VISUALIZING AND PREDICTING HEART DISEASES WITH AN INTERACTIVE DASHBOARD

Documentation

Date	19 November 2022
Team ID	PNT2022TMID52562
Project Name	Visualizing and Predicting Heart Diseases with an Interactive Dashboard.

Guided By,

Senthilvel A N

Team Members:

Team Lead : Dhakshinya B G (1905009)

Member 1: Dharani S (1905010)

Member 2: Janani K (1905019)

Member 3: Shruthi S (1905052)

TABLE OF CONTENT

CHAPTER-1

INTRODUCTION

1.1 PROJECT OVERVIEW

. Regarding the different causes of heart disease, analyzing what causes heart disease has become mainstream nowadays.After an in-depth understanding of data analysis and machine learning-related knowledge, data analysis and data training are carried out on a dataset.This project proposes a prediction model to predict whether a people have a heart disease or not and visualize it using a dashboard.

1.2 PURPOSE

With the ever increasing population of the world, diseases and their possibilities are also increasing at an alarming rate. As time passes by, diagnosing diseases and providing appropriate treatment at the right time has become quite a challenge. Heart diseases, for one, have been a major cause of death worldwide. Finding an efficient way to predict the chances of a heart failure is indeed a need.

CHAPTER - 2
REQUIREMENTS

2.1 EXISTING PROBLEM

Existing solutions are found to have,

1. Inefficient and inaccurate heart disease prediction system

2. No proper assistance of medical professionals in evaluating a patient's heart disease based on the clinical data of the patient.

2.2 PROBLEM STATEMENT DEFINITION

Day by day the cases of heart diseases are increasing at a rapid rate and it's very important and concerning to predict any such diseases beforehand. This diagnosis is a difficult task i.e. it should be performed precisely and efficiently.

2.3 REFERENCES

1   A. Ed-Daoudy and K. Maalmi, "Real-time machine learning for early detection of heart disease using big data approach," 2019 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS), 2019, pp. 1-5, doi: 10.1109/WITS.2019.8723839.

2   K. G. Dinesh, K. Arumugaraj, K. D. Santhosh and V. Mareeswari, "Prediction of Cardiovascular Disease Using Machine Learning Algorithms," 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), 2018, pp. 1-7, doi: 10.1109/ICCTCT.2018.8550857.

3   P. Motarwar, A. Duraphe, G. Suganya and M. Premalatha, "Cognitive Approach for Heart Disease Prediction using Machine Learning," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020, pp. 1-5, doi: 10.1109/ic-ETITE47903.2020.242.

4   A. Rahim, Y. Rasheed, F. Azam, M. W. Anwar, M. A. Rahim and A. W. Muzaffar, "An Integrated Machine Learning Framework for Effective Prediction of Cardiovascular Diseases," in IEEE Access, vol. 9, pp. 106575-106588, 2021, doi: 10.1109/ACCESS.2021.3098688.

5   R. Indrakumari, T. Poongodi, Soumya Ranjan Jena,Heart Disease Prediction using

6   Exploratory Data Analysis,Procedia Computer Science,Volume 173,2020,Pages 130-139,ISSN 1877-0509,https://doi.org/10.1016/j.procs.2020.06.017.

7   Renugadevi, G & Priya, G & Sankari, B & Rajamanickam, Gowthamani. (2021). Predicting heart disease using hybrid machine learning model. Journal of Physics: Conference Series. 1916. 012208. 10.1088/1742-6596/1916/1/012208.

Xinyu Zhang. 2021. Using Data Visualization to Analyze the Correlation of HeartDisease Triggers and Using Machine Learning to Predict Heart Disease. In 2021 3rd International Conference on Intelligent Medicine and Image Processing (IMIP '21). Association for Computing Machinery, New York, NY, USA, 127–132. https://doi.org/10.1145/3468945.3468966

2   Habib, Sumaya, Moin, Maisha Binte, Aziz, Sujana. 2018-12.Heart failure risk prediction and medicine recommendation system using exploratory analysis and big data analytics.BRAC University http://hdl.handle.net/10361/11446
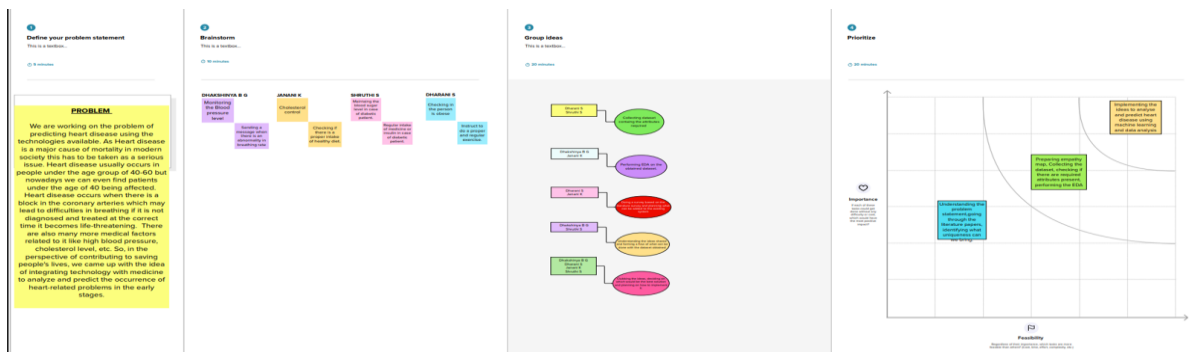
CHAPTER – 3

IDEATION AND PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS

3.2 IDEATION & BRAINSTORMING

3.3 PROPOSED SOLUTION

S.No.	Parameter	Description
1.	Problem Statement (Problem to besolved)	To analyze and predict the possibility of having heart disease in the near future withthe help of an interactive dashboard.
2.	Idea / Solution description	With the help of congos analysis, we can predict the occurrence of heart-related issues in a much early stage and can providean alert to the patient regarding the same.
3.	Novelty / Uniqueness	With greater accuracy in prediction, we alertthe patients to get treatment sooner or to take their medicine which may in turn reduce the mortality rate.
4.	Social Impact / CustomerSatisfaction	⇨ User-friendly dashboard ⇨ Alerts the patient to have a proper consultation before the situation getsworse ⇨ Saves life
5.	Business Model (Revenue Model)	⇨ Frequent updates according to the improving technology and medicalfacilities ⇨ Easy access ⇨ Patient data confidentiality
6.	Scalability of the Solution	⇨ Updation in the system does notdelete/affect the existing one ⇨ Platform independent

3.4 PROBLEM SOLUTION FIT

1.CUSTOMER SEGMENT(S)	2.CUSTOMER CONSTRAINTS	3.AVAILABLE SOLUTIONS
The patients who were affected by heart disease theirdetails can be collected. This helps to discover the characteristics of the customer.It is based on analyzes the customer bases on their data including behavioral data.	You may experience various types of emotional distress or behavioural Disturbance. Problems such as depression and anxiety are especially common. These conditions not only affect your emotions state but can alsoimpact on your symptoms of heart disease.	Surgery procedure may be done to open to open a blocked artery. Surgeries and procedures to treat a heart attack include: coronary angioplasty and stenting. This procedure is done to open clogged heart arteries.

2.JOBS-TO-BE DONE/PROBLEMS	9.PROBLEM ROOT CAUSE	7.BEHAVIOUR
To cure from the disease, regular treatment for the patient is very important. Nowadays, the digital machines are developed and it is used to monitoring the patients health condition such as pulse rate, diabetes, pressure etc,. Checking health condition is very important because by taking health test will able to know the disease in our body.This may lead to various problems that damage our health condition.	Coronary artery disease (CAD)is a main cause of heart attack.A less common cause is a seven spasm, or sudden contraction, ofa coronary, of a coronary artery that can stop blood flow to the heart muscle.	Your lifestyle can increase therisk for heart disease. Eating a diet high in saturated fats, trans fat, and cholesterol bas been linked to heart disease and related conditions, such as astherosclerosis. Also, too muchsalt in the diet can raise blood pressure.

3.TRIGGERS

Leading risk factors for heart disease and stroke are blood pressure,low density lipoprotein, cholesterol, diabetes, smoking and secondhand smoke exposure, obesity, unhealthy diet, and physical inactivity.

4.EMOTIONS BEFORE/AFTER

The range of emotions after a heart attack can be overwhelming andvary based on a survivor's individual experience. You might feel angry, vulnerable, anxioius, depressed hopeful, or even Inspried.

10.YOUR SOLUTION

Heart disease treatment depends on the cause and typesof heart damage. Healthy lifestyles 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.

8.CHANNELS O BEHAVOIUR

8.1 ONLINE
The channels which they published or telecast the program which conveys the guidelines and prevention ofheart disease.

8.2 OFFLINE
Giving awareness and campsto prevention to heart disease.

CHAPTER - 4

REQUIREMENT ANALYSIS

4.1   FUNCTIONAL REQUIREMENT

Purpose
This section includes background, definitions, and system overview.

Overall description
The description document consists of Health vision, Prediction rules, and assumptions.

Specific requirements
The requirements might be database requirements, system attributes, and functional requirements.

Here is a project definition example:
"*Admin dashboard - a web portal allowing Admin to view and manage Applicants and Customers, Check up details, Health reports, manage medications, prices, and review statistics.*"

Use cases
Use cases describe the interaction between the system and external users that leads to achieving particular goals.

Each use case includes the main elements:
• Actors are the users who will interact with your product.
Example:
 "*Applicant - a person who wants to use the portal and applies for registration.*
Patient - a person who was approved for registration and can use the portal.
Doctor - a person, who was registered in the portal and fulfills the requirements for Patient."
• System functional requirements describe the intended behavior of the project.
Example:
"*Visualization System helps Patients for visualization.*"
• Goals describe all interactions between the users and the system.
Example:
"*When the visualization is over, Doctor marks it has ended in their portal.*"

User stories
User stories are documented descriptions of software features from the end-user perspective. The document describes scenarios of how the user engages with the solution.
Example:
"*As a Patient, I want to diagnose visuals from the carousel so that I can complete the order*".
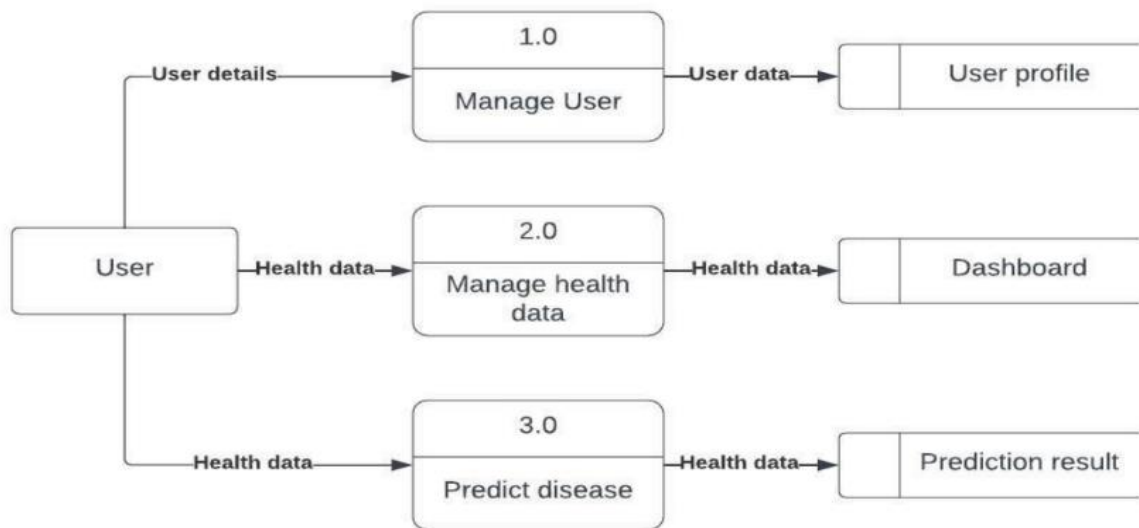
4.2   NON-FUNCTIONAL REQUIREMENTS

i)   Usability

ii)   Security

iii)   Reliability

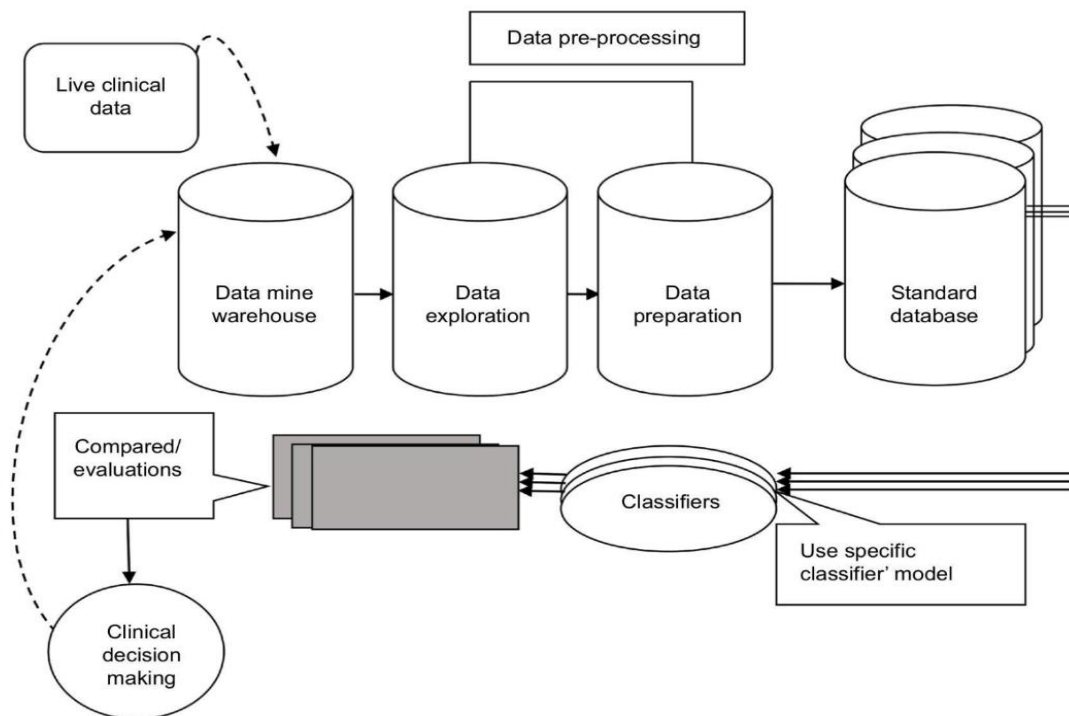iv)   Performance

v)   Availability

vi)   Scalability

CHATER - 5

PROJECT DESIGN

5.1 DATA FLOW DIAGRAM

5.2 SOLUTION ARCHITECTURE

5.3 USER STORIES

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Web user)	Registration	USN-1	As a user, I can register for the application by entering my email, password,etc…	I can access my account / dashboard, enter the details	High	Sprint-1
Customer (Web user)		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email	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	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 (Web user)	Dashboard	USN-6	The User will edit the data and more features	I can see the features add adjusted accordingly.	Medium	Sprint 2

CHAPTER - 6

PROJECT PLANNING & SCHEDULING

6.1 SPRINT PLANNING & ESTIMATION

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Story points	Priority
Sprint1	Registration	USN-1	As a user, I can register for the Cognos Analytics by entering Email, password, and confirm password.	I can access my account / dashboard.	10	High
		USN-2	As a user, I can register for the application through Gmail.	I can register & access the dashboard with gmail login.	10	High
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
Sprint3	Cognos Dashboard d	USN-4	The user is allowed to view or update his/her profile.	I can see the profile.	10	Medium

		USN-5	The user can change the password.	I am able to change the password.	10	Medium
Sprint-4	Classified result	USN-6	Home - Analyse your Heart.	I can predict the heart condition.	5	High
		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
		USN8	The report is generated based on the condition.	The user is able to view or download the report.	5	Medium

6.2 SPRINT DELIVERY SCHEDULE

SPRINT	TOTAL STORY POINTS	DURATION	SPRINT START DATE	SPRINT END DATE	STORY POINTS COMPLETED	SPRINT RELEASE DATE
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		
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022		
Sprint-4	20	6 Days	14 Nov 2022	19Nov 2022		

6.3 REPORTS

CHAPTER - 7

CODING & SOLUTIONING

7.1 FEATURE 1

7.2 FEATURE 2

7.3 FEATURE 3

CHAPTER – 8

TESTING

8.1 TEST CASES

```
df = pd.read_csv('heart.csv')

X = df.drop("target", axis=1)
y = df["target"]
X_test = np.array([58.0,0.0,0.0,100.0,248.0,0.0,0.0,122.0,0.0,1.0,1.0,0.0,2.0])
X_test = X_test.reshape((1,-1))

model = RandomForestRegressor()
model.fit(X, y)

val = model.predict(X_test)

print(int(val*100),"%")
✓ 0.3s
73 %
```

```
df = pd.read_csv('heart.csv')

X = df.drop("target", axis=1)
y = df["target"]
X_test = np.array([52.0,1.0,0.0,125.0,212.0,0.0,1.0,168.0,0.0,1.0,2.0,2.0,3.0])
X_test = X_test.reshape((1,-1))

model = RandomForestRegressor()
model.fit(X, y)

val = model.predict(X_test)

print(int(val*100),"%")
✓ 0.3s
0 %
```
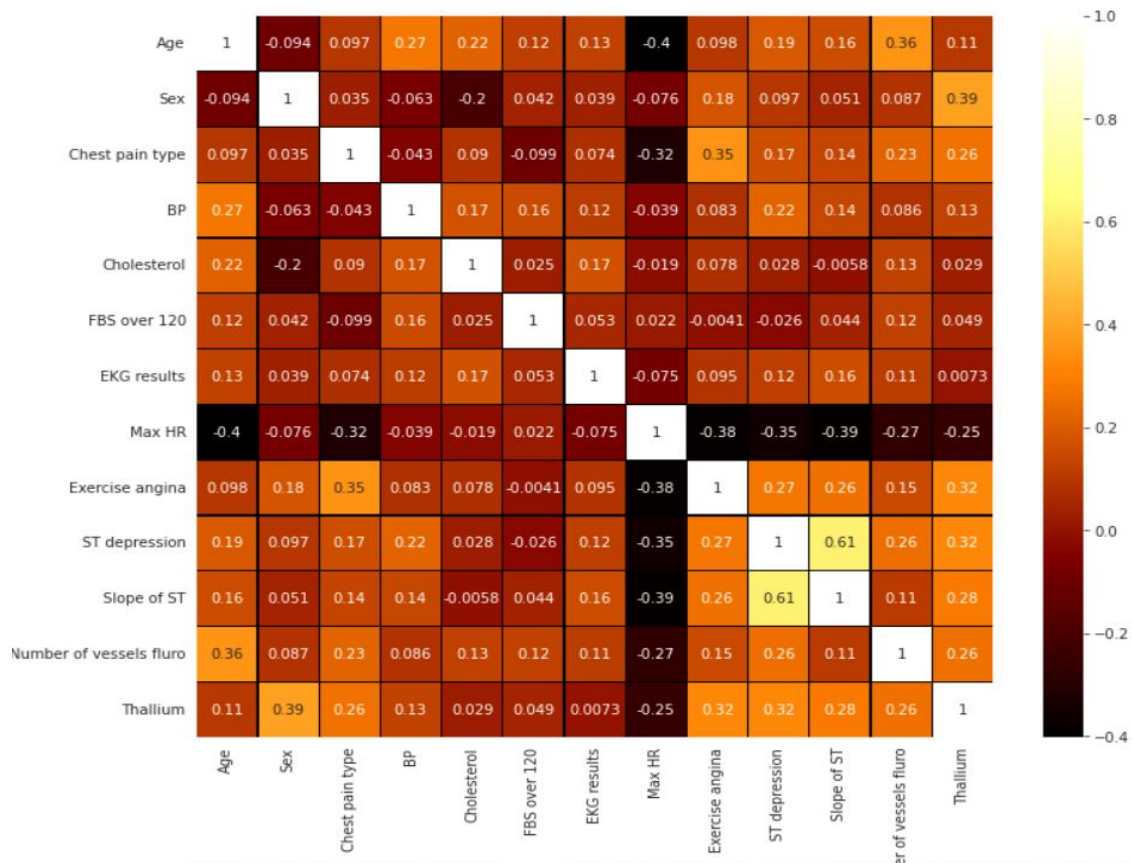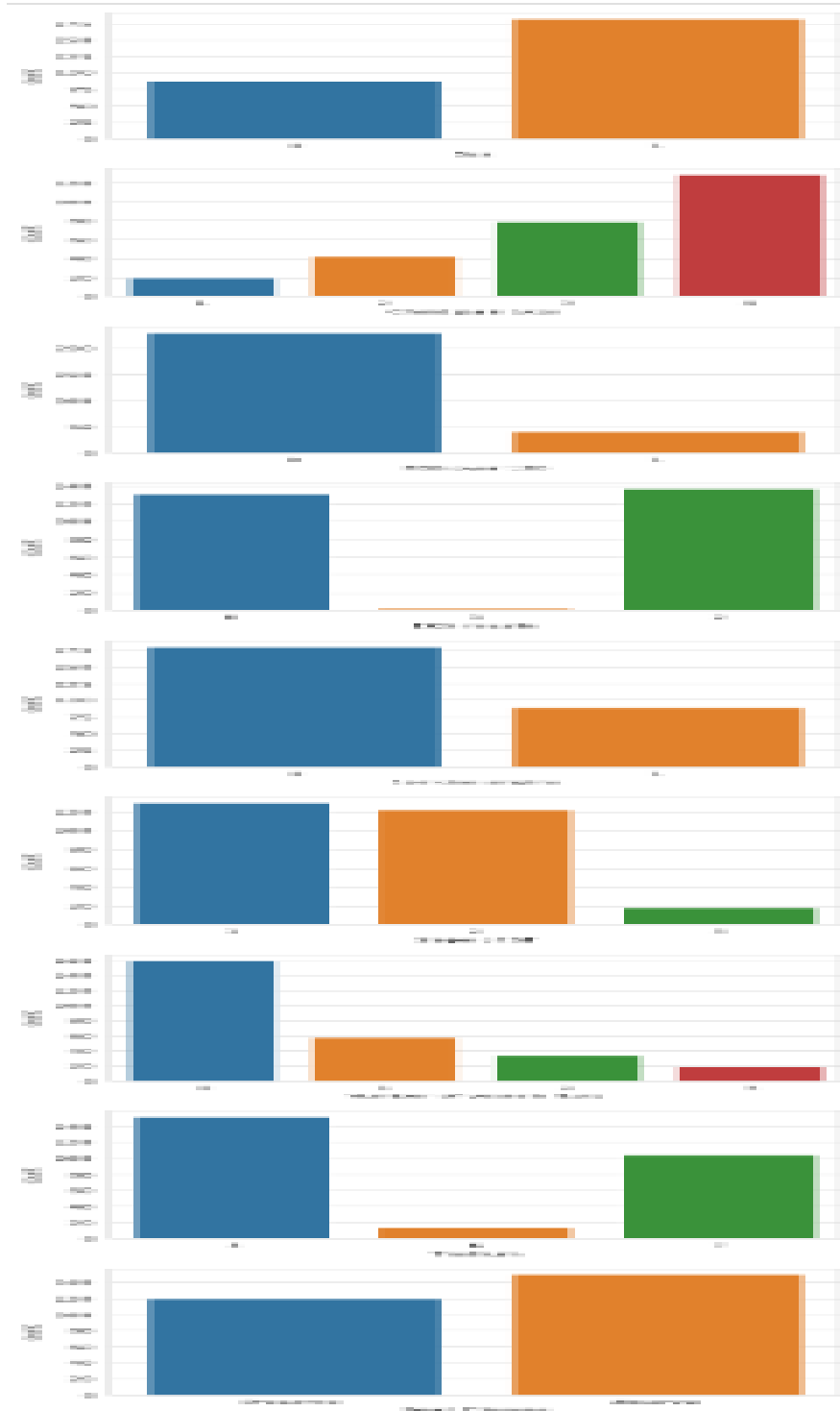
CHAPTER - 9
RESULTS
9.1PERFORMANCE METRICS

Various performance metrics of the model is analyzed and documented as below

HEAT MAP:

	Age	Sex	Chest pain type	BP	Cholesterol	FBS over 120	EKG results	Max HR	Exercise angina	ST depression	Slope of ST	Number of vessels fluro	Thallium
Age	1	-0.094	0.097	0.27	0.22	0.12	0.13	-0.4	0.098	0.19	0.16	0.36	0.11
Sex	-0.094	1	0.035	-0.063	-0.2	0.042	0.039	-0.076	0.18	0.097	0.051	0.087	0.39
Chest pain type	0.097	0.035	1	-0.043	0.09	-0.099	0.074	-0.32	0.35	0.17	0.14	0.23	0.26
BP	0.27	-0.063	-0.043	1	0.17	0.16	0.12	-0.039	0.083	0.22	0.14	0.086	0.13
Cholesterol	0.22	-0.2	0.09	0.17	1	0.025	0.17	-0.019	0.078	0.028	-0.0058	0.13	0.029
FBS over 120	0.12	0.042	-0.099	0.16	0.025	1	0.053	0.022	-0.0041	-0.026	0.044	0.12	0.049
EKG results	0.13	0.039	0.074	0.12	0.17	0.053	1	-0.075	0.095	0.12	0.16	0.11	0.0073
Max HR	-0.4	-0.076	-0.32	-0.039	-0.019	0.022	-0.075	1	-0.38	-0.35	-0.39	-0.27	-0.25
Exercise angina	0.098	0.18	0.35	0.083	0.078	-0.0041	0.095	-0.38	1	0.27	0.26	0.15	0.32
ST depression	0.19	0.097	0.17	0.22	0.028	-0.026	0.12	-0.35	0.27	1	0.61	0.26	0.32
Slope of ST	0.16	0.051	0.14	0.14	-0.0058	0.044	0.16	-0.39	0.26	0.61	1	0.11	0.28
Number of vessels fluro	0.36	0.087	0.23	0.086	0.13	0.12	0.11	-0.27	0.15	0.26	0.11	1	0.26
Thallium	0.11	0.39	0.26	0.13	0.029	0.049	0.0073	-0.25	0.32	0.32	0.28	0.26	1

CONFUSION MATRIX:

CHAPTER - 10

ADVANTAGES & DISADVANTAGES

ADVANTAGE:

- Medical professionals can quickly recognize and respond to potential dangers with the right visualization tools.

- Doctors may better define patient population and allocate resources by displaying health data in real time.

- Doctors may better forecast their patient's health and more accurate diagnoses using data visualization software equipped with predictive analytics technologies.

- Clear and simple charts and infographics improve patient awareness and increase participation, making visualization a helpful tool for patients education.

- Visualization tools can significantly improve presentations and reports.

- Good data visualization ads in making the call to improve health regulations.

DISADVANTAGE:

- Due to the risk of malpractice a visualization intended for use by health care professionals must be so clear and precise as to render misinterpretation virtually impossible.

- Insufficient data or poor structuring of data object in the database may lead to inaccurate interpretation of visualized data.

- Those with heart failure can develop swelling, dizziness, and other symptoms that can affect their ability to complete daily tasks.

CHAPTER 11

CONCLUSION

A cardiovascular disease detection model has been developed using three ML classification modelling techniques. This project predicts people with cardiovascular disease by extracting the patient medical history that leads to a fatal heart disease from a dataset that includes patients' medical history such as chest pain, sugar level, blood pressure, etc. This Heart Disease detection system assists a patient based on his/her clinical information of them been diagnosed with a previous heart disease. Use of more training data ensures the higher chances of the model to accurately predict whether the given person has a heart disease or not . By using these, computer aided techniques we can predict the patient fast and better and the cost can be reduced very much There are a number of medical databases that we can work on as these Machine learning techniques are better and they can predict better than a human being which helps the patient as well as the doctors. Thus, data visualization and data analysis of the target variable, age features, and whatnot along with its univariate analysis and bivariate analysis.

Thus, data visualization and data analysis of the target variable, age features, and whatnot along with its univariate analysis and bivariate analysis. Thus in conclusion, the algorithms used in building the given model are Random Forest Classifier. The accuracy of our model is 73%.

CHAPTER-11
CONCLUSION

Diagnosis of cardiac disease is the sternest challenge in the medical profession. It is based on the thorough review by medical experts of the various clinical and medical data of the patient.The methods which are used for comparison of ML models are confusion matrix, heat map.. For the 13 features which were in the dataset, KNeighbors classifier performed better in the ML approach when data preprocessing is applied.

CHAPTER-12
FUTURE SCOPE

● The current project does for prediction and visualization of heart diseases. It can be extended to add features related to hospital management.

● The dataset size can be increased and then deep learning with various other optimizations can be used and more promising results can be achieved.

● More ways could be found where we could integrate heart-disease-trained ML and DL models with certain multimedia for the ease of patients and doctors.

● If a large dataset is present, the results can increase very much in deep learning and ML as well

● The prediction of heart diseases by using advanced techniques and algorithms in less time complexity