"VISUALIZING AND PREDICTING HEART DISEASES WITH AN

INTERACTIVE DASH BOARD"

NALAIYA THIRAN IBM PROJECT REPORT Team Id : **PNT2022TMID51181**

|  |  |  |
| --- | --- | --- |
|  | *Submitted by* |  |
| **SHANTHOSH VIKASH S** |  | **953719104051** |
| **AKASH KUMAR S** |  | **953719104005** |
| **KATHIRAVAN R** |  | **953719104025** |
| **PANDI SANKAR P** |  | **953719104041** |
| **SATHEESHKUMAR P** |  | **953719104048** |

UNDER THE GUIDANCE OF

**Dr.V.VIVEK**

in partial fulﬁllment for the award of the degree

*of*

BACHELOR OF ENGINEERING

in

Department of Computer Science and Engineering

### AAA COLLEGE OF ENGINEERING AND TECHNOLOGY SIVAKASI

**CONTENT**

### INTRODUCTION

* 1. Project overview
  2. Purpose

### LITERATURE SURVEY

* 1. Existing problem
  2. Problem statement

### IDEATION & PROPOSED SOLUTION

* 1. Empathy Map Canvas
  2. Ideation & Brainstorm
  3. Proposed Solution
  4. Problem Solution Fit

**4 REQUIREMENT ANALYTICS**

* 1. Functional requirement
  2. Non-Functional requirement

**5 PROJECT DESIGN**

Data Flow Diagram

User Stories

Solution & Technical Architecture

### PROJECT PLAINING & SCHEDULING

Sprint Planning & Estimation

Sprint Delivery Schedule

### CODING & SOLUTIONING

Feature 1

Feature 2

### TESTING

Test Cases

User Acceptance Testing

### RESULTS

Performance Metrics

**10 ADVANTAGES AND DISADVANTAGES**

### CONCLUSION

1. **FUTURE SCOPE**
2. **APPENDIX**
3. INTRODUCTION

Heart Disease is even highlighted as a silent killer which leads to the death of the person without obvious symptoms.The early diagnosis of heart disease plays a vital role in making decisions on lifestyle changes in high-risk patients and in turn reduce the complications. This project aims to predict future Heart disease by analyzing data of patients which classiﬁes whether they have heart disease or not using machine-learning algorithms.

## Project overview

In this fast moving world people want to live a very luxurious life so they work like a machine in order to earn lot of money and live a comfortable life therefore in this race they forget to take care of themselves, because of this there food habits change their entire lifestyle change, in this type of lifestyle they are more tensed they have blood pressure, sugar at a very young age and they don’t give enough rest for themselves and eat what they get and they even don’t bother about the quality of the food if sick the go for their own medication as a result of all these small negligence it leads to a major threat that is the heart disease.

## Purpose

The health care industries collect huge amounts of data that contain some hidden information, which is useful for making effective decisions for providing appropriate results and making effective decisions on data, some data mining techniques are used to better the experience and conclusion that have been given.

# LITERATURE SURVEY

The main aim of this paper is to use various classiﬁcation algorithms of data science framework to somehow detect the chances of having a heart disease. Also, the main aim of this research paper is to ﬁnd out the most eﬃcient classiﬁcation algorithm that can help us to detect heart diseases at early stage. This algorithm can be used on heart records of the patient or by using it on classiﬁcation reports. This research was conducted and tested upon various algorithms to test its accuracy like Logistic Regression, Random Forest, Vector Support and XG-Boost. After applying these algorithms of prediction model has been developed.

## Existing problem

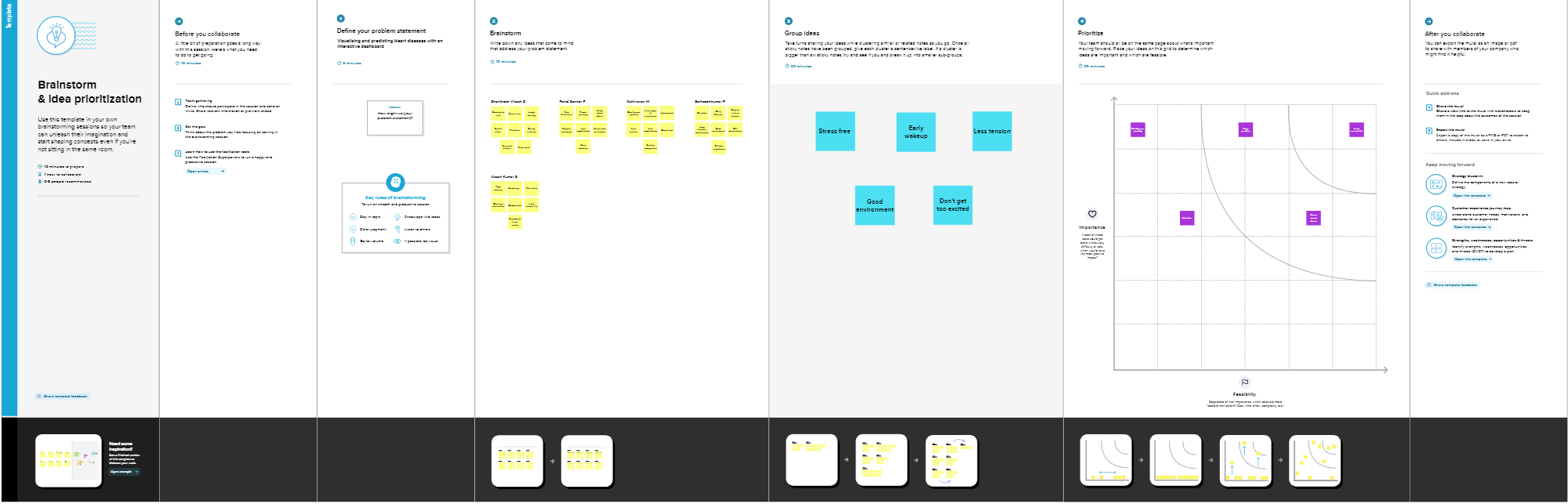


* 1. Problem statement

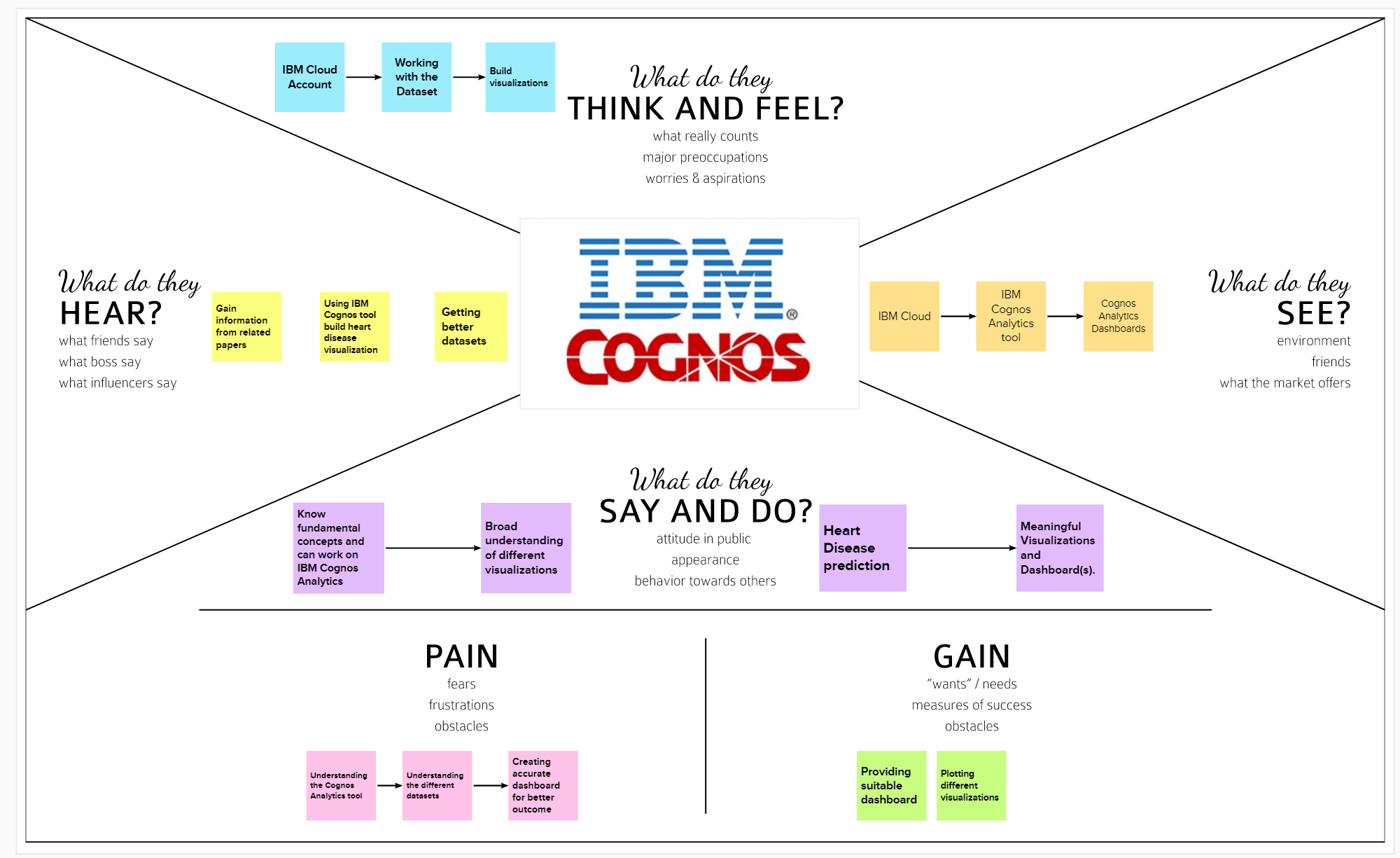
It is not possible to monitor patients every day in all cases accurately and consultation of a patient for 24 hours by a doctor is not available since it requires more sapience, time and expertise. Since we have a good amount of data in today’s world, we can use various machine learning algorithms to analyze the data for hidden patterns. The hidden patterns can be used for health diagnosis in medicinal data.

# IDEATION PHASE

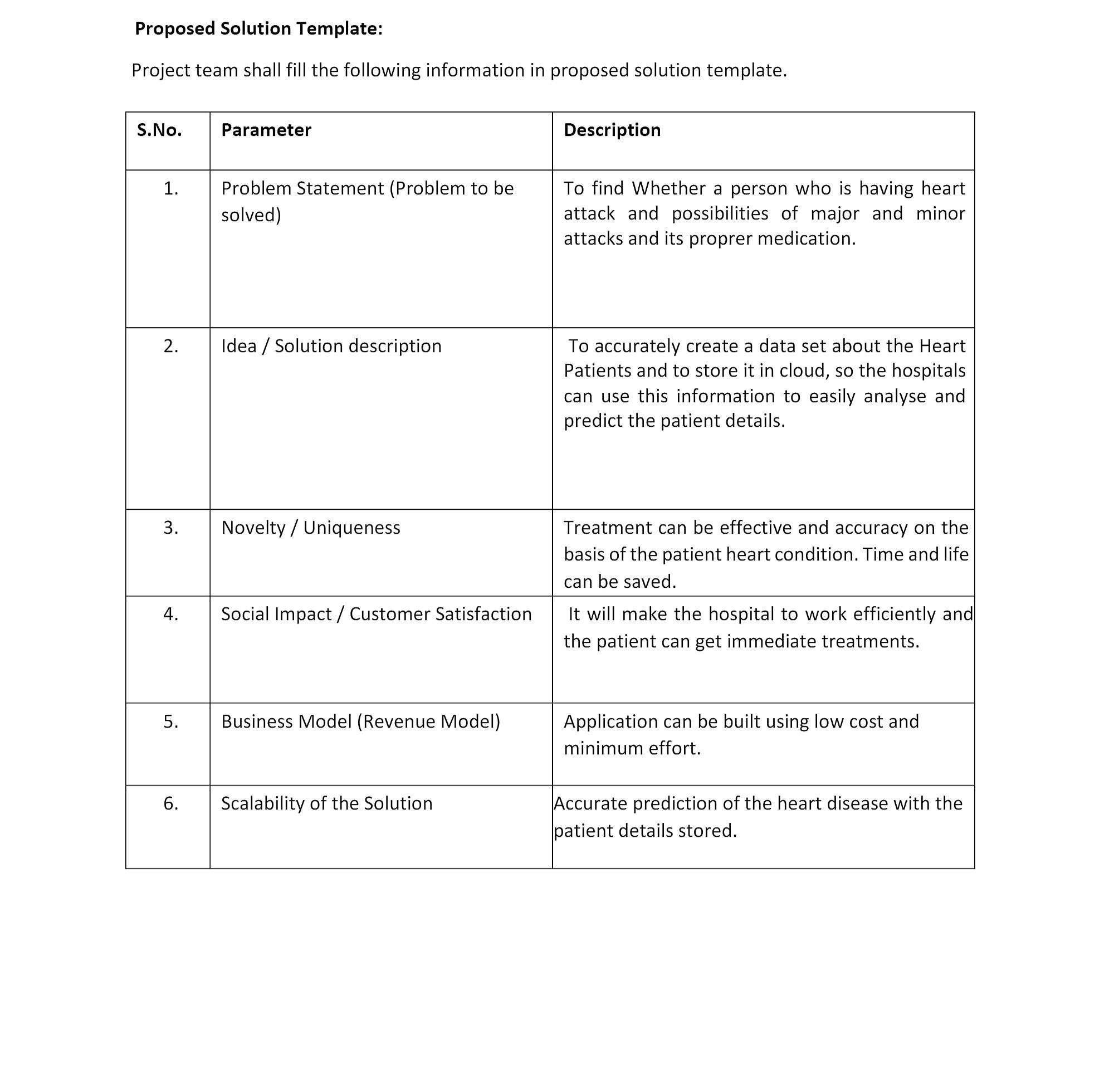
## Brainstorm & idea prioritization



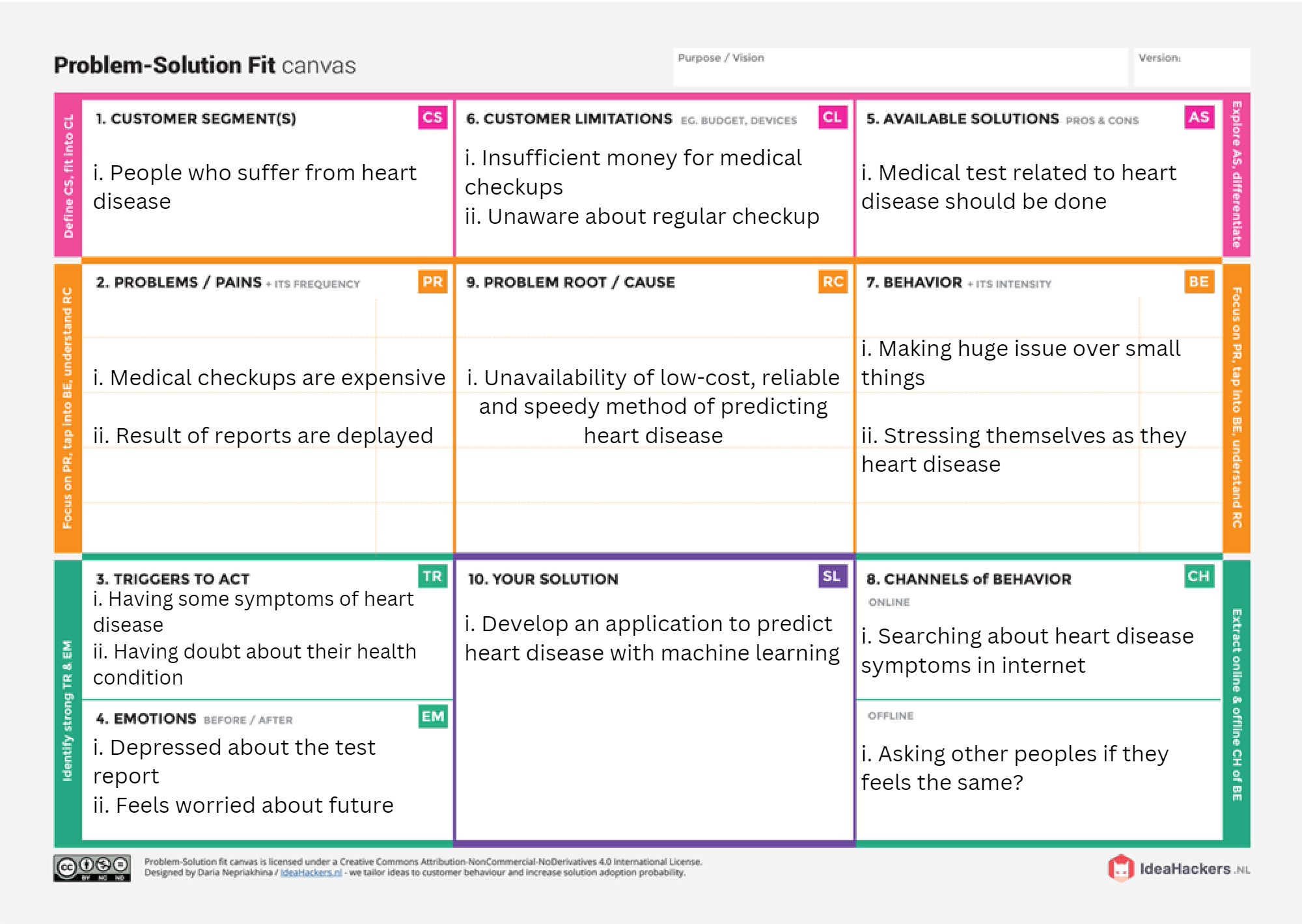
Empathy map

It is a useful tool to helps teams better understand their users. Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user’s perspective along with his or her goals and challenges. 

Proposed Solution



Problem Solution Fit



1. REQUIREMENT ANALYTICS
   1. Functional Requirement

Following are the functional requirements of the proposed solution.

|  |  |  |
| --- | --- | --- |
| **FR No.** | **Functional Requirement (Epic)** | **Sub Requirement (Story / Sub-Task)** |
| FR-1 | User Registration | Enables user to make registration for the application through Gmail |
| FR-2 | User Conﬁrmation | Once after registration, the user will get conﬁrmation  via Email |
| FR-3 | Visualizing Data | User can visualize the trends on the heart disease through Dashboard created using IBM Cognos Analytics |
| FR-4 | Generation Report | User can view his/her health report and can make decisions accordingly |

## Non-Functional Requirement

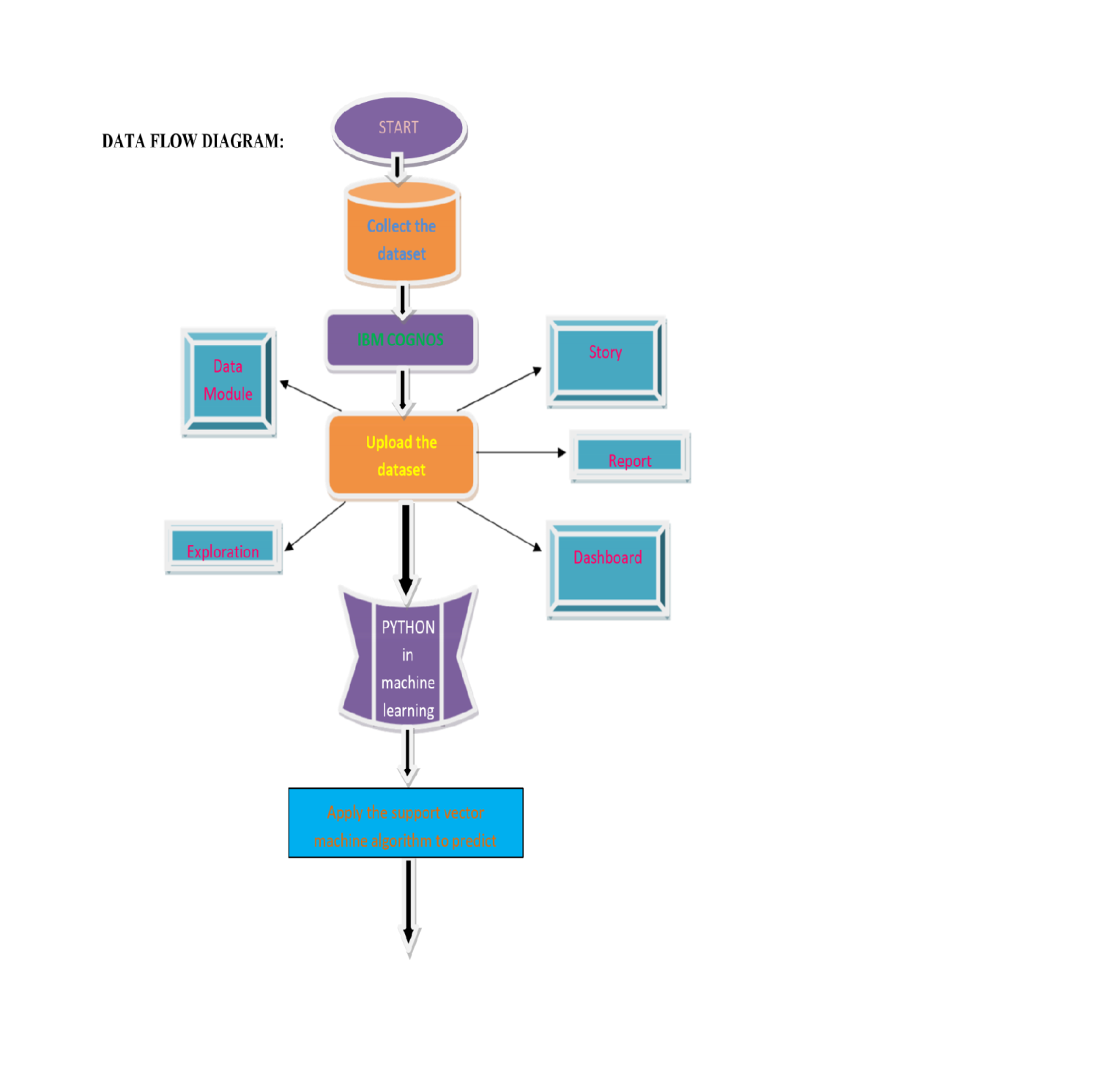
|  |  |  |
| --- | --- | --- |
| **FR No.** | **Non-Functional Requirement** | **Description** |
| NFR-1 | Usability | The application will have a simple and userfriendly graphical interface. Users will be ableto understand and use all the features of the application easily.  Any action has to be performed with just a few clicks |
| NFR-2 | Security | For security of the application the technique known as database replication should be used so that all the important data should be kept safe. Incase of crash, the system should be able to backup and recover the data |
| NFR-3 | Reliability | The application has to be consistent at every  scenario and has to work without failure in any environment |
| NFR-4 | Performance | Performance of the application depends on the  response time and the speed of the data submission. The response time of the application |

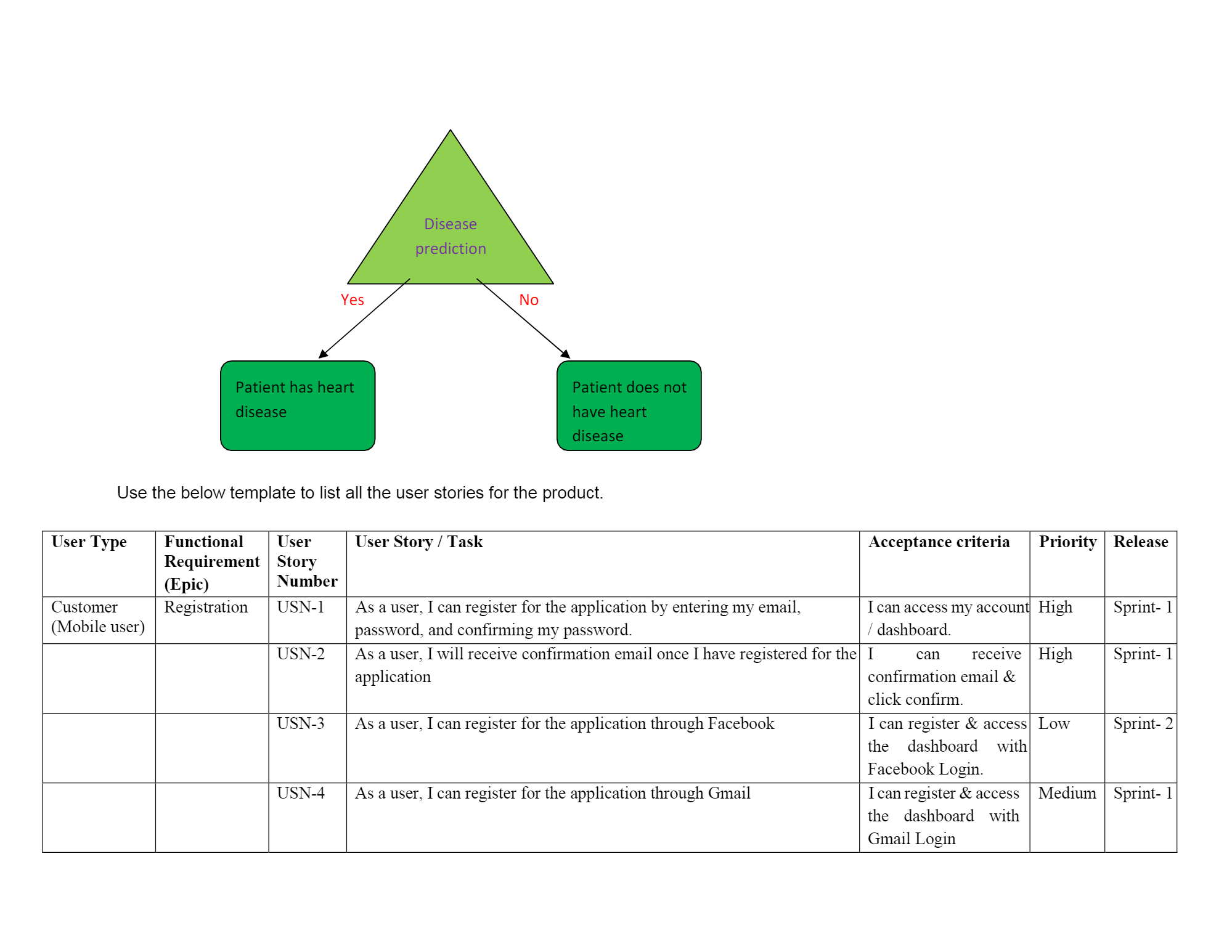
|  |  |  |
| --- | --- | --- |
|  |  | Is direct and faster which depends on the eﬃciency  of implemented algorithm |
| NFR-5 | Availability | The application has to be available 24 x 7 for users without any interruption |
| NFR-6 | Scalability | The application can withstand the increase in the  no. of users and has to be able to develop Higher versions |

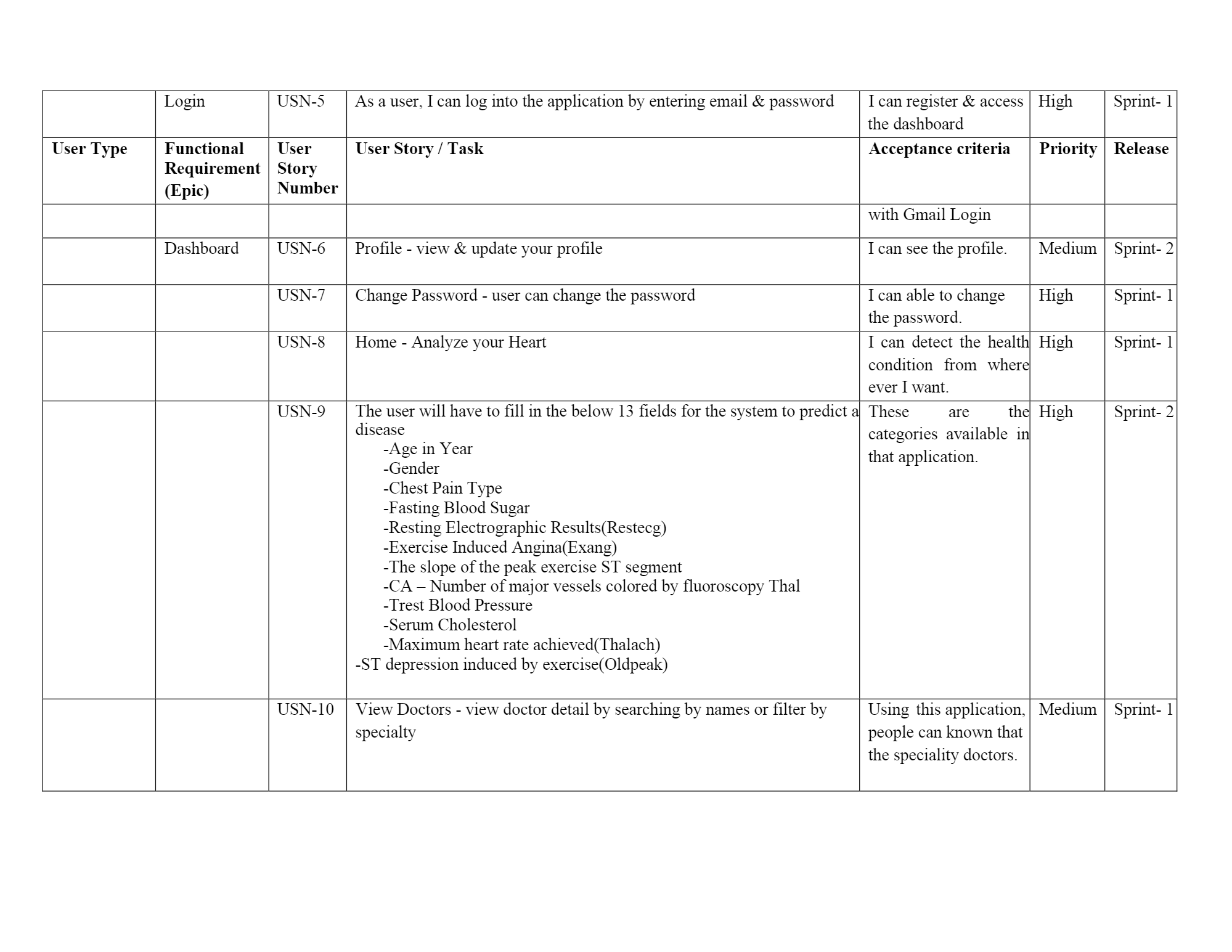
## 5 PROJECT DESIGN

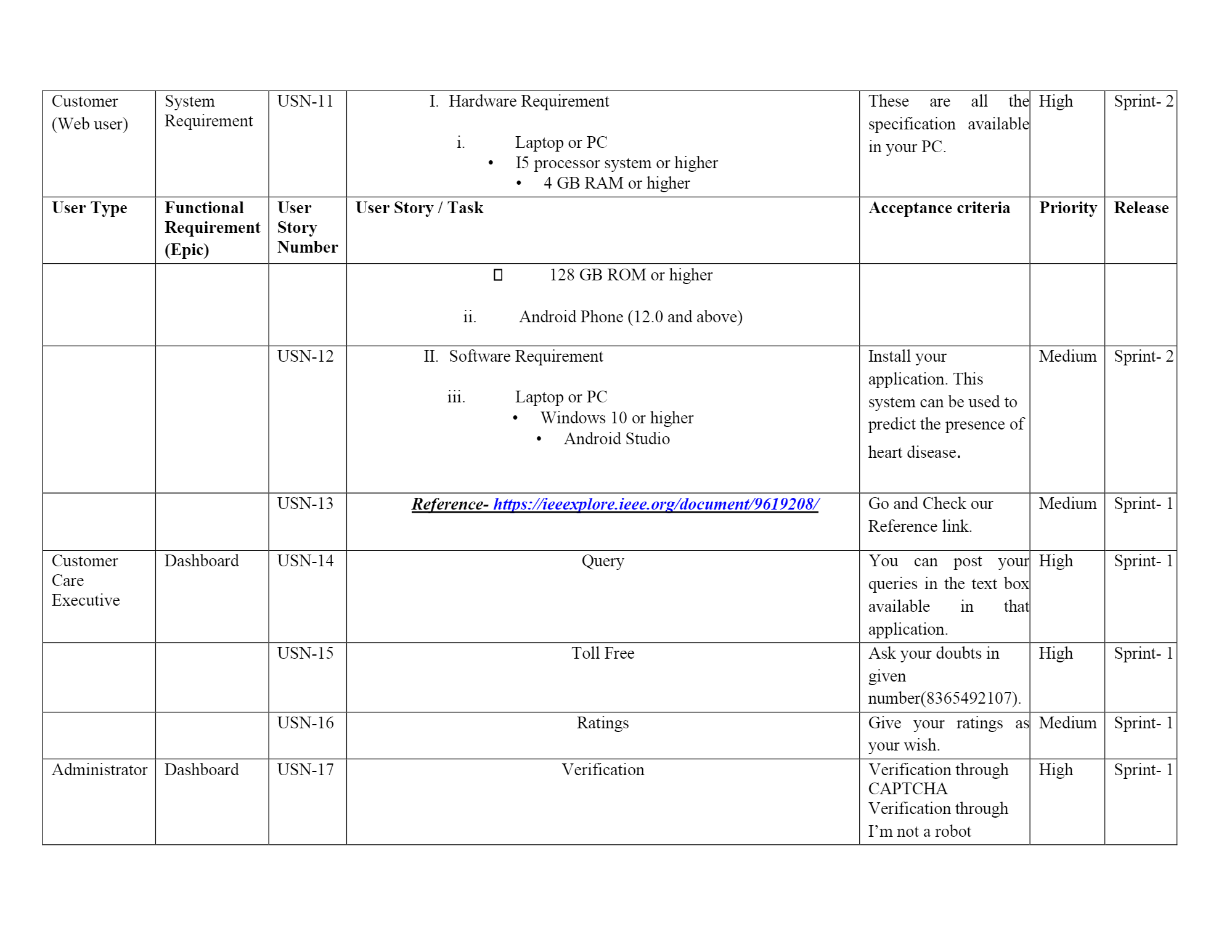
Data Flow Diagram

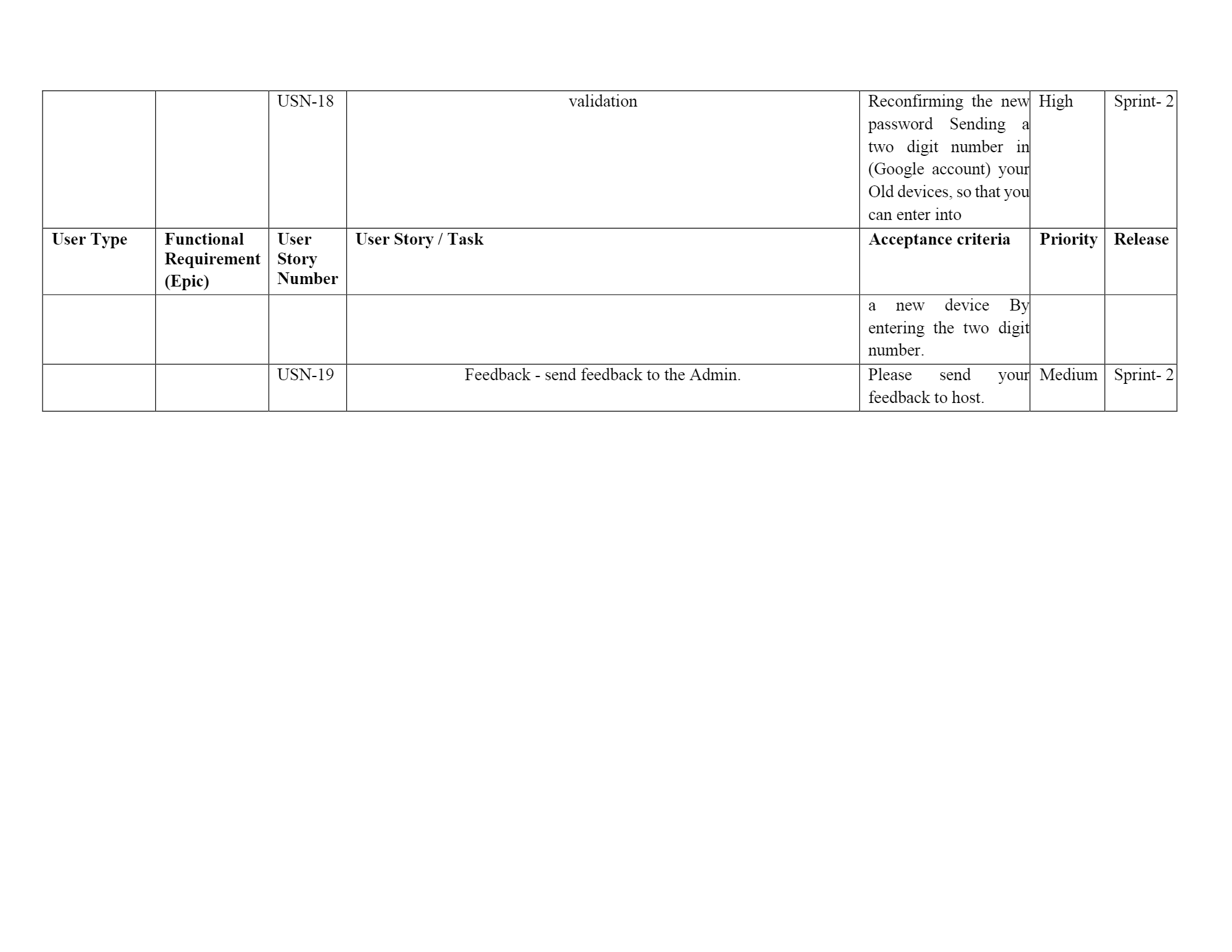
A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored









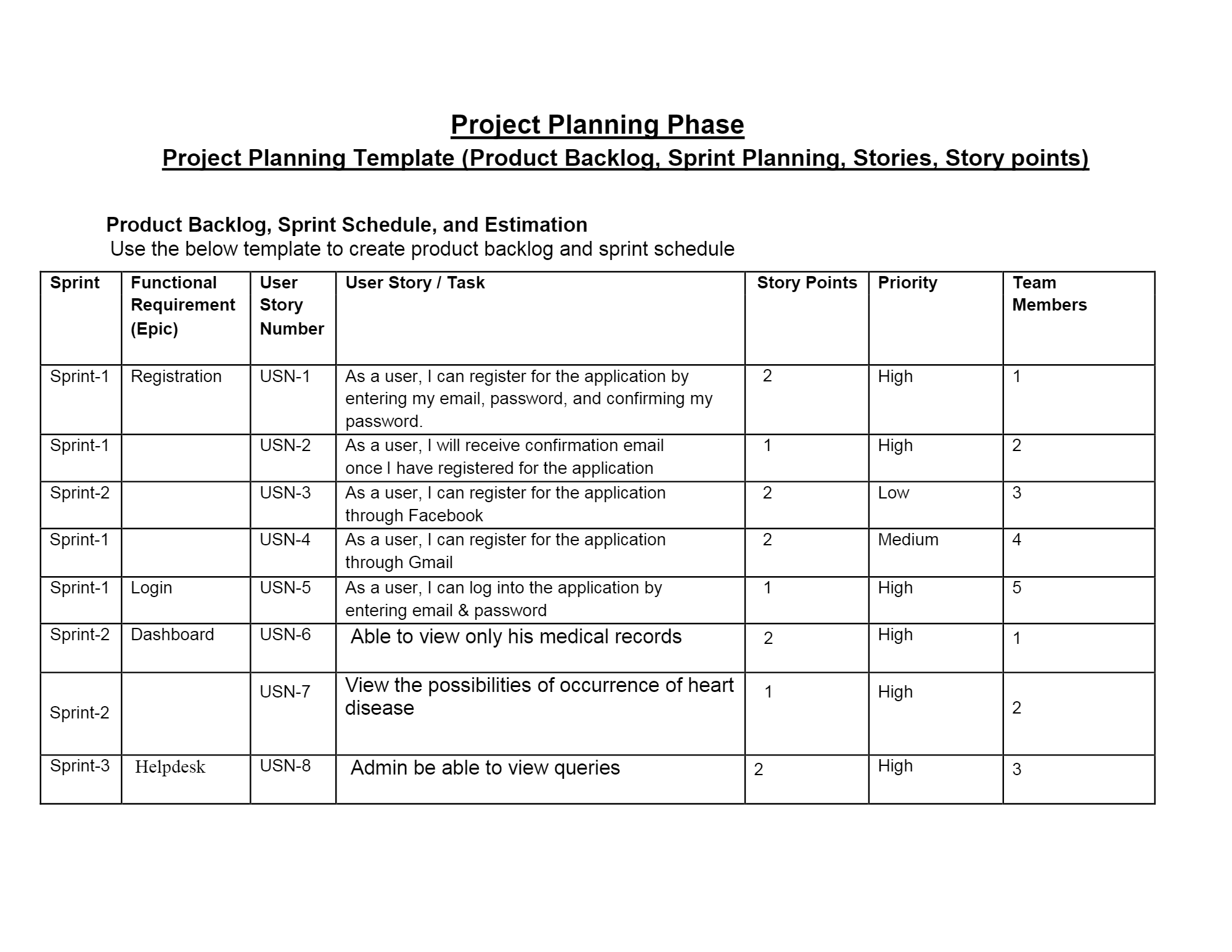


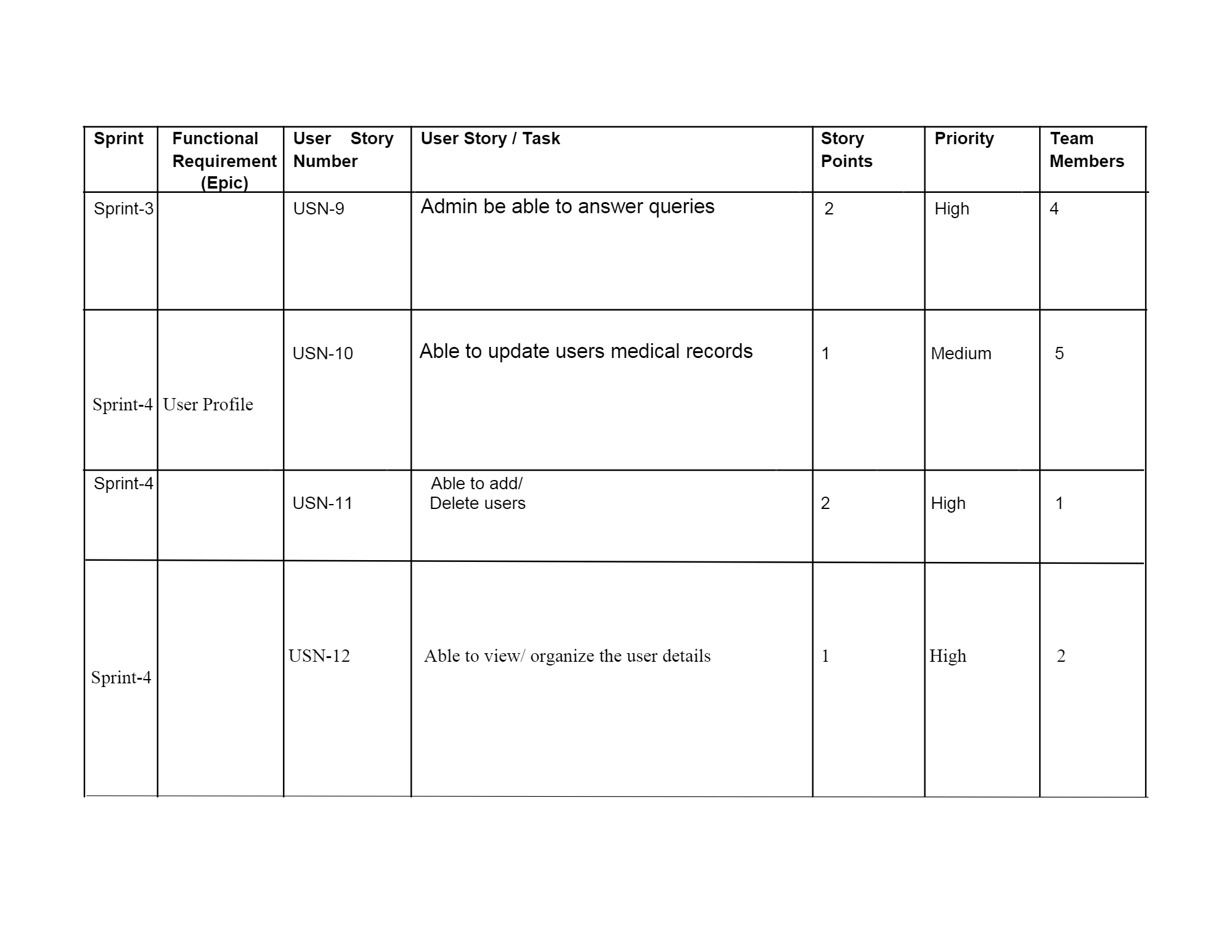
Solution & Technical Architecure



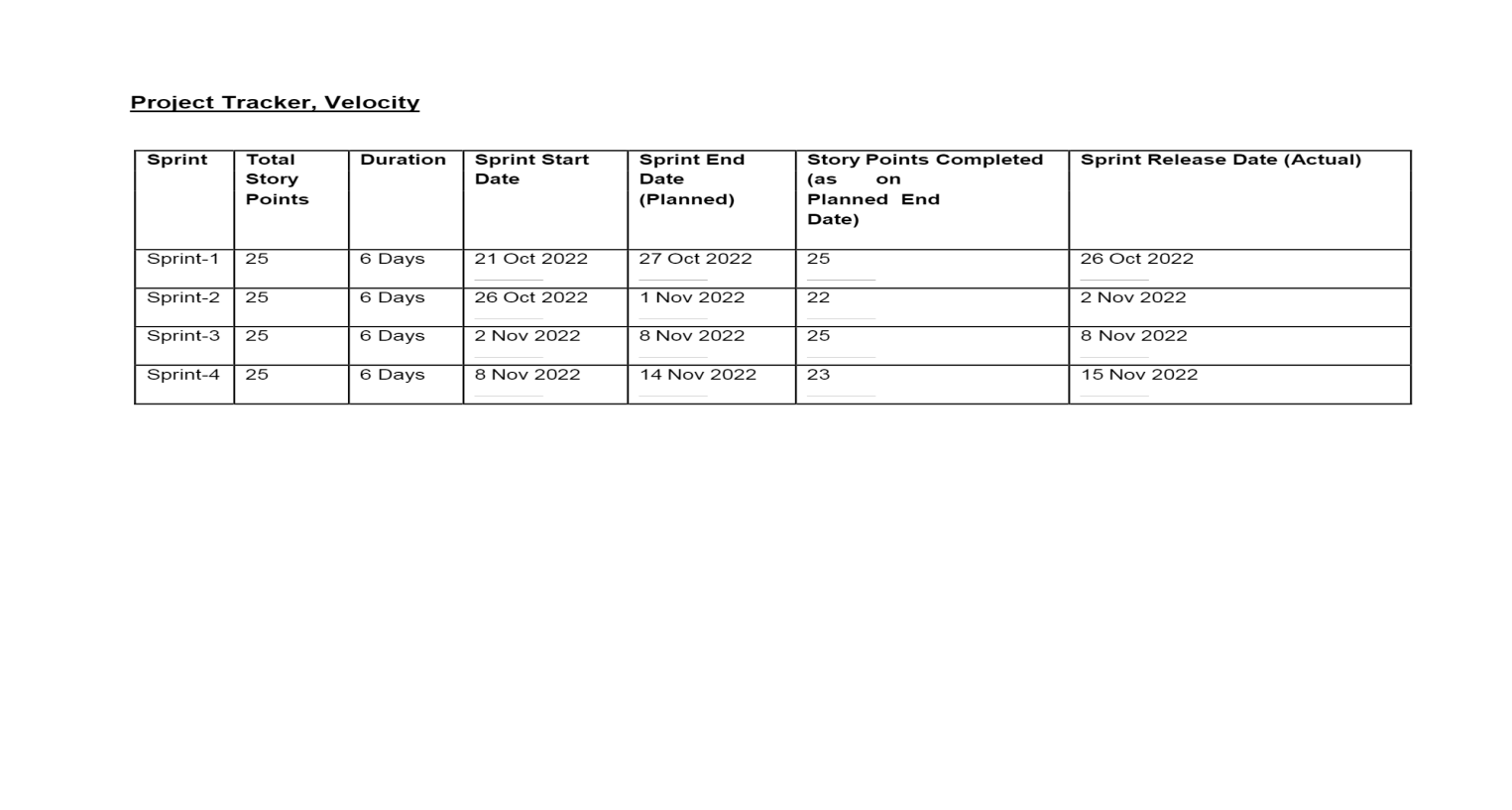
# PROJECT PLAINING & SCHEDULING

## Sprint Planning & Estimation





* 1. Sprint Delivery Schedule

7 CODING & SOLUTIONING

* 1. Machine Learning

Learning which model is best for the given Dataset



Comparing it with the accuracy gotten from Decision Tree: TP = cm[0][0] #cm=Confusion Matrix

TN = cm[1][1]

FN = cm[1][0]

FP = cm[0][1]

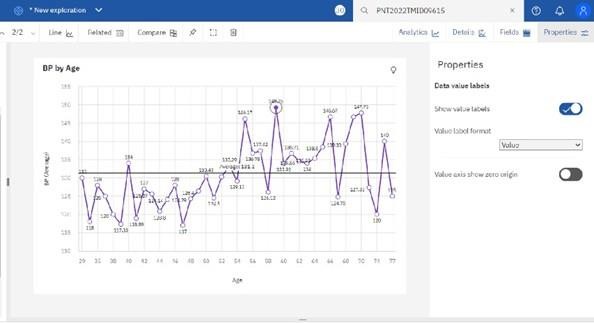
print('Testing Accuracy for Decision Tree:',(TP+TN)/(TP+TN+FN+FP)) print('Testing Sensitivity for Decision Tree:',(TP/(TP+FN)))

print('Testing Speciﬁcity for Decision Tree:',(TN/(TN+FP)))

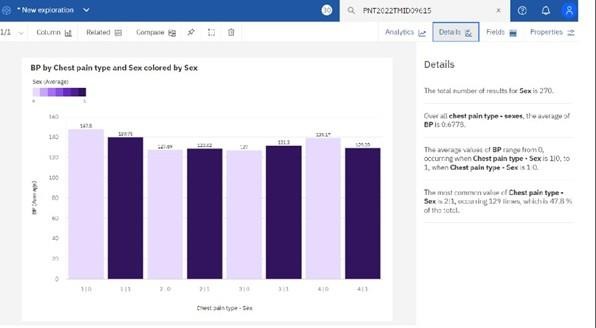
print('Testing Precision for Decision Tree:',(TP/(TP+FP)))



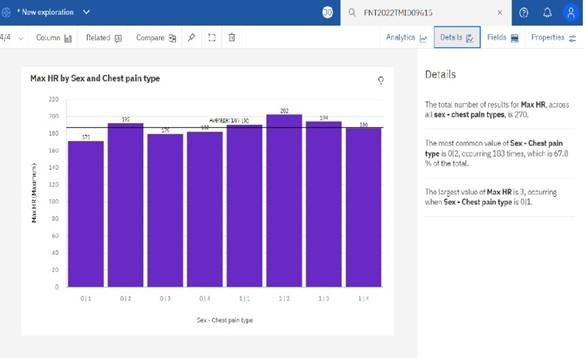
## Dash Board

Average BP during chest pain

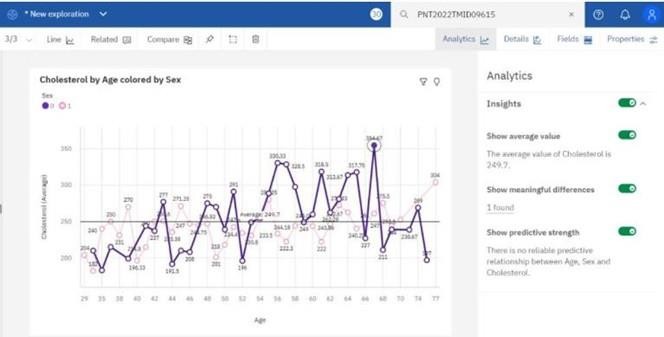
Exploration Of BP vs Chest Pain Type and Gender:



Exploration Of Max Heart Rate During the Chest Pain:



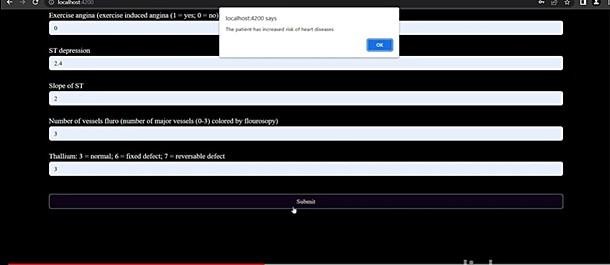
Exploration Of Cholesterol by age and Gender:



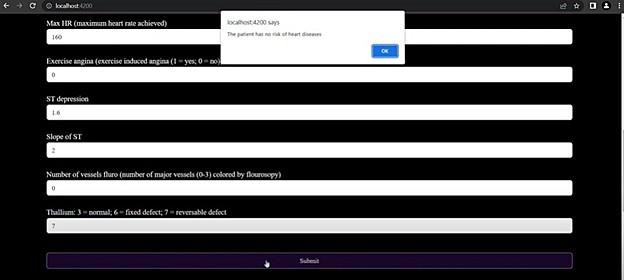
# TESTING

## Test Cases

* 1. User acceptance Testing



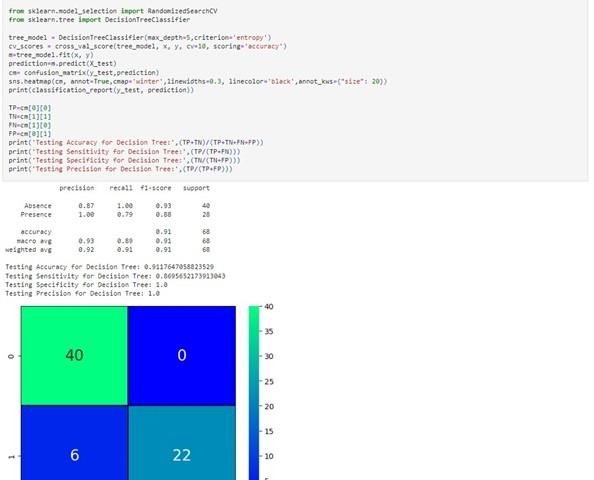
Testing a case where user does not have heart disease



# RESULT

## Performance Metrics

The confusion matrix below shows the performance metrics



# ADVANTAGE AND DISADVANTAGE

Advantages:

1. This is one of the fastest ways to determine if a person is likely to suffer from a heart disease or not.
2. Useful for medical practitioners to easily classify their patients.
3. User Friendly
4. Easy to understand
5. Secure
6. Dashboard provides insightful information Disadvantages:
7. Needs work
8. Users need to know all the ﬁelds
9. Does Not take null value as input
10. Does not provide suggestions to user

## CONCLUSION

Complications of heart disease include heart attack and stroke. You can reduce the risk of complications with early diagnosis and treatment. So the suggestion that we get from the website might help save patients. It is always to get treated in the early stages of heart disease.

1. FUTURE SCOPE

Like the saying goes “Prevention is better than cure”. We have to look into methods to prevent heart diseases altogether other than just predicting it in early stage.To use this website we need to take a lot of tests beforehand. So it would be better if we require less attributes and still give an effective result

1. APPENDIX

Source code:

*https://github.com/IBM-EPBL/IBM-Project-6798-1658837472/upload/main/FINAL%20DELIVERABES*