VISUALIZING AND PREDICTING HEART DISEASES WITH AN INTERACTIVE

Submitted by

ASHIKA NOORIYA.M 110119205012

FARHEEN OMAR 110119205015

YUSRA SABA.M 110119205040

AALIYA FATHIMA.F 110119205301

In partial fulfillment for the award of the degree

Of

BACHELOR OF TECHNOLOGY

IN

INFORMATION TECHNOLOGY

AALIM MUHAMMED SALEGH COLLEGE OF ENGINEERING

ANNA UNIVERSITY:CHENNAI-600025

NOVEMBER-2022

BONAFIDE CERTIFICATE

Certified that this project report is Visualizing and Predicting Heart Diseases with an Interactive DashBoard the bonafide work of "**ASHIKA NOORIYA.M (110119205012),FARHEENOMAR (110119205015),YUSRASABA.M(110119205040), AALIYAFATHIMA.F(110119205301)**"who carried out the project work under my supervision

SIGNATURE
DR.S.ARIF ABDUL
RAHAMAN
HEAD OF DEPARTMENT
Department of Information Technology
Aalim Muhammed Salegh College of Engineering
Chennai-600025

SIGNATURE
ASSTPROF
RAEESHATHULHAFELA K R
SUPERVISOR
ASSISTANT PROFEESOR
Department of Information Technology
Aalim Muhammed Salegh College of Engineering
Chennai-600025

ACKNOLOWLEDEMENT

FIRST and foremost we would like to thank the god of almighty who is our refuge and strength. We would like to express our heartfelt thanks to our BELOVED PARENTS who sacrifice their presence for our better futureWe are very much indebted to thank our college founder ALHAJ,DR.S.M.SHAIK NURDDIN, chairperson JANABAALHAJIYANI M.S.HABIBUNNISA,Aalim Muhammed Salegh CollegeGroup of educational Institutions,Honorable Secretary&CorrespondentJANAB ALHAJ S.SEGU JAMALUDEEN, Aalim Muhammed SaleghCollege Group of educational Institutions for providing necessary facilities all through the course We take this opportunity to put forth our deep sense of gratitude to ourbeloved Principal Prof.DR.M SATHISH for granting permission to undertake the project We also express our gratitude to Prof.DR.S.ARIF ABDULRAHAMAN,Head of Department ,Information Technology, Aalim Muhammed SaleghCollege Engineering,Chennai,for his involvement and constantsupport,valuable guidance to effectively complete our project in ourcollegeWe would like to express our sincere thanks to our project guide RAEESHATHULHAFELA K R Assistant Professor,Information Technology,Aalim Muhammed Salegh College Engineering,who perusuaded us totake this project and never ceased to lend his encouragement and supportFinally,we take immense pleasure to thank our family member,facultiesand friends for their constant support and encouragement in doing our project

CERTIFICATE OF EVALUATION

COLLEGE NAME : Aalim Muhammed Salegh College OF Engineering

BRANCH : Information Technology

PROJECT TITLE:Visualizing and Predicting Heart Diseases with an Interactive Dash Board

NAME OF THE STUDENT	REGISTERATION NUMBER	NAMEOF THE SUPERVISOR
ASHIKA NOORIYA.M	110119205012	
FARHEEN OMAR	110119205015	RAEESHATHUL
YUSRA SABA.M	110119205040	HAFELA K R
AALIYA FATHIMA.F	110119205301	

The report of the mini project work submitted by the above students inpartial fulfilment for award of Bachelor ofTechnology Degree inINFORMATION TECHNOLOGY of Anna University were evaluated and confirmed to be reports of the work done by above student during the academic year 2019-2023

TABLE OF CONTENT

CHAPTER 1:INTRODUCTION

1. Project Overview
2. Purpose

CHAPTER 2:LITERATURE SURVEY

3. Existing problem
4. References
5. Problem Statement Definition

CHAPTER 3:IDEATION & PROPOSED SOLUTION

6. Empathy Map Canvas
7. Ideation & Brainstorming
8. Proposed Solution
9. Problem Solution fit

CHAPTER 4:REQUIREMENT ANALYSIS

10 Functional requirement

11 Non-Functional requirements

CHAPTER 5 :PROJECT DESIGN

12 Data Flow Diagrams

13 Solution & Technical Architecture

14 User Stories

CHAPTER 6 :PROJECT PLANNING & SCHEDULING

15 Sprint Planning & Estimation

16 Sprint Delivery Schedule

17 Reports from JIRA

CHAPTER 7:CODING & SOLUTIONING

(Explain the features added in the project along with code)

18 Feature 1

19 Feature 2

20 Database Schema (if Applicable)

CHAPTER 8:TESTING

21 Test Cases

22 User Acceptance Testing

CHAPTER 9:RESULTS

23 Performance Metrics

CHAPTER 10:ADVANTAGES & DISADVANTAGES

CHAPTER 11:CONCLUSION

CHAPTER 12:FUTURE SCOPE

CHAPTER 13:APPENDIX

Source Code

GitHub & Project Demo Link

CHAPTER 1
INTRODUCTION

This report represents to seventh semester students for the partial fulfillment of COMP 484, Machine Learning, given by the department of Information Technology, KU. Cardiovascular diseases are the most common cause of death worldwideover the last few decades in the developed as well as underdeveloped and developing countries. Early detection of cardiac diseases and continuous supervision of clinicians can reduce the mortality rate. However, it is not possible to monitor patients every day in all cases accurately and consultation of a patient for 24 hours by a doctor is not available since it require more sapience, time and expertise. In this project, we have developed and researched about models for heart disease prediction through the various heart attributes of patient and detect impending heart disease using Machine learning techniques like backward elimination algorithm, logistic regression and REFCV on the dataset available publicly in Kaggle Website, further evaluating the results using confusion matrix and cross validation. The early prognosis of cardiovascular diseases can aid in making decisions on lifestyle changes in high risk patients and in turn reduce the complications, which can be a great milestone in the field of medicine.

Exploratory Data Analysis(EDA)is classified into Graphical or non-graphical and Univariate or multivariate Univariate data consider one data column at a time while multivariate method considers more than two variables while analyzing. The diagnostic methods of diseases are of two types namely, Invasive and Non-invasive

PROJECT OVERVIEW

This report represents the mini-project assigned to seventh semester students for the partial fulfillment of COMP 484, Machine Learning, given by the department of computer science and engineering, KU. Cardiovascular diseases are the most common cause of death worldwide

over the last few decades in the developed as well as underdeveloped and developing countries. Early detection of cardiac diseases and continuous supervision of clinicians can reduce the mortality rate. However, it is not possible to monitor patients every day in all cases accurately and consultation of a patient for 24 hours by a doctor is not available since it requires more sapience, time and expertise.

In this project, we have developed and researched about models for heart disease prediction through the various heart attributes of patient and detect impending heart disease using Machine learning techniques like backward elimination algorithm, logistic regression and REFCV on the dataset available publicly in Kaggle Website, further evaluating the results using confusion matrix and cross validation.

The early prognosis of cardiovascular diseases can aid in making decisions on lifestyle changes in high risk patients and in turn reduce the complications, which can be a great milestone in the field of medicine

keywords: Machine Learning, Logistic regression, Cross-Validation, Backward Elimination, REFCV, Cardiovascular Disease

PURPOSE:

Heart disease refers to several types of abnormalities in heart conditions. The leading cause of death is heart disease. It is infeasible for a common man to frequently undergo tests for ECG and so on. Hence,there needs a replacement for this, which must be handy and reliable.

Idea / Solution description The idea behind the proposed solution is to propose an interactive dashboard for visualizing and predicting heart diseases in which user can view his/her medical report

analysis and the predicted final result. The dashboard will be generated using IBM Cognos. The heart disease will be predicted using Naïve Bayes Algorithm.

Novelty / Uniqueness The novelty behind the proposed system is to provide suggestions to the user based on his/her medical analysis. It will

provide the preventive measures to take care of the user himself.

Social Impact / Customer Satisfaction The system helps the user as well as the doctor to make better decisions to predict heart disease. It is useful in predicting the disease in an earlier stage and makes the user alert about his current condition periodically.

Business Model (Revenue Model) This interactive dashboard for heart disease prediction can be deployed in Health care centres and Hospitals, so that it makes the analysis in a fast manner.

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.

CHAPTER 2
LITERATURE SURVEY

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.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. The use of analytics in healthcare improves care by facilitating preventive care and EDA is a vital step while analyzing data.

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.In the conclusion, the visualized data shows that the prediction is accurate.

EXISTING PROBLEM

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

potential for the healthcare industry to enable health systems to systematically use data and analytics to identify inefficiencies and best practices that improve care and reduce costs. According to (Wurz & Takala, 2006) the opportunities to improve care and reduce costs concurrently could apply to as much as 30% of overall healthcare spending. The successful application of data mining in highly visible fields like e-business, marketing and retail has led to its application in other industries and sectors. Among these sectors just discovering is healthcare. The healthcare environment is still „information rich" but „knowledge poor". There is a wealth of data available within the healthcare systems. However, there is a lack of effective analysis tools to discover hidden relationships and trends in the data for African genres

REFERENCES

[1] Indrakumari, R., Poongodi, T. and Jena, S.R., 2020. Heart disease prediction using exploratory data analysis. Proceed Computer Science, 173, pp.130-139.

[2] Narayanan, V.S., Raj, K.N., Kumar, K. and Kumar, M., 2021. SurvivalPrediction of Heart Failure Patients Using Lasso Algorithm and Gaussian Naive Bayes Classifier (No. 6912). EasyChair.

[3] Suvarna, C., Sali, A. and Salmani, S., 2017, July. Efficient heart disease prediction system using optimization technique. In 2017 International Conference on Computing Methodologies and Communication (ICCMC) (pp. 374-379). IEEE.

[4] Dinesh, K.G., Arumugaraj, K., Santhosh, K.D. and Mareeswari, V., 2018, March. Prediction of cardiovascular disease using machine learning algorithms. In 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT) (pp. 1-7). IEEE.

[5] Raju, C., Philipsy, E., Chacko, S., Suresh, L.P. and Rajan, S.D., 2018, March. A survey on predicting heart disease using data mining techniques. In 2018 conference on emerging devices and smart systems (ICEDSS) (pp.

253-255). IEEE.
[6] Kavitha, M., Gnaneswar, G., Dinesh, R., Sai, Y.R. and Suraj, R.S., 2021, January. Heart disease prediction using hybrid machine learning model. In 2021 6th International Conference on Inventive Computation Technologies (ICICT) (pp. 1329-1333). IEEE.
[7] Repaka, A.N., Ravikanti, S.D. and Franklin, R.G., 2019, April. Design and implementing heart disease prediction using naives Bayesian. In2019 3rd International conference on trends in electronics and informatics (ICOEI) (pp. 292-297). IEEE.
[8] Pouriyeh, S., Vahid, S., Sannino, G., De Pietro, G., Arabnia, H. and Gutierrez, J., 2017, July. A comprehensive investigation and comparison of machine learning techniques in the domain of heart disease. In 2017 IEEE symposium on computers and communications (ISCC) (pp. 204-207). IEEE.
[9] Sowmiya, C. and Sumitra, P., 2017, March. Analytical study of heart disease diagnosis using classification techniques. In 2017IEEEInternational Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS) (pp. 1-5). IEEE.
[10] Islam, M.R., Liu, S., Razzak, I., Kabir, M.A., Wang, X., Tilocca, P. and Xu, G., 2020, November. Mhivis: Visual analytics for exploring mental illness of policyholders in life insurance industry. In 2020 7th international conference on behavioral and social Computing (BESC) (pp. 1-4). IEEE.
[11] Rahman, M., Islam, M.R., Akter, S., Akter, S., Islam, L. and Xu, G., 2021. DiaVis: Exploration and Analysis of Diabetes through Visual Interactive System. Human-Centric Intelligent Systems, 1(3-4), pp.75-85.
[12] Kwon, B.C., Anand, V., Severson, K.A., Ghosh, S., Sun, Z., Frohnert, B.I., Lundgren, M., Ng, K., 2020. Dpvis: Visual analytics with hidden markov models for disease progression pathways. IEEE transactions on visualization and computer graphics.
[13] Ayon, S.I., Islam, M.M. and Hossain, M.R., 2020. Coronary artery heart disease prediction: a comparative study of computational

intelligence techniques. IETE Journal of Research, pp.1-20.
[14] Kondababu, A., Siddhartha, V., Kumar, B.B. and Penumutchi, B., 2021. A comparative study on machine learning based heart disease prediction. Materials Today: Proceedings.
[15]Rise in cases of asymptomatic heart attacks amongst middle aged people", [Online] Available https://www.maxhealthcare.in/blogs/rise-cases-asymptomatic-heart-attacks-amongst-middle-aged-people
[16]"Multiplier Design using RBSD Number System", A. A. S Awwal and J. U.Ahmed, Wright State University, Computer Science & Engineering Department, Dayton, Ohio
[17]"High Speed VLSI Multiplication Algorithm with Addition Tree", H.Yasuura, N.Takagi and S.Yajima, IEEE Trans. Comp., C-34, pp. 789-795, 1985
[18]"On-Line Error Detectable High-Speed Multiplier using Redundant Binary Representation and Three-Rail Logic", N. Takagi and S. Yajima, IEEE Tran. Comp, Vol. C-36, No. 11, Nov. 1987.

[19]"Fast Computing using Signed Digit Number System", Neelam Sharma, Reena Rani and L.K. Singh, IEEE Explorer2009.

PROBLEM STATEMENT

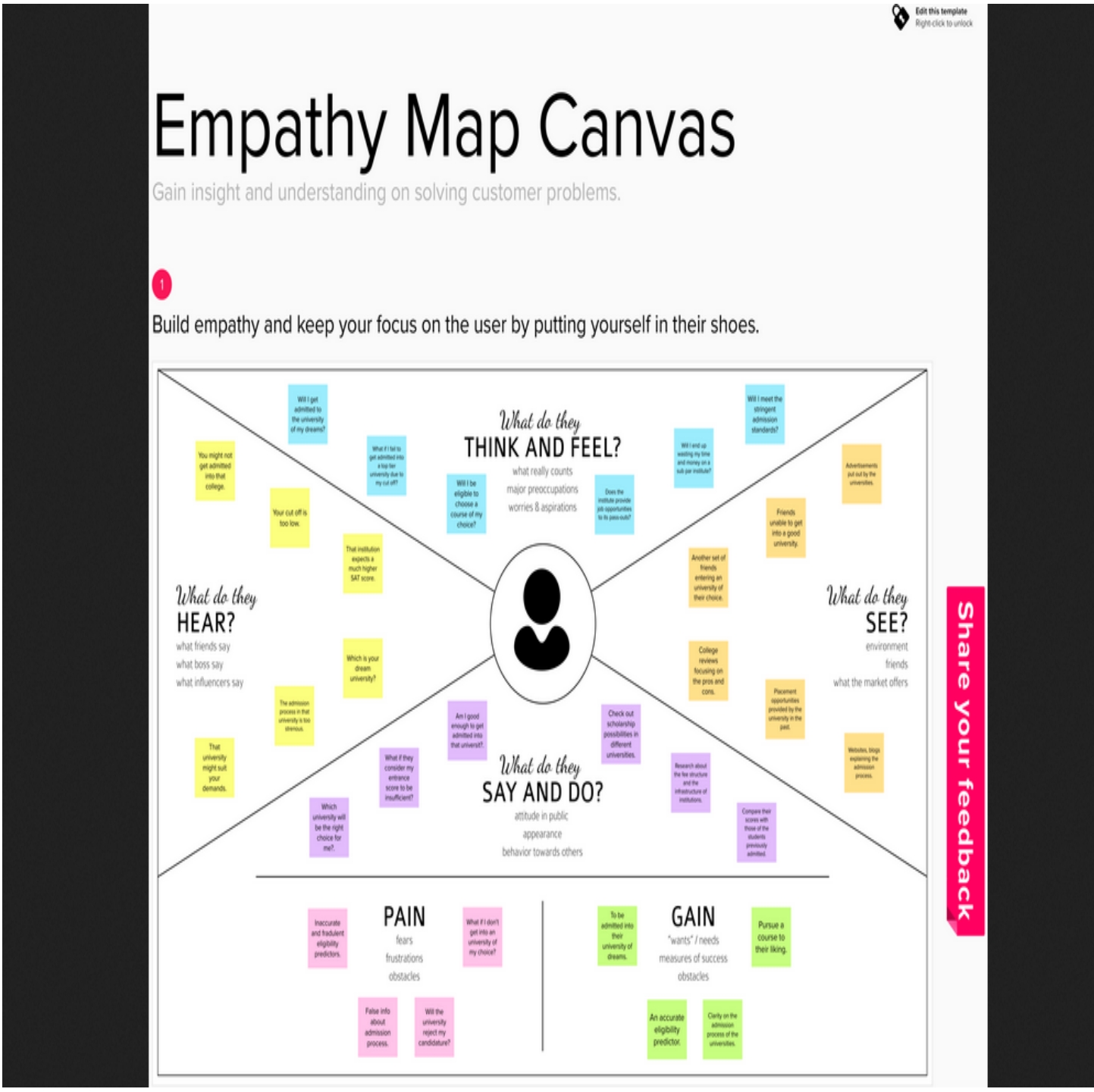
The major challenge in heart disease is its detection. There are instruments available which can predict heart disease but either it are expensive are not efficient to calculate chance of heart disease in human. Early detection of cardiac diseases can decrease the mortality rate and overall complications. However,it is not possible to monitor patients everyday in all cases accurately and consultation of a patient for 24 hours by a doctor is not available since it requires more sapience, time and expertise.Since we have a good amount of data in today's world, we can use various machine learning algorithms to

analyze the data for hidden patterns. The hidden patterns can be used for health diagnosis in medicinal data.

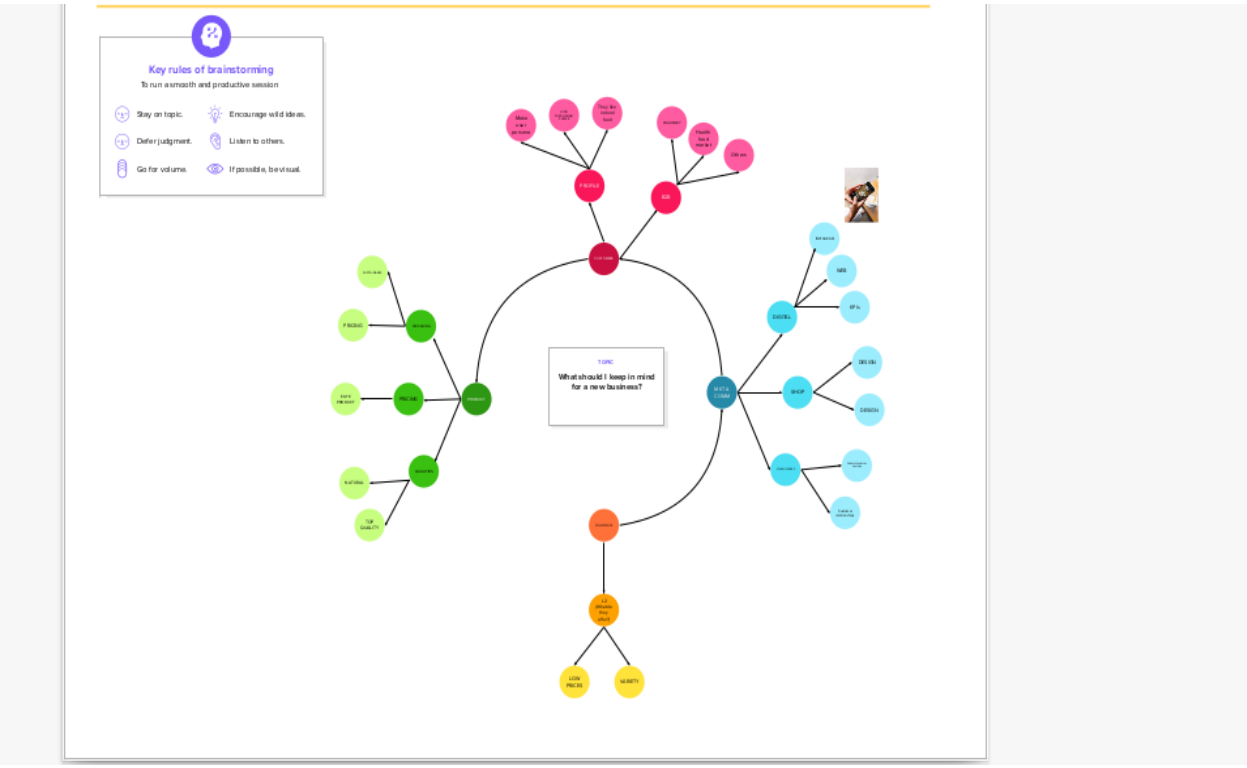
CHAPTER 03

IDEATION & PROPOSED SOLUTION

Empathy Map Canvas



Ideation & Brainstorming



Proposed Solution

Project team shall fill the following information in proposed solution template.
problem statement(problem to me solved)
➢ To develop an interactive dashboard to predict the heart disease accurately with few tests and attributes the presence of heart disease.
2. Idea / Solution description ➢ Analyzing data and identifying the heart
3. Novelty / Uniqueness ➢ Hoping to achieve maximum accuracy to provide prior treatment to the patients and reduce the fatality rate.
4. Social Impact / Customer Satisfaction ➢ Saving lives, User friendly interactive dashboard.Reduces the biases and mistakes caused by the decisions of doctors based on their intuitions and experiences.
5. Business Model (Revenue Model) ➢ Data security.
➢ Easy to use.

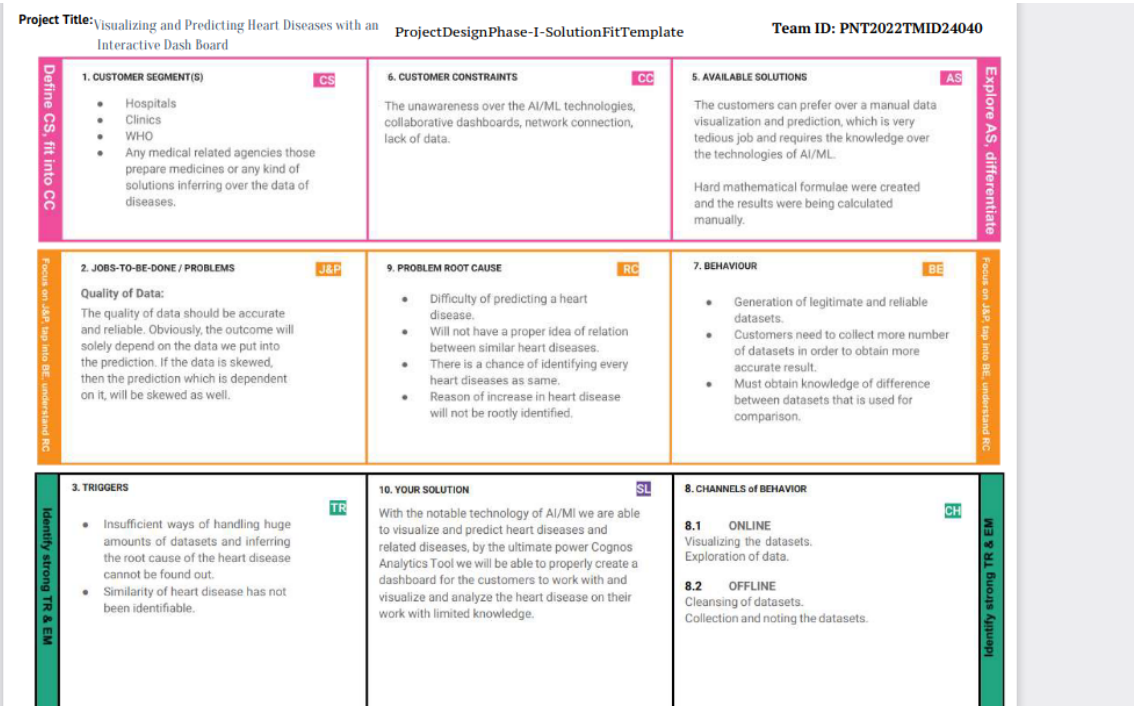
➢ Constant updates according to necessity.

6. Scalability of the Solution ➢ Can be used in any platform (Windows, mac, etc.,)

➢ Adding new feature doesn’t affect the performance of the system.

➢ Scalable dataset.

PROBLEM SOLUTION FIT



Project Title: Visualizing and Predicting Heart Diseases with an Interactive Dash Board

ProjectDesignPhase-I-SolutionFitTemplate

Team ID: PNT2022TMID24040

Define CS, fit into CC

1. CUSTOMER SEGMENT(S) CS

- Hospitals
- Clinics
- WHO
- Any medical related agencies those prepare medicines or any kind of solutions inferring over the data of diseases.

6. CUSTOMER CONSTRAINTS CC

The unawareness over the AI/ML technologies, collaborative dashboards, network connection, lack of data.

5. AVAILABLE SOLUTIONS AS

The customers can prefer over a manual data visualization and prediction, which is very tedious job and requires the knowledge over the technologies of AI/ML.

Hard mathematical formulae were created and the results were being calculated manually.

Explore AS, differentiate

Focus on J&P, tap into BE, understand RC

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

Quality of Data:

The quality of data should be accurate and reliable. Obviously, the outcome will solely depend on the data we put into the prediction. If the data is skewed, then the prediction which is dependent on it, will be skewed as well.

9. PROBLEM ROOT CAUSE RC

- Difficulty of predicting a heart disease.
- Will not have a proper idea of relation between similar heart diseases.
- There is a chance of identifying every heart diseases as same.
- Reason of increase in heart disease will not be rootly identified.

7. BEHAVIOUR BE

- Generation of legitimate and reliable datasets.
- Customers need to collect more number of datasets in order to obtain more accurate result.
- Must obtain knowledge of difference between datasets that is used for comparison.

Focus on J&P, tap into BE, understand RC

Identify strong TR & EM

3. TRIGGERS TR

- Insufficient ways of handling huge amounts of datasets and inferring the root cause of the heart disease cannot be found out.
- Similarity of heart disease has not been identifiable.

10. YOUR SOLUTION SL

With the notable technology of AI/Ml we are able to visualize and predict heart diseases and related diseases, by the ultimate power Cognos Analytics Tool we will be able to properly create a dashboard for the customers to work with and visualize and analyze the heart disease on their work with limited knowledge.

8. CHANNELS of BEHAVIOR CH

8.1 ONLINE

Visualizing the datasets.
Exploration of data.

8.2 OFFLINE

Cleansing of datasets.
Collection and noting the datasets.

Identify strong TR & EM

4. EMOTIONS: BEFORE / AFTER EM **Before** -> It creates a huge ambiguity in knowing the proper or accurate reasons for a heart disease. **After** -> There is a large chance understanding of the heart disease and root cause of it. which makes a better solution and finding a preventive way over it.			

CHAPTER 04

REQUIREMENT ANALYSIS

FUNCTIONALREQUIREMENT

Followingarethefunctionalrequirements of the proposed solution.

User Registration :Enables user to make registration for the Applicationthroughthegmailaccount

User Confirmation:Once after the registration the user will getconformationviaEmail

Visualizing Data :User can visualize the trends on the heart through the interactive Dashboard created using IBM Cog nos Analytics

Generating Report:User can view the his or her report easily and can makedecisionsaccordingly

NON-FUNCTIONALREQUIREMENT:

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

Usability:The application will have a simple and user-friendly graphical interface. Users will be able to understand and use allthefeaturesoftheapplicationeasily.

Security:For security of the application the technique known as database replication should be used Incase of crash, the system should be able kept safe in case of cash the system should able to backupandrecoverthedata

Reliability: The application has to be consistent at every scenario and has to work without failure in any environment

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

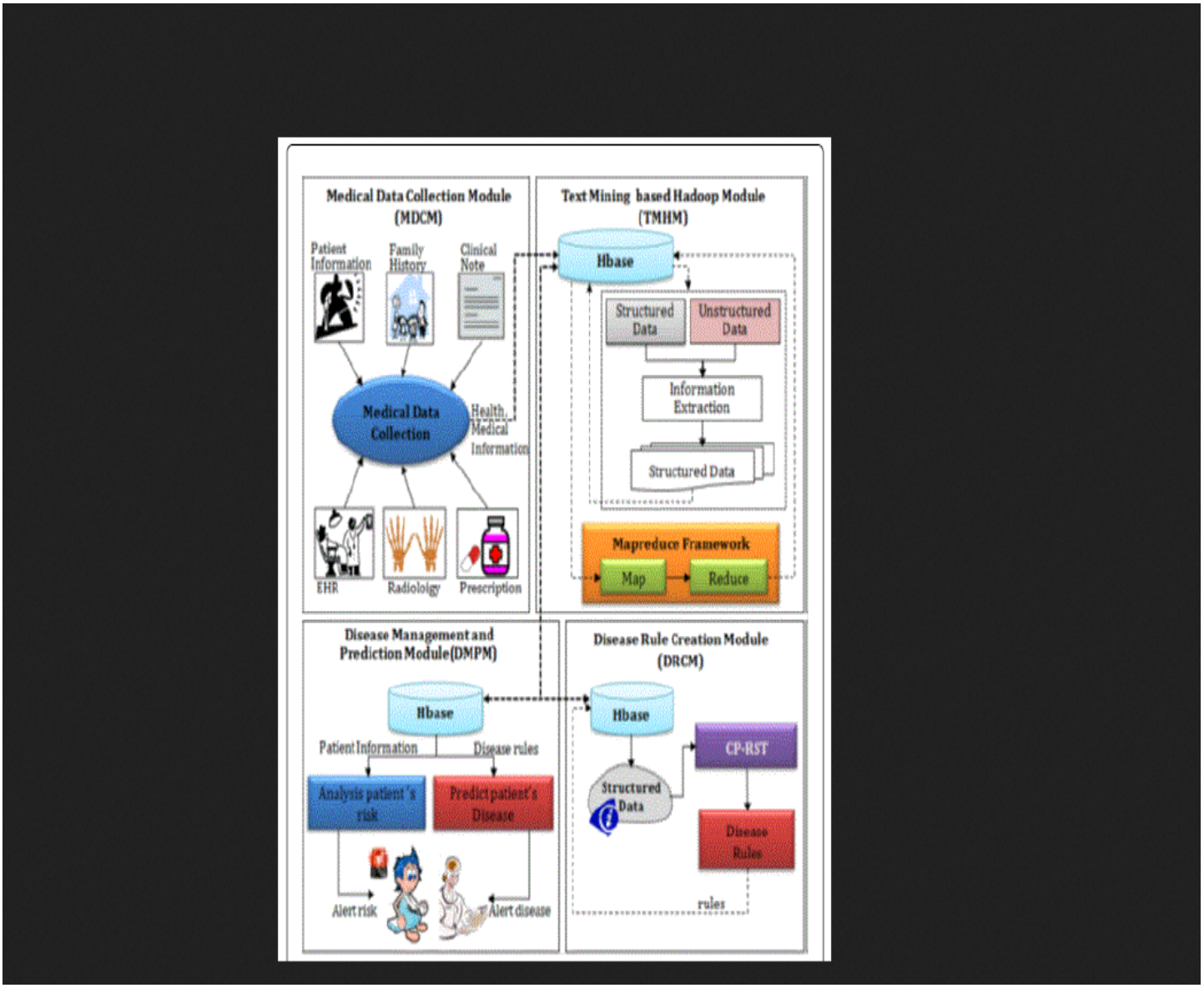
Availability: The application has to be available 24 x 7 for userswithoutanyinterruption

Scalability : The application can withstand the increase in the no. of users and has to be able to develop Higher versions

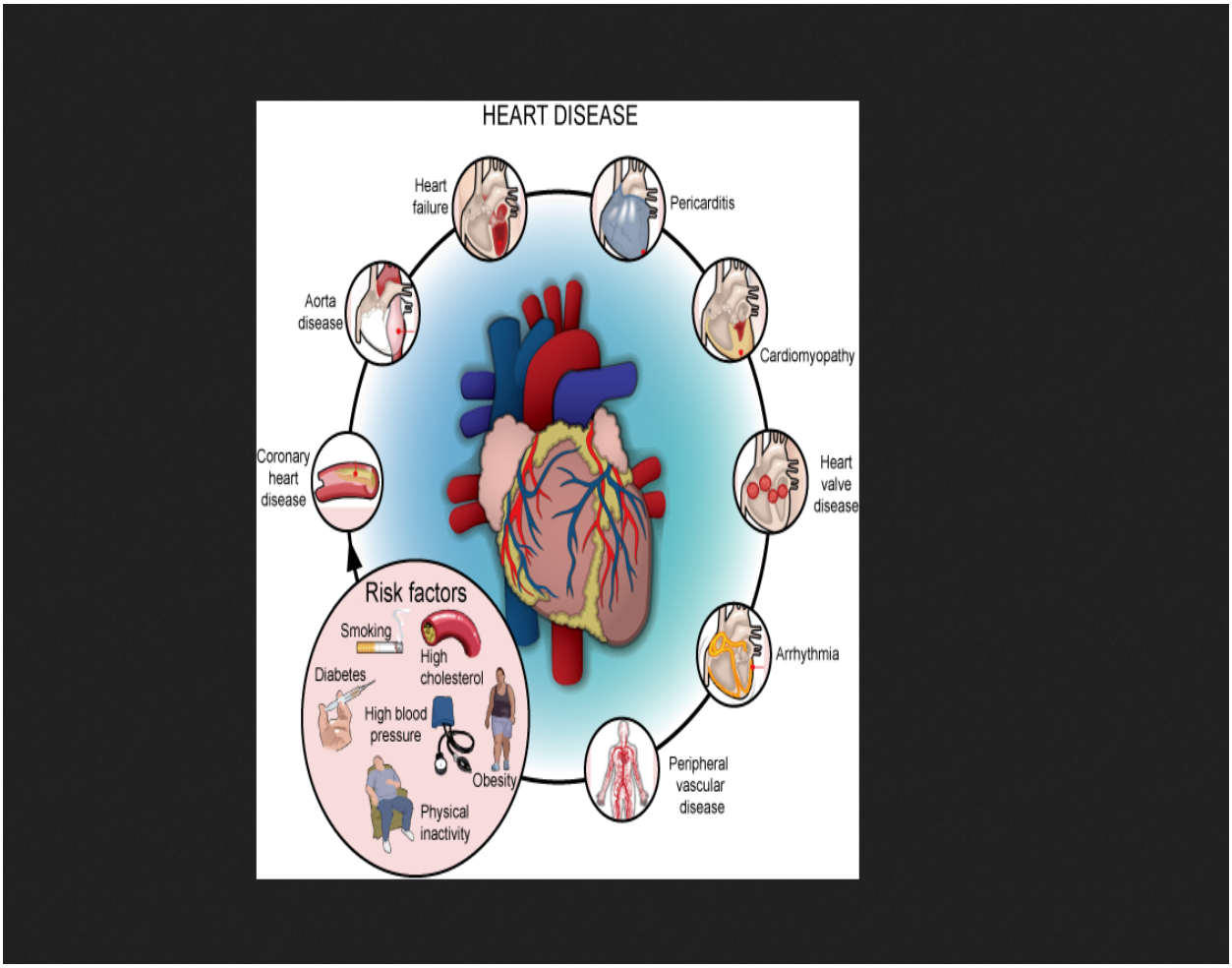
CHAPTER-05

PROJECT DESIGN

DATA FLOW DIAGRAM



SOLUTION AND TECHNICAL ARCHITECTURE:



USER STORIES:

Every day, our heart and vascular health care team makes a difference in the lives of our patients. The stories below profile some of the patients who came to UPMC for heart and vascular care.

Note: These patients' treatment and results may not be representative of all similar cases.

CHAPTER-06

PROJECT PLANNING & SCHEDULING

Sprint Planning & Estimation:

Use the below template to create product backlog and sprint schedule

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.	2	High	1
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	1	High	2
Sprint-2		USN-3	As a user, I can register for the application through Facebook	2	Low	4
Sprint-1		USN-4	As a user, I can register for the application through Gmail	2	Medium	3
Sprint-1	Login	USN-5	As a user, I can log into the application by entering email & password	1	High	2
Sprint-2	Dashboard	USN-6	Profile - view & update your profile	2	High	4
Sprint-1		USN-7	Change Password - user can change the password	1	High	2

Sprint-1		USN-8	Home - Analyze your Heart	2	High	4
Sprint-3		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)	2	High	4
		USN-10	View Doctors - view doctor detail by searching by names or filter by specialty	1	Medium	4
Sprint-3	System Requirement	USN-11	I.Hardware Requirement i.LaptoporPC	2	High	2

			I5processor system or higher			
sprint-4	system	USN-12	processor system low or high	3	High	2

Project Tracker, Velocity

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	31oct 2022	05 Nov 2022	19	05 Nov 2022
Sprint-3	20	6 Days	07Nov 2022	12Nov 2022	20	12Nov 2022
Sprint-14	20	6 Days	14 Nov 2022	19 Nov 2022	18	19 Nov 2022

Velocity:

Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let’s calculate the team’s average velocity (AV) per iteration unit (story points per day)

Sprint Delivery Schedule:

Sprint 1:

```
<htmllang="en">

<head>
```

<title>Heart Disease Prediction</title>

</head>

<bodybackground="heart.jpg">

<h1style="font-size: 3rem; color: #FFFFFF; text-align:right ">Heart Diseases Prediction</h1>

</body>

</html>

<!DOCTYPE html>

<htmllang="en">

<head>

<metacharset="UTF-8">

<metahttp-equiv="X-UA-Compatible" content="IE=edge">

<metaname="viewport" content="width=device-width, initial-scale=1.0">

<title>LOGIN PASSWORD VALIDATION | PRAROZ TUTORIAL</title>

<linkrel="stylesheet" href="style.css">

<scriptsrc="valid.js"></script>

```
</head>

<body>

<divclass="form">

<h1>LOGIN HERE</h1>

<p>Username :</p>

<inputtype="text" name="" placeholder="Name Here">

<p>Password :</p>

<inputtype="password" name="" placeholder="Password Here" id="pass">

<inputtype="checkbox" onclick="myfunction()">

<inputtype="submit" name="" value="LOGIN" onclick="validate()">

</div>



<div>

<pid="length"></p>

</div>



</body>
</html>

{
```

```
margin:0;

padding:0;

font-family: sans-serif;

}


body{

background:linear-gradient(rgba(0,0,0,0.4)50%,rgba(0,0,0,0.4)50%),url(ML.jpg);

background-position: center;

background-size: cover;

height:100vh;

}


.form{

width:250px;

height:330px;

color:#fff;

background:linear-gradient(to top,rgba(0,0,0,0.8)50%,rgba(0,0,0,0.8)50%);
```

```
position: absolute;

top:50%;

left:50%;

transform:translate(-50%,-50%);

padding:40px25px;

border-radius:10px;

}


.formh1{

width:220px;

text-align: center;

padding-left:11px;

font-size:35px;

color:#66ff00;

margin-bottom:20px;

}


.formp{
```

```
padding-bottom:-15px;

}


.forminput{

width:100%;

height:35px;

padding-top:5px;

margin-bottom:30px;

background: transparent;

border-bottom:1px solid #fff;

border-top: none;

border-left: none;

border-right: none;

color:#fff;

outline: none;

font-size:15px;

letter-spacing:1px;

}
```

```
.forminput[type="submit"]

{

width:60%;

margin-left:50px;

border: none;

height:40px;

color:#000;

background:#fff;

font-size:16px;

font-weight: bold;

border-radius:15px;

}


.forminput[type="submit"]:hover{

cursor: pointer;

background:#66ff00;

color:#fff;
```

```
font-weight: bold;
}

functionmyfunction(){

varx=document.getElementById("pass");



if(x.type==="password"){

x.type="text";

}

else{

x.type="password";

}

}



function validate(){

var password=document.getElementById("pass");

var length=document.getElementById("length");



if(password.value.length>=8){

alert("Login Succesfull");
```

```
window.location.replace("heart1.html");

return false;

}

else{

alert("Login Failed");

}

}




functionalpage(){

window.location.replace("Landingpage.html")
}

functionmyfunction(){

varx=document.getElementById("pass");



if(x.type==="password"){

x.type="text";
```

```
}
else{
x.type="password";
}
}

functionvalidate(){
varpassword=document.getElementById("pass");
varlength=document.getElementById("length");

if(password.value.length>=8){
alert("Login Succesfull");
window.location.replace("heart1.html");
return false;
}
else{
alert("Login Failed");
}
```

functionlpage(){

window.location.replace("Landingpage.html")
}

Sprint 2:

CSS

```
<!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>LOGIN PASSWORD VALIDATION | PRAROZ TUTORIAL</title>

                            <link rel="stylesheet" href="style.css">
                            <script src="valid.js"></script>

                        </head>
                        <body>
                            <div class="form">
                                <h1>LOGIN HERE</h1>
                                <p>Username :</p>
                                <input type="text" name="" placeholder="Name Here">
                                <p>Password :</p>
                                <input type="password" name="" placeholder="Password Here" id="p
                                <input type="checkbox" onclick="myfunction()">
                                <input type="submit" name="" value="LOGIN" onclick="validate()">
```

```
                    </div>

                    <div>
                        <p id="length"></p>
                    </div>

                </body>
                </html>
```

74 Project Development Phase/Sprint-2/style.css

```
@@ -0,0 +1,74 @@
*{
    margin: 0;
    padding: 0;
    font-family: sans-serif;
}

body{
    background: linear-gradient(rgba(0,0,0,0.4)50%,rgba(0,0,0,0.4)50%), url(ML.jpg)
    background-position: center;
    background-size: cover;
    height: 100vh;
}

.form{
    width: 250px;
    height: 330px;
    color: #fff;
    background: linear-gradient(to top, rgba(0,0,0,0.8)50%,rgba(0,0,0,0.8)50%);
    position: absolute;
    top: 50%;
    left: 50%;
    transform: translate(-50%,-50%);
    padding: 40px 25px;
    border-radius: 10px;
}

.form h1{
    width: 220px;
    text-align: center;
```

```
    padding-left: 11px;
    font-size: 35px;
    color: #66ff00;
    margin-bottom: 20px;
}

.form p{
    padding-bottom: -15px;
}

.form input{
    width: 100%;
    height: 35px;
    padding-top: 5px;
    margin-bottom: 30px;
    background: transparent;
    border-bottom: 1px solid #fff;
    border-top: none;
    border-left: none;
    border-right: none;
    color: #fff;
    outline: none;
    font-size: 15px;
    letter-spacing: 1px;
}

.form input[type="submit"]
{
    width: 60%;
    margin-left: 50px;
    border: none;
    height: 40px;
    color: #000;
    background: #fff;
    font-size: 16px;
    font-weight: bold;
    border-radius: 15px;
}

.form input[type="submit"]:hover{
    cursor: pointer;
```

```
        background: #66ff00;
        color: #fff;
        font-weight: bold;
    }
```

30 Project Development Phase/Sprint-2/valid.js

```
    @@ -0,0 +1,30 @@
    function myfunction(){
        var x =document.getElementById("pass");

        if(x.type === "password"){
            x.type = "text";
        }
        else{
            x.type = "password";
        }
    }

    function validate(){
        var password = document.getElementById("pass");
        var length = document.getElementById("length");

        if(password.value.length >= 8){
            alert("Login Succesfull");
            window.location.replace("heart1.html");
            return false;
        }
        else{
            alert("Login Failed");
        }
    }



    function lpage(){
        window.location.replace("Landingpage.html")
    }
```

Heart 1.html

```
<html lang="en">
                              <head>
                                <title>Heart Disease Prediction</title>
                              </head>
                              <body background="heart.jpg">
                              <h1 style="font-size: 3rem; color: #FFFFFF; text-align:right ">Heart
                              </body>
                             </html>

{
                    margin: 0;
                    padding: 0;
                    font-family: sans-serif;
                }

                body{
                    background: linear-gradient(rgba(0,0,0,0.4)50%,rgba(0,0,0,0.4)50%), url(ML.jpg)
                    background-position: center;
                    background-size: cover;
                    height: 100vh;
                }

                .form{
                    width: 250px;
                    height: 330px;
                    color: #fff;
                    background: linear-gradient(to top, rgba(0,0,0,0.8)50%,rgba(0,0,0,0.8)50%);
                    position: absolute;
                    top: 50%;
                    left: 50%;
                    transform: translate(-50%,-50%);
                    padding: 40px 25px;
                    border-radius: 10px;
                }

                .form h1{
                    width: 220px;
                    text-align: center;
                    padding-left: 11px;
                    font-size: 35px;
                    color: #66ff00;
                    margin-bottom: 20px;
```

```
}

.form p{
    padding-bottom: -15px;
}

.form input{
    width: 100%;
    height: 35px;
    padding-top: 5px;
    margin-bottom: 30px;
    background: transparent;
    border-bottom: 1px solid #fff;
    border-top: none;
    border-left: none;
    border-right: none;
    color: #fff;
    outline: none;
    font-size: 15px;
    letter-spacing: 1px;
}

.form input[type="submit"]
{
    width: 60%;
    margin-left: 50px;
    border: none;
    height: 40px;
    color: #000;
    background: #fff;
    font-size: 16px;
    font-weight: bold;
    border-radius: 15px;
}

.form input[type="submit"]:hover{
    cursor: pointer;
    background: #66ff00;
    color: #fff;
    font-weight: bold;
}
```

```
function
myfunction(){
                                    var x =document.getElementById("pass");

                                    if(x.type === "password"){
                                        x.type = "text";
                                    }
                                    else{
                                        x.type = "password";
                                    }
                                }

                                function validate(){
                                    var password = document.getElementById("pass");
                                    var length = document.getElementById("length");

                                    if(password.value.length >= 8){
                                        alert("Login Succesfull");
                                        window.location.replace("heart1.html");
                                        return false;
                                    }
                                    else{
                                        alert("Login Failed");
                                    }
                                }



                                function lpage(){
                                    window.location.replace("Landingpage.html")
                                }
```

ıogin .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>LOGIN PASSWORD VALIDATION | PRAROZ TUTORIAL</title>
```

```
    <link rel="stylesheet" href="style.css">
    <script src="valid.js"></script>

</head>
<body>
    <div class="form">
        <h1>LOGIN HERE</h1>
        <p>Username :</p>
        <input type="text" name="" placeholder="Name Here">
        <p>Password :</p>
        <input type="password" name="" placeholder="Password Here"
id="pass">
        <input type="checkbox" onclick="myfunction()">
        <input type="submit" name="" value="LOGIN"
onclick="validate()">
    </div>

    <div>
        <p id="length"></p>
    </div>

</body>
</html>
```

STYLE CSS

```
*{
    margin: 0;
    padding: 0;
    font-family: sans-serif;
}

body{
    background: linear-gradient(rgba(0,0,0,0.4)50%,rgba(0,0,0,0.4)50%),
url(ML.jpg);
    background-position: center;
    background-size: cover;
    height: 100vh;
}

.form{
    width: 250px;
```

```
    height: 330px;
    color: #fff;
    background: linear-gradient(to top,
rgba(0,0,0,0.8)50%,rgba(0,0,0,0.8)50%);
    position: absolute;
    top: 50%;
    left: 50%;
    transform: translate(-50%,-50%);
    padding: 40px 25px;
    border-radius: 10px;
}

.form h1{
    width: 220px;
    text-align: center;
    padding-left: 11px;
    font-size: 35px;
    color: #66ff00;
    margin-bottom: 20px;
}

.form p{
    padding-bottom: -15px;
}

.form input{
    width: 100%;
    height: 35px;
    padding-top: 5px;
    margin-bottom: 30px;
    background: transparent;
    border-bottom: 1px solid #fff;
    border-top: none;
    border-left: none;
    border-right: none;
    color: #fff;
    outline: none;
    font-size: 15px;
    letter-spacing: 1px;
}

.form input[type="submit"]
```

```
{
    width: 60%;
    margin-left: 50px;
    border: none;
    height: 40px;
    color: #000;
    background: #fff;
    font-size: 16px;
    font-weight: bold;
    border-radius: 15px;
}

.form input[type="submit"]:hover{
    cursor: pointer;
    background: #66ff00;
    color: #fff;
    font-weight: bold;
}
```

VALID JS

```
function
myfunction(){
                        var x =document.getElementById("pass");

                        if(x.type === "password"){
                            x.type = "text";
                        }
                        else{
                            x.type = "password";
                        }
                    }

                    function validate(){
                        var password =
                    document.getElementById("pass");
                        var length =
                    document.getElementById("length");

                        if(password.value.length >= 8){
                            alert("Login Succesfull");
```

```
window.location.replace("heart1.html");
        return false;
    }
    else{
        alert("Login Failed");
    }
}



function lpage(){

window.location.replace("Landingpage.html")
}
```

SPRINT 3:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import GridSearchCV
from sklearn import metrics
import io
%matplotlib inline

import matplotlib.pyplot as plt
from pandas.plotting import scatter_matrix

from google.colab import files


uploaded = files.upload()
```

Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

```
Saving heart.csv to heart (3).csv

dataset = pd.read_csv(io.BytesIO(uploaded['heart.csv']))
dataset.head()

dataset.info()
```

```
dataset.shape
```

```
(303, 14)
```

```
pd.set_option("display.float", "{:.2f}".format)
dataset.describe()
```

```
dataset.isna().sum()
```

```
correlation=dataset.corr()
correlation
```

```
sns.countplot(dataset['target'])
plt.title("COUNT vs RESULT")
plt.xlabel("RESULT")
plt.ylabel("COUNT")
plt.show()
```

```
cor=dataset.corr()
top_corr_feature=cor.index
plt.figure(figsize=(20,20))
sns.heatmap(dataset[top_corr_feature].corr(),annot=True,cmap="RdYlGn")
plt.title("HEATMAP")
plt.show()
```

```
sns.catplot(x = 'sex', kind = 'count', hue = 'target', data= dataset, palette =
'ch:.25')
```

```
sns.catplot(x = 'cp', kind = 'count', hue = 'target', data= dataset, palette =
'ch:.25')
```

```
sns.catplot(x = 'fbs', kind = 'count', hue = 'target', data= dataset, palette =
'ch:.25')
```

```
sns.catplot(x = 'restecg', kind = 'count', hue = 'target', data= dataset,
palette = 'ch:.25')
```

```
sns.catplot(x = 'exang', kind = 'count', hue = 'target', data=
dataset, palette = 'ch:.25')
```

```
sns.catplot(x = 'slope', kind = 'count', hue = 'target', data= dataset, palette
= 'ch:.25')
```

```
sns.catplot(x = 'ca', kind = 'count', hue = 'target', data= dataset, palette =
'ch:.25')
```

```
sns.catplot(x = 'thal', kind = 'count', hue = 'target', data= dataset, palette
= 'ch:.25')

dataset[['age', 'trestbps', 'chol', 'thalach', 'oldpeak']].describe()

sns.displot(x='age', multiple = 'stack', hue = 'target', data=
dataset, palette = 'ch:.25')
sns.displot(x='trestbps', multiple = 'stack', hue = 'target', data=
dataset, palette = 'ch:.25')

sns.displot(x='chol', multiple = 'stack', hue = 'target', data= dataset,
palette = 'ch:.25')

sns.displot(x='thalach', multiple = 'stack', hue = 'target', data= dataset,
palette = 'ch:.25')

sns.displot(x='oldpeak', multiple = 'stack', hue = 'target', data= dataset,
palette = 'ch:.25')

x = dataset.drop('target', axis =1)
y = dataset['target']
x_train, x_test, y_train, y_test = train_test_split(x,y,test_size = 0.2,
random_state = 42)
sc = StandardScaler().fit(x_train)
x_train = sc.transform(x_train)
x_test = sc.transform(x_test)

knn = KNeighborsClassifier()
parameters = {'n_neighbors': [3,5,7,9,11], 'weights': ['uniform',
'distance']}
grid = GridSearchCV(knn, parameters, cv = 4, scoring = 'accuracy')
grid.fit(x_train,y_train)

GridSearchCV(cv=4, estimator=KNeighborsClassifier(),
             param_grid={'n_neighbors': [3, 5, 7, 9, 11],
                         'weights': ['uniform', 'distance']},
             scoring='accuracy')

print(grid.best_params_)
{'n_neighbors': 5, 'weights': 'uniform'}

model = grid.best_estimator_

model.score (x_test,y_test)
```

```
0.9016393442622951

predictions = model.predict(x_test)
cm = metrics.confusion_matrix(y_test, predictions)
cm = pd.DataFrame(cm)
sns.heatmap(cm, annot=True)
plt.show()
TP = 28
FP = 2
TN = 27
FN = 4
sensitivity = TP / (TP + FN) *100
specificity = TN / (TN + FP) * 100
ppv = TP / (TP + FP) * 100
npv = TN / (TN + FN) * 100
print('Sensitivity:', sensitivity,'% ','Specificity:', specificity,'%
','positive predictive value:',ppv,'% ','negative predictive
value:',npv,'%' )
Sensitivity: 87.5 %  Specificity: 93.10344827586206 %  positive predictive
value: 93.33333333333333 %  negative predictive value: 87.09677419354838 %

probs = model.predict_proba(x_test)[:, 1]
auc = metrics.roc_auc_score(y_test, probs)
print(auc)
0.9154094827586208

fpr, tpr, _ = metrics.roc_curve(y_test,probs)
plt.figure()
plt.grid()
plt.plot(fpr, tpr)
plt.plot([0, 1], [0, 1])
plt.show()
```

SPRINT 04:

AppImb:

```
import
pickle
              import time

              import joblib
```

```
import matplotlib
import matplotlib.pyplot as plt
import numpy as np
import pandas
import requests
from flask import Flask, render_template, request


app = Flask(__name__)
model=pickle.load(open('heart.pkl','rb'))
scale=pickle.load(open('scale.pkl','rb'))
@app.route('/')
def home():
    return render_template('index.html')
@app.route('/predict',methods=["POST","GET"])
def predict():
    if request.method == 'POST':

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

        data =
np.array([[age,gender,cp,trestbps,chol,fbs,restecg,thalach,exang,oldpeak,slope,ca,th
        prediction = model.predict(data)
    if prediction == "Yes":
        return render_template("chance.html")
    else:
        return render_template("nochance.html")
if __name__ == "__main__":
  app.run(debug=True)
```

SPRINT DELIVERY SCHEDULE:

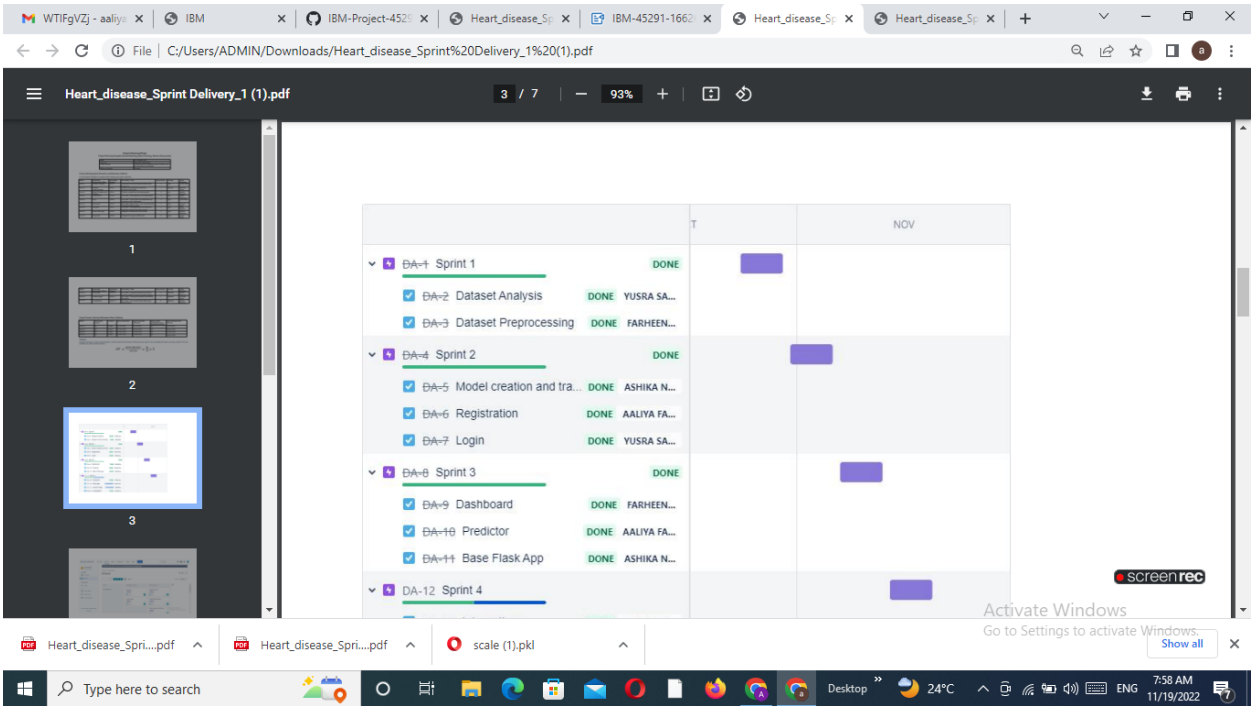
SPRINT -01 WE HAVE SUBMITTED ON 13-11-2022

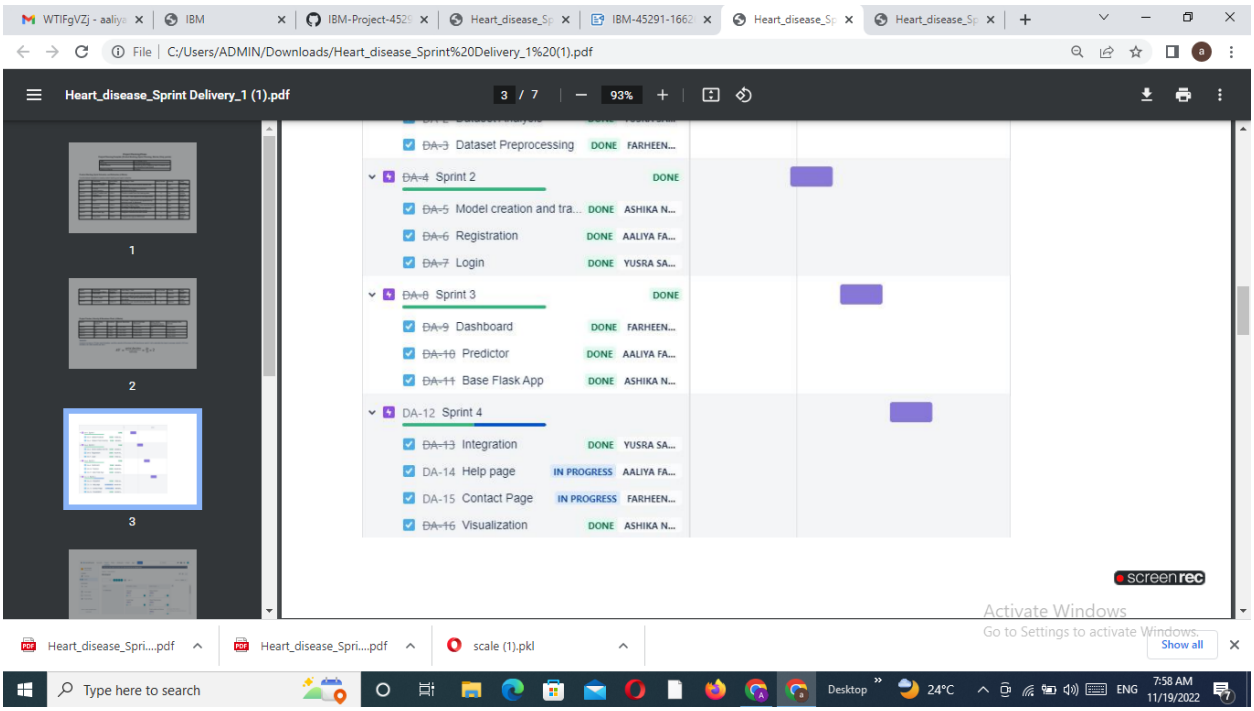
SPRINT-02 WE HAVE SUBMITTED ON 13-11-2022

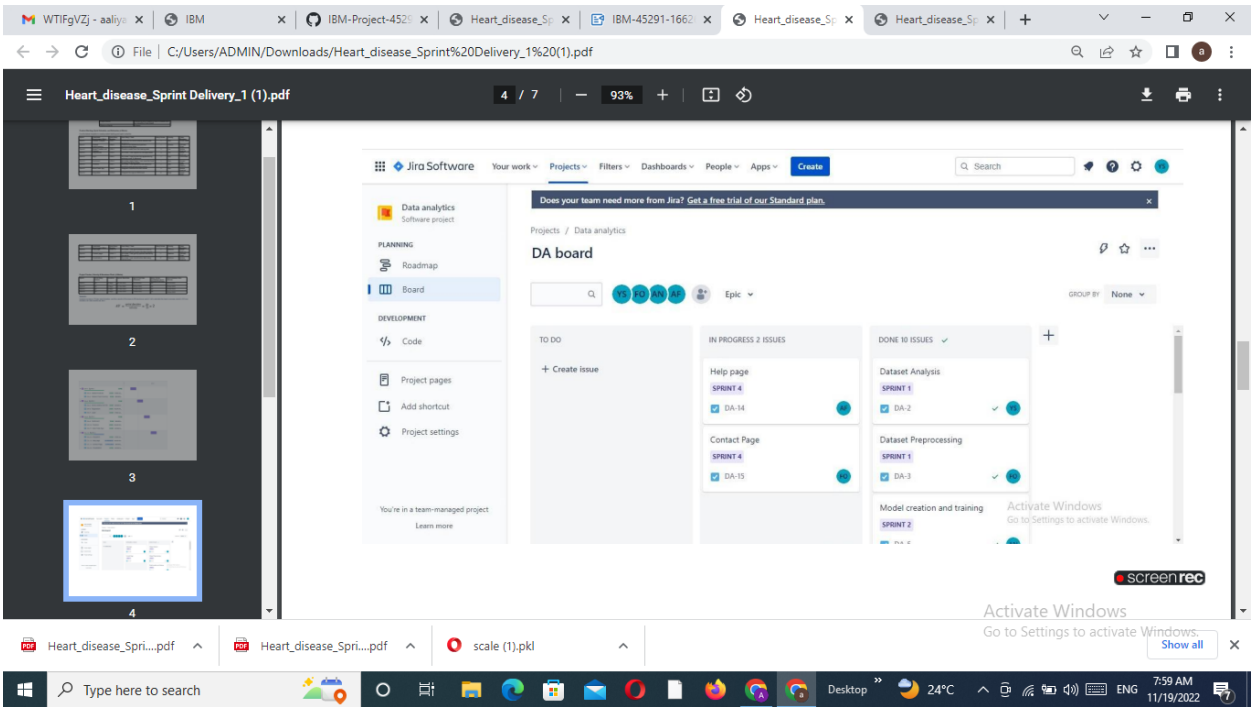
SPRINT -03 WE HAVE SUBMITTED ON 17-11-2022

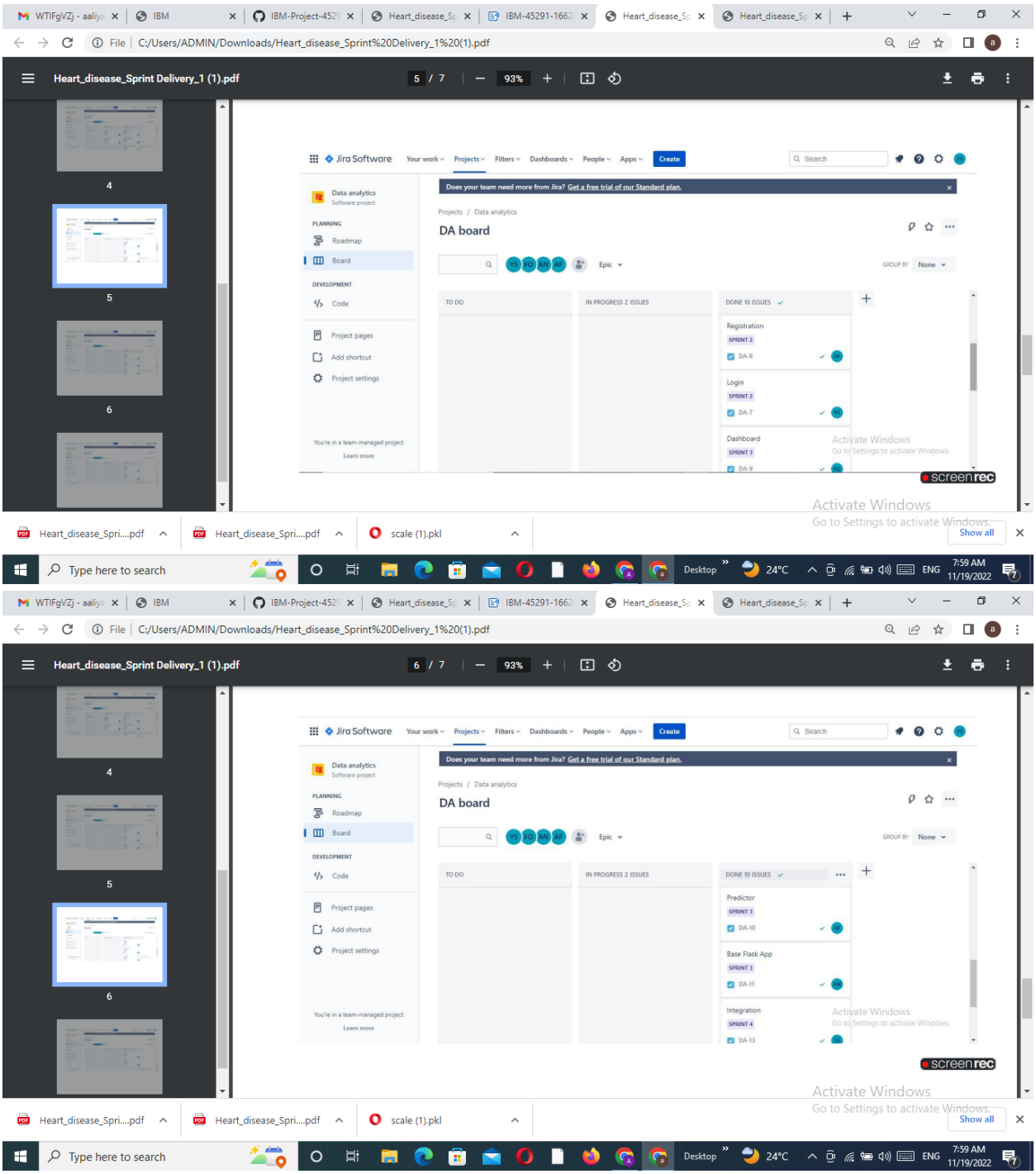
SPRINT-04 WE HAVE SUBMITTED ON 18-11-2022

REPORTS FROM JIRA:

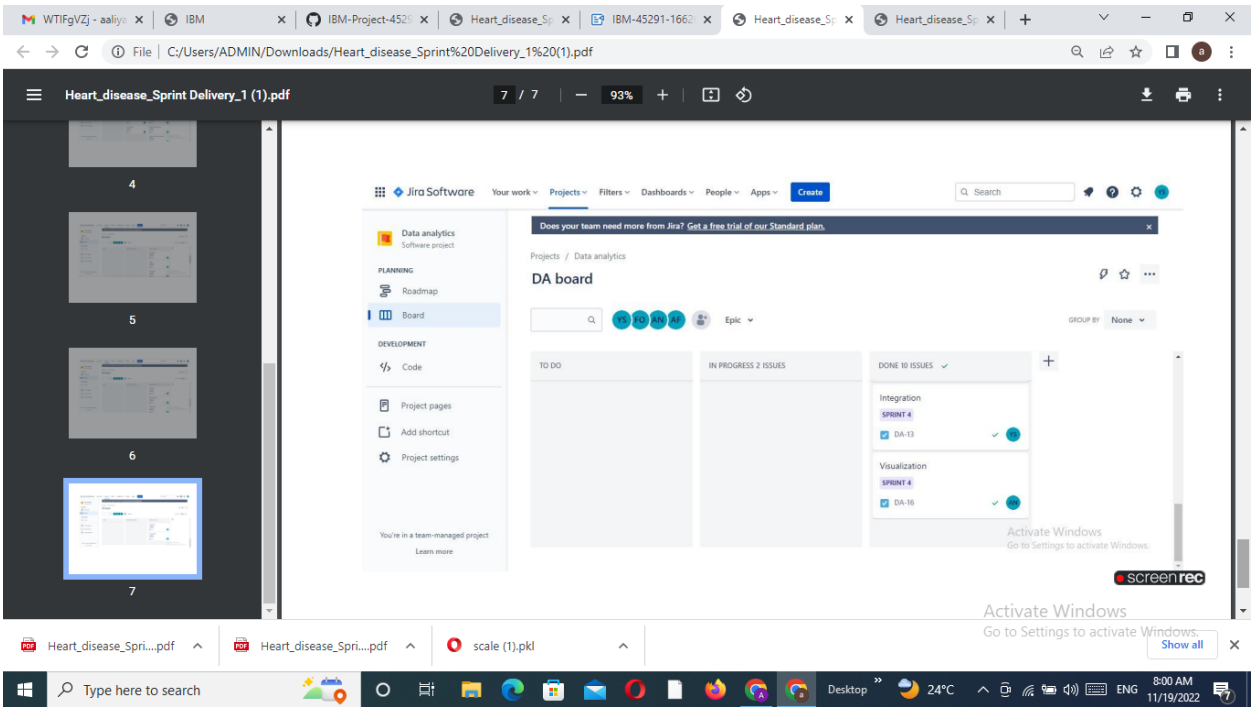



Heart_disease_Sprint Delivery_1 (1).pdf
3 / 7
93%
DA-3 Dataset Preprocessing DONE FARHEEN...
DA-4 Sprint 2 DONE
DA-5 Model creation and tra... DONE ASHIKA N...
DA-6 Registration DONE AALIYA FA...
DA-7 Login DONE YUSRA SA...
DA-8 Sprint 3 DONE
DA-9 Dashboard DONE FARHEEN...
DA-10 Predictor DONE AALIYA FA...
DA-11 Base Flask App DONE ASHIKA N...
DA-12 Sprint 4
DA-13 Integration DONE YUSRA SA...
DA-14 Help page IN PROGRESS AALIYA FA...
DA-15 Contact Page IN PROGRESS FARHEEN...
DA-16 Visualization DONE ASHIKA N...
screenrec
Activate Windows
Go to Settings to activate Windows.
Heart_disease_Spri....pdf
Heart_disease_Spri....pdf
scale (1).pkl
Show all
Type here to search
Desktop
24°C
ENG
7:58 AM
11/19/2022


Heart_disease_Sprint Delivery_1 (1).pdf
4 / 7
93%
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Data analytics
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / Data analytics
DA board
Epic
GROUP BY None
TO DO
+ Create issue
IN PROGRESS 2 ISSUES
Help page
SPRINT 4
DA-14
Contact Page
SPRINT 4
DA-15
DONE 10 ISSUES
Dataset Analysis
SPRINT 1
DA-2
Dataset Preprocessing
SPRINT 1
DA-3
Model creation and training
SPRINT 2
screenrec
Activate Windows
Go to Settings to activate Windows.
Heart_disease_Spri....pdf
Heart_disease_Spri....pdf
scale (1).pkl
Show all
Type here to search
Desktop
24°C
ENG
7:59 AM
11/19/2022


Heart_disease_Sprint Delivery_1 (1).pdf
5 / 7
93%
Jira Software
DA board
Activate Windows
Go to Settings to activate Windows.
screenrec
Heart_disease_Spri....pdf
Heart_disease_Spri....pdf
scale (1).pkl
Show all
Type here to search
Desktop
24°C
ENG
7:59 AM
11/19/2022

Heart_disease_Sprint Delivery_1 (1).pdf
6 / 7
93%
Jira Software
DA board
Activate Windows
Go to Settings to activate Windows.
screenrec
Heart_disease_Spri....pdf
Heart_disease_Spri....pdf
scale (1).pkl
Show all
Type here to search
Desktop
24°C
ENG
7:59 AM
11/19/2022



CHAPTER-07

CODING & SOLUTIONING

FEATURE 1:

Be sure to eat plenty of fresh fruits and vegetables and fewer processed foods. Eating lots of foods high in saturated fat and trans fat may contribute to heart disease. Eating foods high in fiber and low in saturated fats, trans fat, and cholesterol can help prevent high cholesterol.

A low-fat, high-fibre diet is recommended, which should include plenty of fresh fruit and vegetables (5 portions a day) and whole grains.

You should limit the amount of salt you eat to no more than 6g (0.2oz) a day as too much salt will increase your blood pressure. 6g of salt is about 1 teaspoonful.

There are 2 types of fat: saturated and unsaturated. You should avoid food containing saturated fats, because these will increase the levels of bad cholesterol in your blood.

Foods high in saturated fat include:

- meat pies
- sausages and fatty cuts of meat
- butter
- ghee – a type of butter often used in Indian cooking
- lard
- cream
- hard cheese
- cakes and biscuits
- foods that contain coconut or palm oil

However, a balanced diet should still include unsaturated fats, which have been shown to increase levels of good cholesterol and help reduce any blockage in your arteries.

Foods high in unsaturated fat include:

- oily fish
- avocados
- nuts and seeds
- sunflower, rapeseed, olive and vegetable oils

You should also try to avoid too much sugar in your diet, as this can increase your chances of developing diabetes, which is proven to significantly increase your chances of developing CHD.

Read more about:

- healthy eating
- eating less saturated fat
- the facts about sugar

Be more physically active

Combining a healthy diet with regular exercise is the best way of maintaining a healthy weight. Having a healthy weight reduces your chances of developing high blood pressure.

Regular exercise will make your heart and blood circulatory system more efficient, lower your cholesterol level, and also keep your blood pressure at a healthy level.

Exercising regularly reduces your risk of having a heart attack. The heart is a muscle and, like any other muscle, benefits from exercise. A strong heart can pump more blood around your body with less effort.

Any aerobic exercise, such as walking, swimming and dancing, makes your heart work harder and keeps it healthy.

Read more about fitness and exercise.

Keep to a healthy weight

A GP or practice nurse can tell you what your ideal weight is in relation to your height and build. Alternatively, find out what your body mass index (BMI) is by using our BMI calculator.

Read more about losing weight.

Give up smoking

If you smoke, giving up will reduce your risk of developing CHD.

Smoking is a major risk factor for developing atherosclerosis (furring of the arteries). It also causes the majority of cases of coronary thrombosis in people under the age of 50.

Research has shown you're up to 3 times more likely to successfully give

up smoking if you use NHS support together with stop-smoking medicines, such as patches or gum.

Ask a doctor about this or visit NHS Smokefree.

Read more about stopping smoking.

CODING :

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import GridSearchCV
from sklearn import metrics
import io
%matplotlib inline

import matplotlib.pyplot as plt
from pandas.plotting import scatter_matrix

sns.countplot(dataset['target'])
plt.title("COUNT vs RESULT")
plt.xlabel("RESULT")
plt.ylabel("COUNT")
plt.show()



from google.colab import files

cor=dataset.corr()
top_corr_feature=cor.index
plt.figure(figsize=(20,20))
sns.heatmap(dataset[top_corr_feature].corr(),annot=True,cmap="RdYlGn")
```

```
plt.title("HEATMAP")
plt.show()
```

```
uploaded = files.upload()
```

Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

```
Saving heart.csv to heart (3).csv
```

In [7]:

```
dataset = pd.read_csv(io.BytesIO(uploaded['heart.csv']))
dataset.head()
```

FEATURE 02:

Reduce your alcohol consumption

If you drink, do not exceed the maximum recommended limits.

- men and women are advised not to regularly drink more than 14 units a week
- spread your drinking over 3 days or more if you drink as much as 14 units a week

Always avoid binge drinking, as this increases the risk of a heart attack.

Read more about drinking and alcohol.

Keep your blood pressure under control

You can keep your blood pressure under control by eating a healthy diet low in saturated fat, exercising regularly and, if needed, taking medicine to

lower your blood pressure.

Your target blood pressure should be below 140/90mmHg. If you have high blood pressure, ask a GP to check your blood pressure regularly.

Read more about high blood pressure.

Keep your diabetes under control

You have a greater chance of developing CHD if you have diabetes. Being physically active and controlling your weight and blood pressure will help manage your blood sugar level.

If you have diabetes, your target blood pressure level should be below 130/80mmHg.

Read more about diabetes.

Take any prescribed medicine

If you have CHD, you may be prescribed medicine to help relieve your symptoms and stop further problems developing.

If you do not have CHD but have high cholesterol, high blood pressure or a history of family heart disease, your doctor may prescribe medicine to

prevent you developing heart-related problems.

If you're prescribed medicine, it's vital you take it and follow the correct dosage. Do not stop taking your medicine without consulting a doctor first, as doing so is likely to make your symptoms worse and put your health at risk.

CODE:

```
model= RandomForestClassifier(n_estimitor=20)

model.fit(x_trainer_scaler,y_trainer)

y_pred=model.predict(x_tester_scaler)

p=model.score(y_test_scaler,y_tester)

print(p)
```

CHAPTER-08

TESTING

Test Cases:

Machine Learning is used across many spheres around the world. The healthcare industry is no exception. Machine Learning can play an essential role in predicting presence/absence of Locomotor disorders, Heart diseases and more. Such information, if predicted well in advance, can provide important insights to doctors who can then adapt their diagnosis and treatment per patient basis

I saved it to my working directory with the name `dataset.csv`. Next, I used `read_csv()` to read the dataset and save it to the `dataset` variable.

Before any analysis, I just wanted to take a look at the data. So, I used the `info()` method.

User Acceptance Testing:

We will present number of Machine Learning Classifications algorithm for our analysis [36], in which they will be used for model performance comparison, classification and prediction of HD dataset.

K-nearest neighbour (K-NN)

The idea of the Nearest Neighbor (NN) classifier is to take a test data point and comparing it with all training data points and to predict the label (class) of the test data point based on the closest training class using the L_1 distance given by:

Where I_1, I_2 are the vectors representation of points *1* and *2* respectively, and

d_1 denote the distance and $\sum$ is taken over all points.

NaiveBayes

Given the Bayes theorem:

$$P(A|B)=P(B|A)P(A)P(B) \tag{2}$$

For a given elements A and B and their probability of occurrence $P(X)$ is calculated, where $P(A)$ is the probability of occurrence of element A, $P(B)$ is the probability of occurrence of element A and $P(A|\ B)$ is the conditional probability of element A given element B occur, and such theorem will be used to perform the classification. So for independent features, the mentioned theorem would perform a direct multiplication of the probability of each feature happening.

Decision tree J48

A decision tree model is a model that run number of comparison questions to divide the dataset into different smaller sets based on a given questions (Boolean for instance), and it keeps repeating the task with different set of questions for different level of the available subsets until it covers all available attributes in the dataset. We can have different type of decision tree classifiers based on the nature of the provided questions and their decision rules and based on the nature of the data set. Decision tree J48 is a special case based on the C4.5 algorithm, and it is used for a unified variable associated with the dataset.

Decision tree JRip

JRip (RIPPER) is a rule learner classifier found in decision tree algorithm, and uses a repeated incremental pruning for error reduction, and uses four

distinctive phases, building, growing, pruning and optimization [37].

Stochastic gradient descent (SGD)

Gradient descent is an algorithm that optimizes many loss functions, such as Support Vector Machine (SVM), and Logistic Regression models, and is usually used to optimize the linear function, and the stochastic concept is introduced here based on the roots finding nature of the optimization task. In Stochastic Gradient Descent, for each iteration, samples are selected randomly using a term "batch" for number of samples, instead of the whole data set, and these batches are used to calculate the gradient for each iteration.

Support vector machine (SVM)

Given a set of data with N attributes, Support Vector Machine (SVM) classifier is to find a suitable hyper plane in N-Dimensional space that clearly classify the dataset with a maximum margin between data points, where it segregates the two main classes hyper-plane and line to separate the available sets of points, and it is considered a supervised machine learning algorithm which can be used for classification.

Adaboost

Is a type of estimator that starts by selecting a set of the original data for fitting on a classifier, and then update the set based on the weight changes of the incorrectly classified instances, until best estimation is achieved [38], and has number of parameters to be considered, such as, base estimator, number of estimators and learning rate.

Statistical terminologies

The following statistical parameters were used in comparing the evaluation

performance of the used classifiers to classify the HD dataset:

Relative absolute error (RAE)

RAE Is the value of the relative error divided by the exact value, where the absolute error is the magnitude of the difference between exact value and approximation:

$$RAE = \left| \left| \left| V_A - V_E V_E \right| \right| \right| \quad (3)$$

Where V_A is the approximation and V_E is the exact value respectively.

Mean absolute error (MAE)

MAE is a value of the relative error divided by the number of instances, *n*, in a dataset:

$$MAE = \sum_{ni=1} |V_{Ai} - V_{Ei}| n \quad (4)$$

Kappa

Kappa statistic is the value of how close an instance is classified, where the higher Kappa value implies a better classification for a given classifier is performed.

Area under curve (ROC)

is a classification parameter to distinguish how well a classifier is [performing in term of the accuracy of identifying data point, and the ideal ROC value for perfect classification is equal to 1.

Conclusion

In this paper, a comparative analysis of different classifiers was done for the

classification of the Heart Disease dataset for positive and negative diagnosed participants. The algorithms were used K- Nearest Neighbor (K-NN), Naive Bayes, Decision tree J48, JRip, SVM, Adaboost, Stochastic Gradient Decent (SGD) and Decision Table (DT) classifiers. It was shown that using different classification algorithms for the classification of the HD dataset produced very promising results in term of the classification accuracy for the K-NN (K = 1), Decision tree J48 and JRip classifiers compared to Naïve Bayes, SGD, SVM, Decision Table and Adaboost classifiers, with accuracy of classification of 99.7073, 98.0488 and 97.2683% respectively, which outperformed other used references in this paper in term of the classification accuracy of 82.5% in [39], 86.43% in [43], 98% in [46], 81.23% in [A], 95% in [D], 93.08% in [G] and 94.62% in [H]. Also results shows Kappa statistic value of 0.9941, 0.961 and 0.9454 respectively. Sensitivity analysis for the Decision tree J48 classifier was applied to study its performance to classify HD dataset with respect to some changes in its pruning confidence factor parameter, and results shows an enhancement in the classification accuracy for the PCF = 0.30 and 0.35, with an accuracy of 98.1463% compared to the original results obtained for PCF = 0.25, and an enhancement in the Kappa statistic, MAE and RAE for the values of 0.9629, 0.0189 and 0.1268 respectively. A feature extraction method was performed using Classifier Subset Evaluator on the HD dataset to evaluate the classification performance after selecting the relevant attributes per classification algorithm. Results show enhanced performance of increasing of the classifications accuracy for K-NN ($N = 1$) and Decision Table classifiers from 99.7073, 93.6585% before applying feature selection to 100 and 93.8537% respectively, compared to 90.40% perdition accuracy in [E], were relevant attributes can be used for high accuracy classification for K-NN and Decision tree J48 classifiers to predict a HD case by only applying a combination of up to 4 attributes instead of 13 attributes of the full dataset. We can clearly see the advantages of this analysis in term of comparing different classifiers to classify the HD dataset, and the benefit of having a reliable feature selection method for HD disease prediction with using

minimal number of attributes instead of having to consider all available ones.

CHAPTER 09

RESULTS

Results

- When performing various methods of feature selection, testing it was found that backward elimination gave us the best results among others. The various methods tried were BackwardElimination with and without KFold, Recursive Feature Elimination with Cross Validation. The accuracy that was seen in them ranged around 85% with 85.5% being maximum.
- Though both methods gave similar accuracy but it was seen that in Backward Elimination we found that the number of misclassifications of True Negative was more and it was observed that the accuracy had more variance compared to RFEV. The precision of Backward Elimination and RFEV are 84%and 86% respectively. And the recalls are 0.99 and 1 respectively.
- The precision and recall also shows that the number of misclassifications is less in RFECV than in Backward

Elimination.Evaluation Metrics Backward Elimination RFECVAccuracy 83% 85%Recall 0.99 0.99Precision 0.84 0.86

CHAPTER 10

ADVANTAGES & DISADVANTAGES

Advantages

- User can search for doctor’s help at any point of time.
- User can talk about their Heart Disease and get instant diagnosis.
- Doctors get more clients online. Very useful in case of emergency.

Disvantage

The system is not fully automated, it needs data from user for full diagnosis.

CHAPTER 11

CONCLUSION

CONCLUSION

Heart diseases when aggravated spiral way beyond control. Heart diseases are complicated and take away lots of lives every year .When the early symptoms of heart diseases are ignored, the patient might end up with drastic consequences in a short span of time. Sedentary lifestyle and excessive stress in today's world have worsened the situation. If the disease is detected early then it can be kept under control. However, it is always advisable to exercise daily and discard unhealthy habits at the earliest. Tobacco consumption and unhealthy diets increase the chances of stroke and heart diseases. Eating at least 5 helpings of fruits and vegetables a day is a good practice. For heart disease patients, it is advisable to restrict the intake of salt to one teaspoon per day. It has been observed that a properly cleaned and pruned dataset provides much better accuracy than an unclean one with missing values. Selection of suitable techniques for data cleaning along with proper classification algorithms will lead to the development of prediction

systems that give enhanced accuracy. In future an intelligent system may be developed that can lead to selection of proper treatment methods for a patient diagnosed with heart disease. A lot of work has been done already in making models that can predict whether a patient is likely to develop heart disease or not. There are several treatment methods for a patient once diagnosed with a particular form of heart disease. Data mining can be of very good help in deciding the line of treatment to be followed by extracting knowledge from such suitable databases.

REFERENCES [1] https://en.wikipedia.org/wiki/Card

CHAPTER 12
FUTURE SCOPE

The objective of this project is **to check whether the patient is likely to be diagnosed with any cardiovascular heart diseases based on their medical attributes such as gender, age, chest pain, fasting sugar level, etc**. A dataset is selected from the UCI repository with patient's medical history and attributes.

CHAPTER 13

APPENDIX

SOURCE CODE:

https://github.com/IBM-EPBL/IBM-Project-45291-1660729313/tree/main/Final%20Deliverables

GITHUB AND DEMO LINK

GITHUB LINK:

https://github.com/IBM-EPBL/IBM-Project-45291-1660729313/tree/main/Final%20Deliverables

DEMO LINK:

https://drive.google.com/file/d/1cbzOy4uPTc1asTKcmC1jXAYhDNenIfoU/view

PROP PPROSED