict the heart disease. they are thecustomers.	The constrains that if they not predict the heartdisease that may lead to immediate health issuesand loss of people. By this predictive idea we cansaveandletpeoplecangainknowledgeaboutheirbodyparametersandlevels.	collectingthedataaccordingtothe human age. The data collected must includethe parameters such as bloodpressure , sugar ,cholesterolandhabitual.Givingthe validate correct medicines and involvingdietforthoseindividuals makethemgetridofheartdisease.It canbepredictedeasily. Reducedheartratevariabilityhasalsobeennoobserved in depressed patients when ... withheart disease has been shown in numerousstudiesto berelatedtodecreasedheartrate.

<p><u>2. Job to be done:-</u></p> <p>To collect data from the people give the accurate information of prevention of their health from the heart disease. visualizing about the risk and prediction ways.</p>	<p><u>9. Problem route cause:-</u></p> <p>Important to fill the advantages and disadvantages of leaving this as simple. The disease major cause. The predictive analysis.</p>	<p><u>7. Behavior:-</u></p> <p>Finding the right and accurate data is major importance, because the accuracy is important for let people know about their disease,</p>
---	--	---

<p>3. Triggers:- some of the triggers are creating awareness among the people</p> <p>4. Emotions:- With most of the people afraid of heart disease that instantly makes them low. Predictive system helps them to know their needs and health before.</p>	<p><u>10. Solution:-</u></p> <p>All the humans cannot predict the heart disease. We should develop an predictive data analytics system method and convey the modulation of body changes according to the parameters, such as sugar and BP level. and visualizing the risk factors and preventing methods in their period of health</p>	<p>8. Channel of behavior:- ON LINE: Through online we can advertise the modulation of change that one body carries and create awareness OFFLINE: Not everyone can benefit with online channel of behaviour. In urban or slum areas the doctor studying students can go for the camp and directly test the people in all those areas and collect all the data about their health and can give them a great awareness and all the helping predictive method needed by them.</p>
---	---	---

4.RequirementAnalysis

Following are the functional requirements of the proposed solution.

FRNo.	FunctionalRequirement(Epic)	SubRequirement(Story/Sub-Task)
FR-1	PersonRegistration	Registration through FormRegistration through GmailRegistrationthroughLin kedIN
FR-2	Confirmation	Confirmation via EmailConfirmationvia OTP
FR-3	Authentication	As this predicting method has variety of secure purpose the authentication is a major requirement like passkey.
FR-4	AccurateData	We need a require data's of like high blood pressure, high cholesterol, family history, being postmenopausal(women), obesity , stress and physical inactive
FR-5	TransitionRequirements	The transition takes place in short period of time as the data changes the accurate details of individuals is needed.

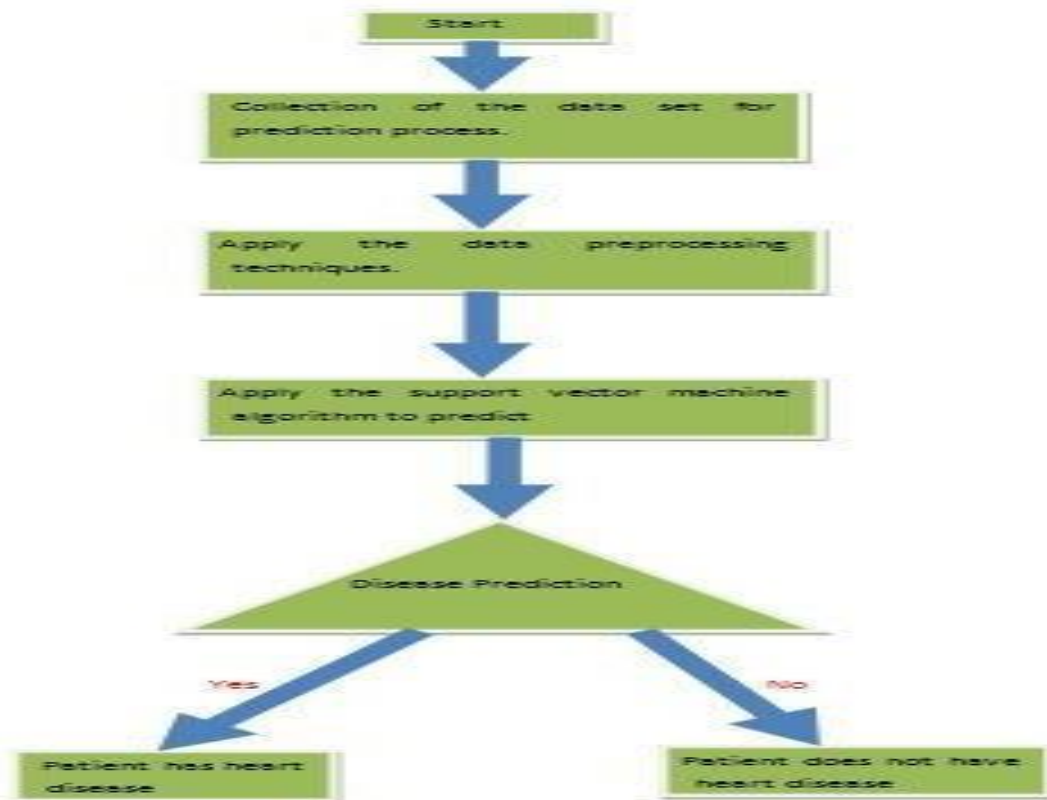
Non-functional Requirements:

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

FRNo.	Non-Functional Requirement	Description
NFR-1	Usability	Visualizing and predicting techniques can be wide range of uses and it can be used with a portable device. Self-analysis can also be done. The conveying method for self-analysis is smartwatch. Using smartwatch, the heart beat and more levels can be identified.
NFR-2	Security	This process is very secure because by predicting the heart rate regularly by using advanced technology and especially cloud technology. By predicting the heart rate if the heart rate detected low then it can be easily got treatment before any dangerous causes.
NFR-3	Reliability	The study of investigating about various heart disease causing parameters such as HRV measurement reliability in patients with chronic obstructive pulmonary disease. Collecting all data and giving the appropriate solutions can be done. This study employed a limited number of HRV parameters and statistical analysis to be used of many clinical decisions.
NFR-4	Performance	Performance is accurate and there will be abrupt changes as the data prediction has changes. As data analytics is done in any time and in any way we would make use of it for the better performance. The powerful data driven recurrent neural network algorithm for the analysis of accelerometer data to make future prediction.
NFR-5	Availability	This predicting method is available at all levels and the collection of data is simple. Availability can be very essential and easy manner.
NFR-6	Scalability	Our predicting analysis is most scalable as it could easily work on with wide range of availability, densities and every one could use this data collection and visualizing the effects of that and can make treatment then lead a healthy life.

1. ProjectDesign

DataFlowDiagram





Project Planning and Scheduling

Script Planning and Execution

User Type	Functional Requirement (Epic)	User Story Number	User Story/Task	Acceptance Criteria	Priority	Release
Customer (Mobile 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
		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user, I can register for the application through Gmail	I can register & access the dashboard with Gmail Login	Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password	I can register & access the dashboard with Gmail Login	High	Sprint-1
	DashBoard	USN-6	Profile-view & update your profile	I can see the profile.	Medium	Sprint-2
		USN-7	Change Password - user can change the password	I can able to change the password	High	Sprint-1
		USN-8	Home - Analyze your Heart	I can detect the health condition from where ever I want.	High	Sprint-1

		USN-9	The user will have to fill in the below 13 fields for the system to predict a disease -Age in Year - Gender -Chest Pain Type -Fasting Blood Sugar -Resting Electrographic Results(Restecg) -Exercise Induced Angina(Exang) -The slope of the peak exercise ST segment -CA – Number of major vessels colored by fluoroscopy -Thal-Trest Blood Pressure -Serum Cholesterol -Maximum heart rate achieved(Thalach)-ST depression induced by exercise(Oldpeak)	These are the categories available in that application.	High	Sprint-2
Customer Care Executive	Dashboard	USN-10	Query	You can post your queries in the text box available in that application.	High	Sprint-1
Administrator	DashBoard	USN-11	Verification	Verification through CAPTCHA Verification through I'm not a robot	High	Sprint-1

	Points		Date	Date(Planned)	Completed(as On Planned End Date)	Rel eas e Dat e (Ac tual)
Sprint-1	20	6 Days	6 November 2022	11 November 2022	20	12 Nove mber 2022
Sprint-2	20	6 Days	7 November 2022	12 November 2022	17	12 Nove mber 2022
Sprint-3	20	6 Days	8 November 2022	13 November 2022	18	13 Nove mber 2022
Sprint-4	20	6 Days	10 November 2022	15 November 2022	13	15 Nove mber 2022

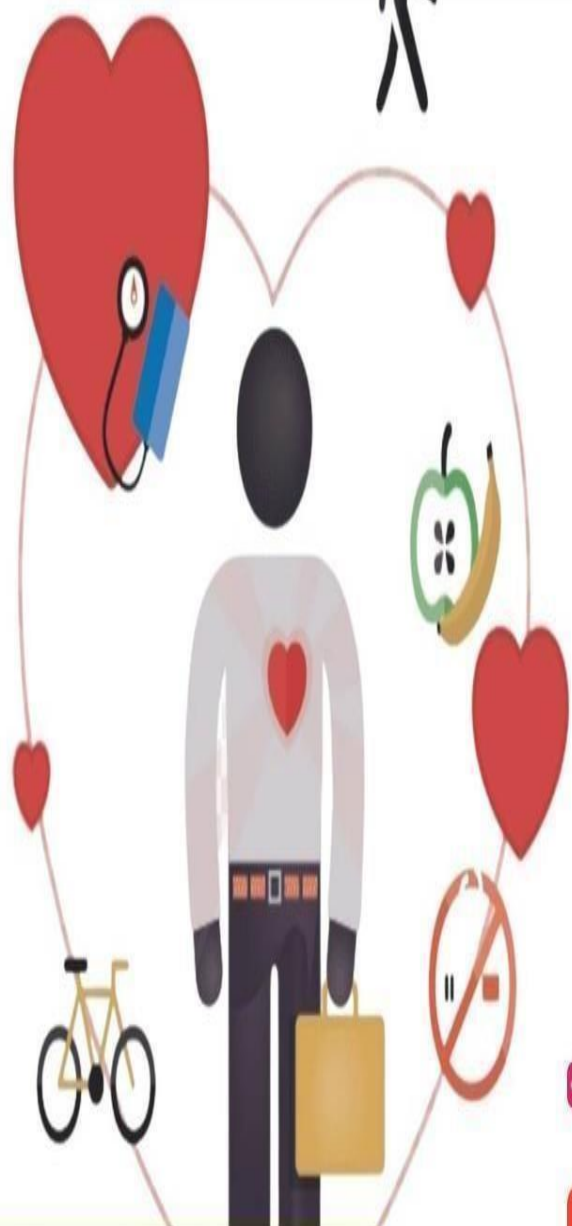
Burndown Chart



6.3JiraReport

Sprint-1

Registration-Can register for the application by entering my email, password, and confirming my password



Welcome to our Project for Heart Analysis Dashboard

[Dashboard](#) [Profile](#) [Sign Up](#) [Sign In](#)

Enter your email and password to register

Name

Email

Password

Phone Number

Blood Group

Address

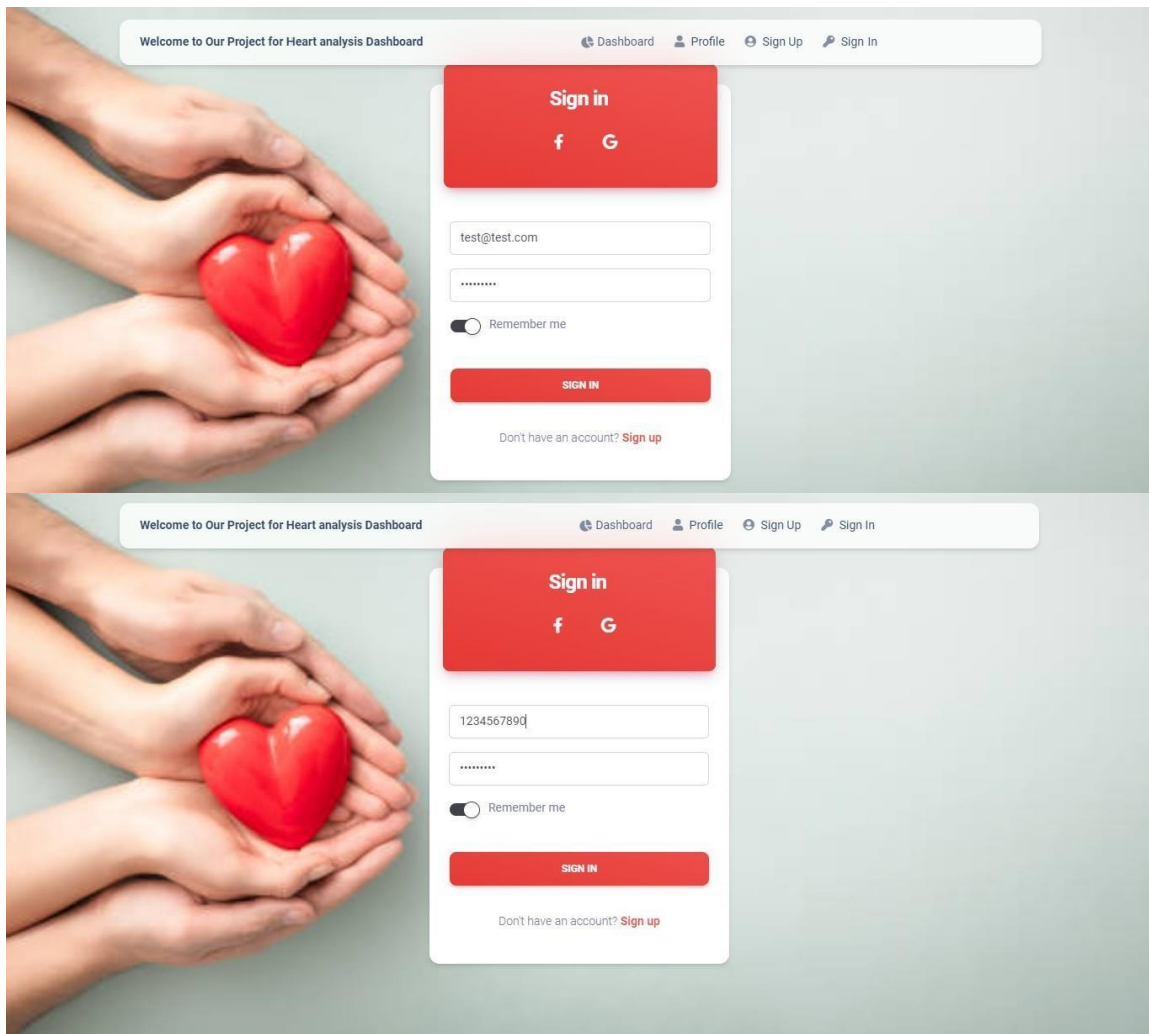
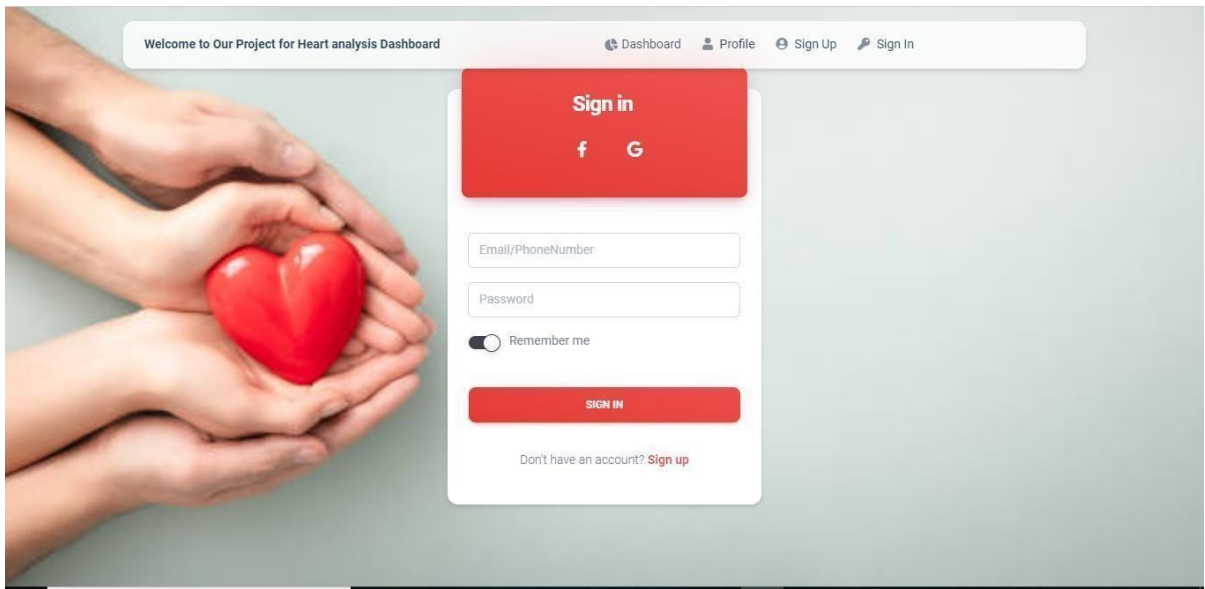
Age

☒ I agree the [Terms and Conditions](#)

SIGN UP

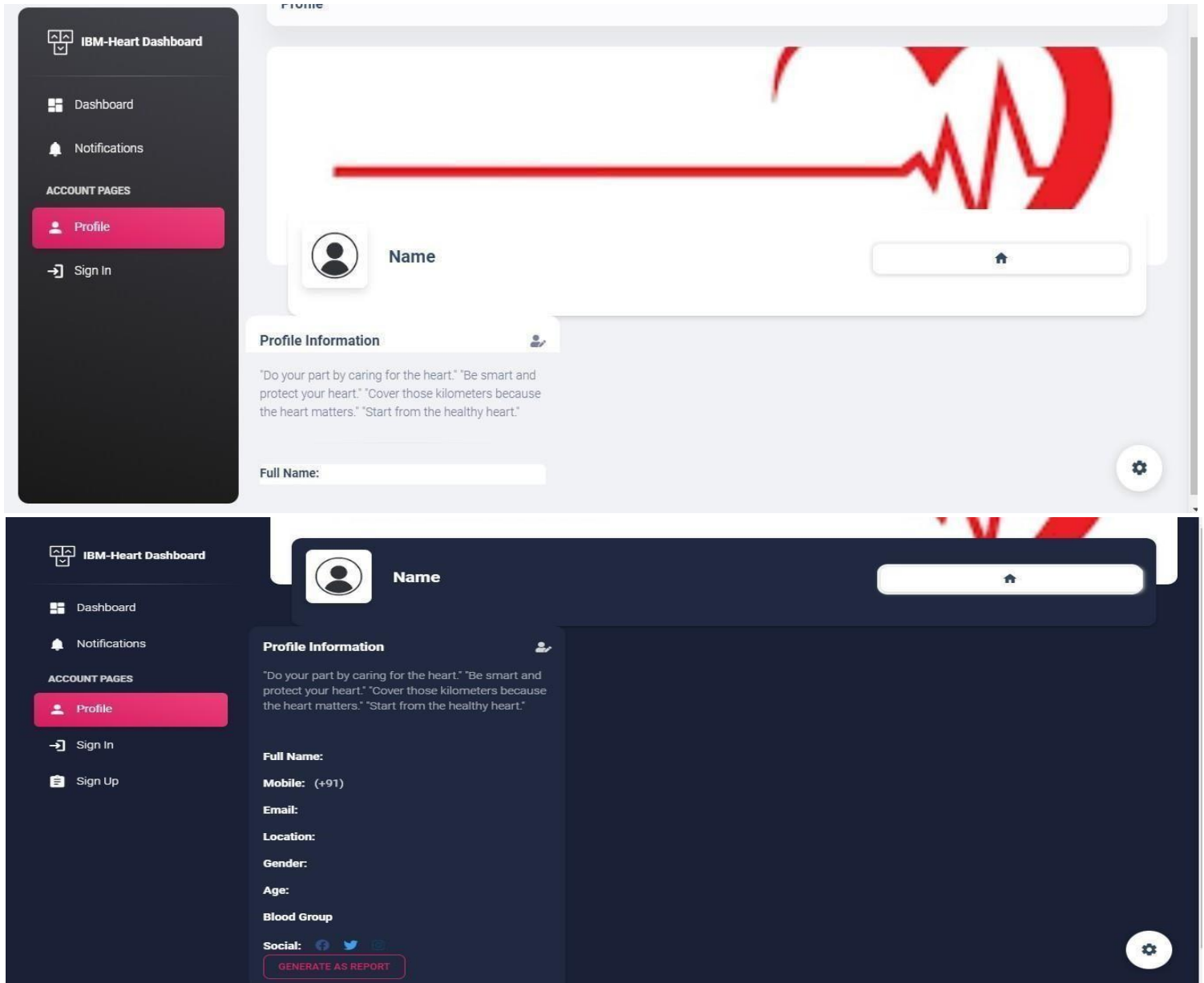
file:///C:/Users/user/Downloads/material-dashboard-master/pages/sign-up.html

Login - Can log into the application by entering email/PhoneNumber & password



Sprint-2

Profile- To Know the User about Him/Her Information and provide to Generate the Report for his Analysis



Dataset collection - The data required for analysis and prediction must be collected from various sources, Collecting Dataset from Different Site.

7 CodingAndSolutioning

MachineLearning

Learningwhichmodelisbest forthegivenDataset

Out[]:	Estimators	Accuracy
0	Linear Regression	0.565830
3	K-Nearest Neighbor	0.729167
4	Random Forest	0.854167
5	Bagging Decision Tree	0.854167
6	Hard coting classifier	0.854167
2	Gaussian Naive Bayes	0.875000
1	Logistic Regression	0.895833

From the above result we can conclude that Logistic Regression has the hisgest accuracy for this particular dataset.

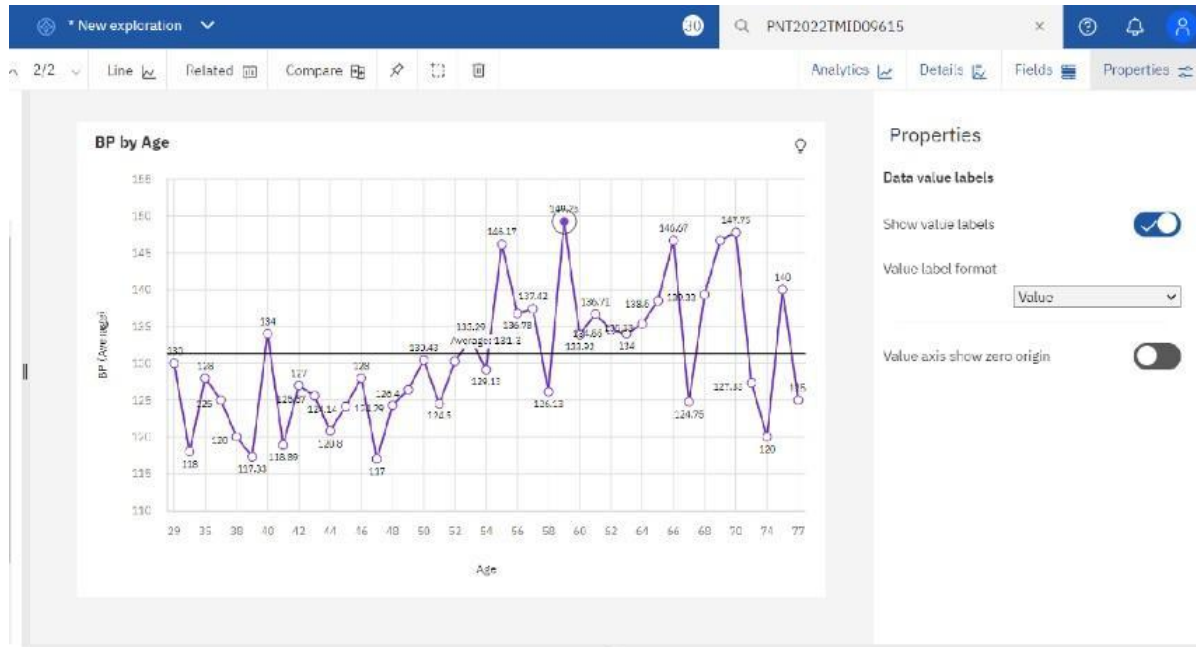
Comparingitwiththeaccuracygotten fromDecisionTree:

```
TP=cm[0][0]#cm=ConfusionMatrix
TN=cm[1][1]
FN=cm[1][0]
FP=cm[0][1]
print("TestingAccuracyforDecisionTree:',(TP+TN)/(TP+TN+FN+FP))pr
int("Testing Sensitivity for Decision Tree:',(TP/(TP+FN)))print("Testing
Specificity for Decision Tree:',(TN/(TN+FP)))print("TestingPrecision
forDecision Tree:',(TP/(TP+FP)))
```

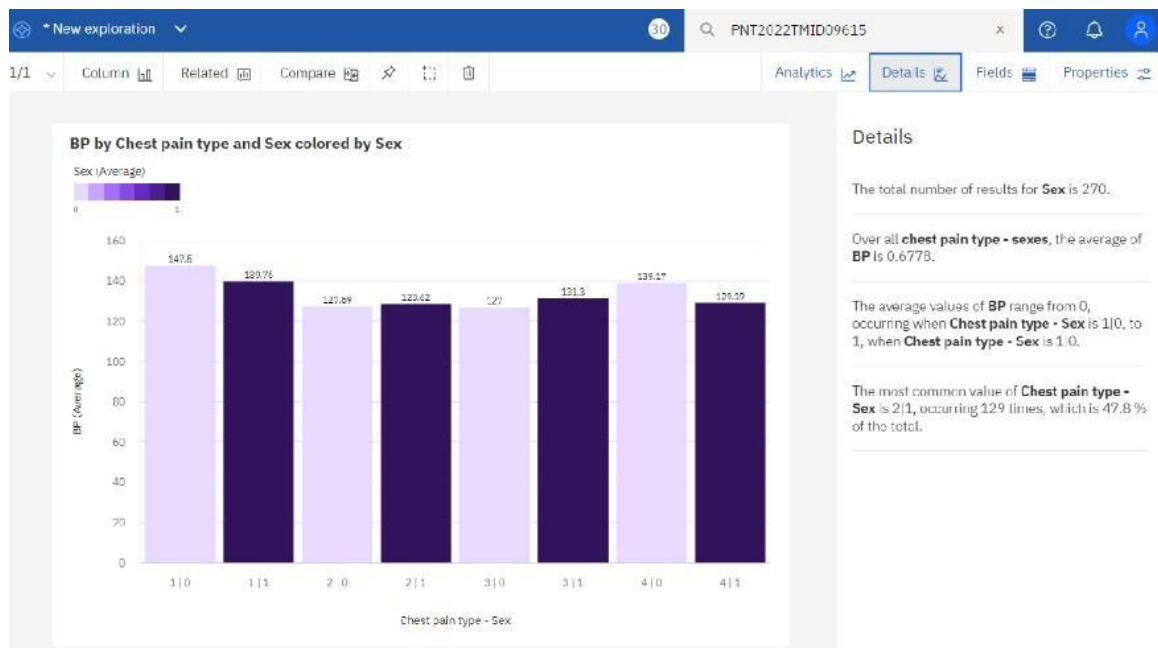
```
Testing Accuracy for Decision Tree: 0.9264705882352942
Testing Sensitivity for Decision Tree: 0.8888888888888888
Testing Specificity for Decision Tree: 1.0
Testing Precision for Decision Tree: 1.0
```

Dashboard

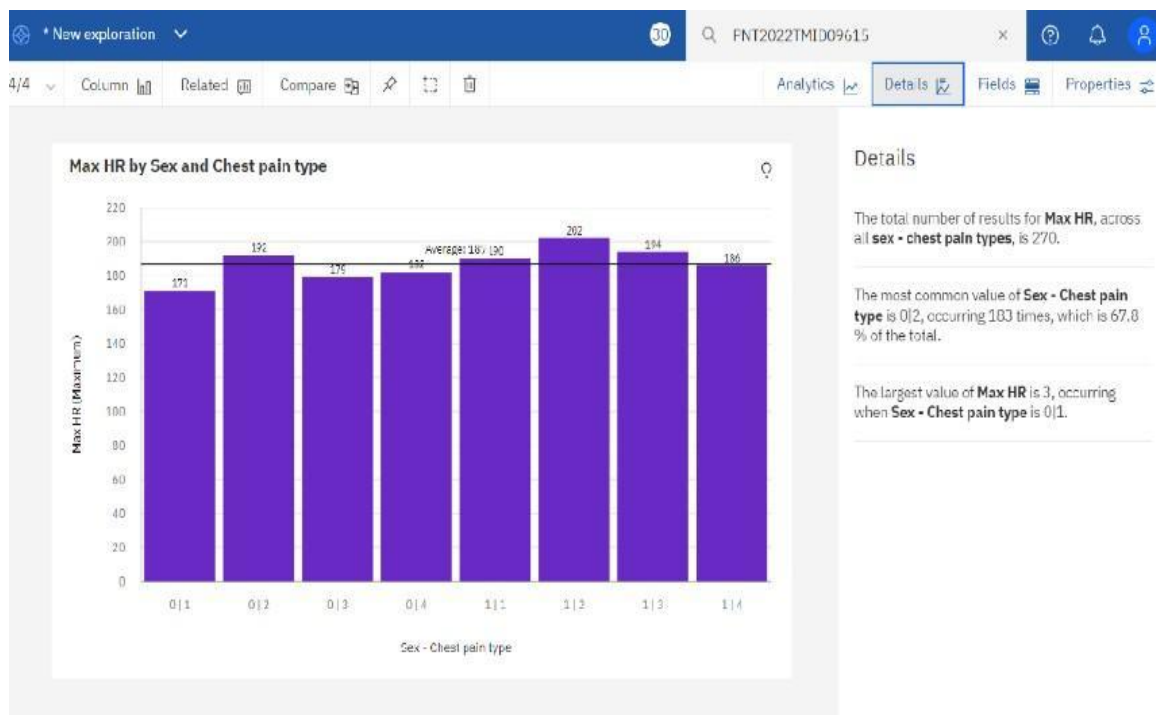
AverageBPduringchestpain



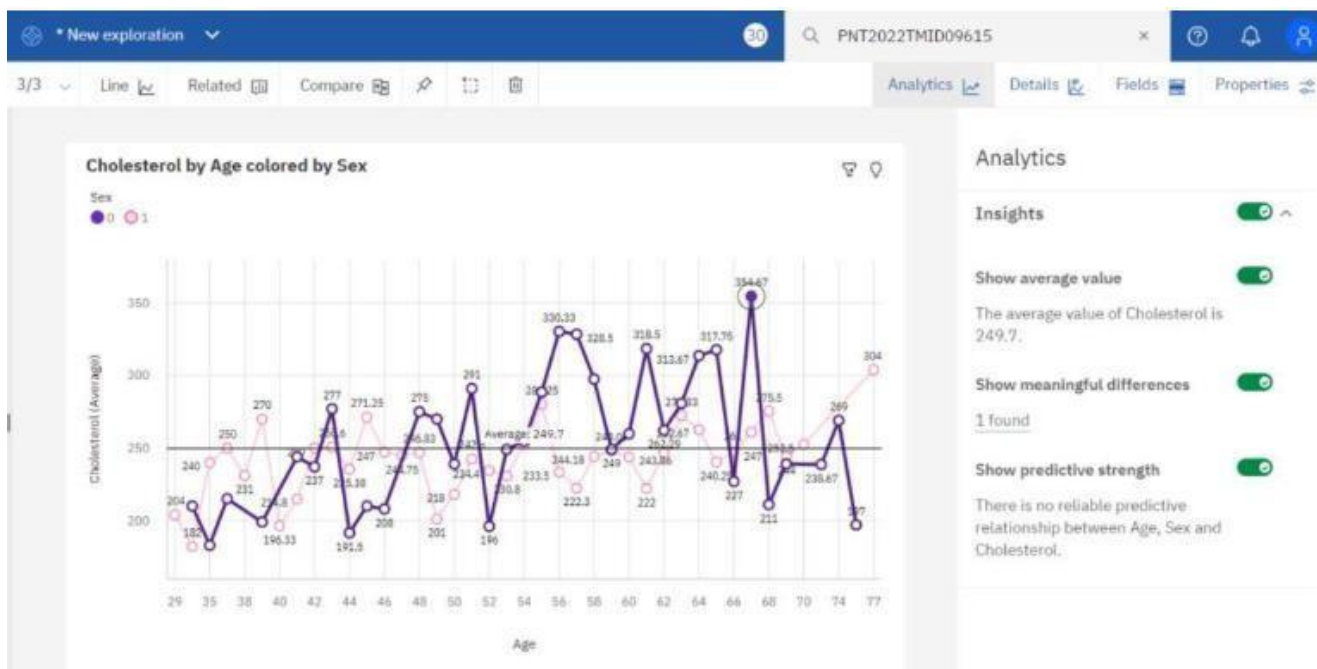
ExplorationOfBPvsChestPainTypeAndGender:



ExplorationOfMax HeartRateDuring TheChestPain:



ExplorationOfCholesterol byageandGender:

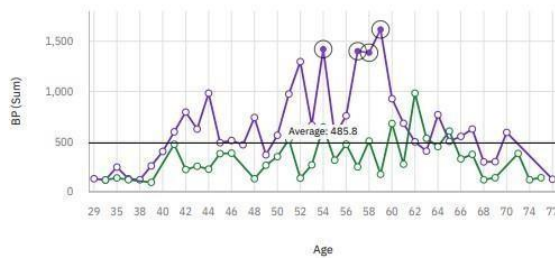


DashboardShowingDifferentTypesOfVisuals:

Tab 8

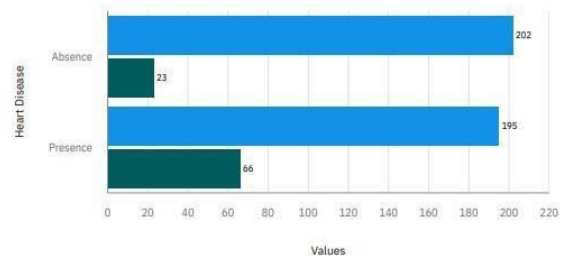
BP by Age colored by Sex

Sex
● 0 ● 1



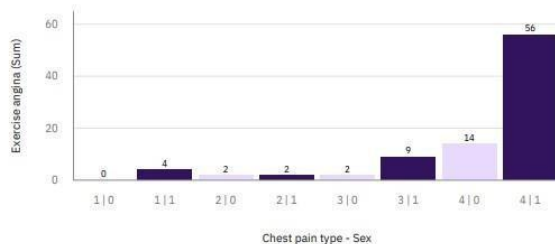
Max HR and Exercise angina by Heart Disease

Measures
● Max HR ● Exercise angina



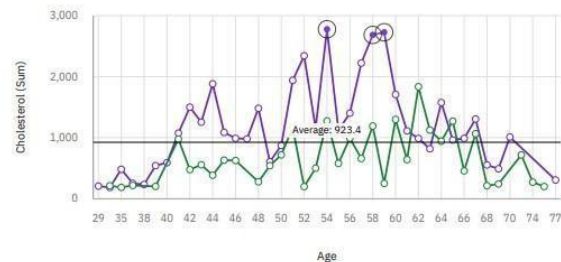
Exercise angina by Chest pain type and Sex colored by Sex

Sex (Average)
0 1



Cholesterol by Age colored by Sex

Sex
● 0 ● 1



8. Testing

TestCases

Testing the data model for various input values.

```
In [ ]: from sklearn.metrics import accuracy_score
input=(63,1,3,145,200,150,98,0,0,0,0,0)
input_as_numpy=np.asarray(input)
input_resaped=input_as_numpy.reshape(1,-1)
pre1=tree_model.predict(input_resaped)
print(pre1)
a1=accuracy_score(pre1,model1.predict(input_resaped))*100
print(a1)

['Absence']
100.0

In [ ]: 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_resaped=input_as_numpy.reshape(1,-1)
pre1=tree_model.predict(input_resaped)
print(pre1)
a1=accuracy_score(pre1,model1.predict(input_resaped))*100
print(a1)

['Presence']
100.0
```

User acceptance Testing

Testing case where user has heart disease

localhost:4200

Exercise angina (exercise induced angina (1 = yes; 0 = no))

0

ST depression

2.4

Slope of ST

2

Number of vessels fluro (number of major vessels (0-3) colored by flourosopy)

3

Thallium: 3 = normal; 6 = fixed defect; 7 = reversable defect

3

Submit

localhost:4200 says:
The patient has increased risk of heart diseases

OK

Testing case where user does not have heart disease

localhost:4200

Max HR (maximum heart rate achieved)

160

Exercise angina (exercise induced angina (1 = yes; 0 = no))

0

ST depression

1.6

Slope of ST

2

Number of vessels fluro (number of major vessels (0-3) colored by flourosopy)

0

Thallium: 3 = normal; 6 = fixed defect; 7 = reversable defect

7

Submit

localhost:4200 says:
The patient has no risk of heart diseases

OK

9. Result

PerformanceMetrics

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

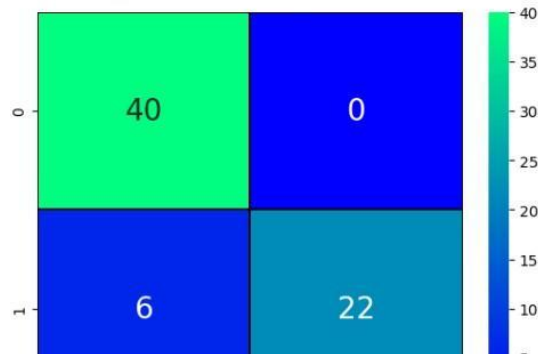
```
from sklearn.model_selection import RandomizedSearchCV
from sklearn.tree import DecisionTreeClassifier

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. AdvantagesDisadvantages

Advantages:

- Thisisoneofthefastest waystodetermineif apersonislikelytosufferfromaheartdiseaseor not.
- Usefulformedicalpractitionerstoeasilyclassifytheirpatients.
- UserFriendly
- Easytounderstand
- Secure
- Dashboardprovidesinsightfulinformations

Disadvantages:

- Needswork
- Usersneedtoknowallthe fields
- DoesNottakenull value asinput
- Doesnotprovide suggestionstouser

11. Conclusion

Complications of heart disease include heart attack andstroke. You can reduce the risk ofcomplications with early diagnosis and treatment. So the suggestion that we get from the websitemighthelp savepatients. Itis alwaysto get treatedin theearly stagesof heartdisease.

12. FutureScope

Like the saying goes “Prevention is better than cure”. We have to look into methods to preventheartdiseases altogether otherthan justpredictingit in early stages.

To use this website we need to take a lot of tests beforehand. So it would be better if we requirelessattributes and still givean effectiveresult

13. Appendix

SourceCode:

https://github.com/IBM-EPBL/IBM-Project-41593-1660643229/tree/main/PROJECT_1/Final%20deliverables/source%20code

Demovideolink:https://drive.google.com/file/d/1EG_4OwBlu5CPvxFeDNdWq5OAv8ID04Yt/view?usp=share_link