

L. D. College of Engineering

A Report on

SEPSIS ANALYSIS AND PREDICTION

Under subject of Project-1

B. E. IV Semester VII Computer Engineering

Submitted By:

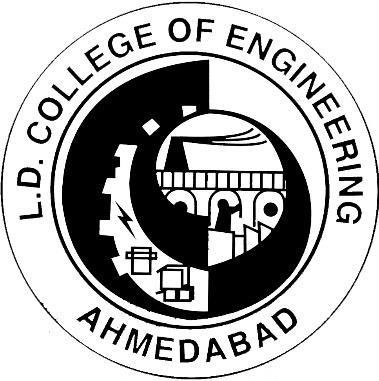
|  |  |  |
| --- | --- | --- |
| Sr No. | Enrolment Number | Name |
| 1 | 160280107097 | Adarsh Shah |
| 2 | 160280107105 | Rushi Shah |
| 3 | 160280107033 | Jay Khatri |

Prof. Zishan Noorani Faculty Guide

Dr.S.M. Shah (Head of Department)

Academic Year 2017 - 18

## CERTIFICATE



Computer Engineering 2019-2020

##### Date:

This is to certify that the final year project entitled “**SEPSIS ANALYSIS & PREDICTION”** has been carried out by Mr. Adarsh Shah (160280107097), Mr. Rushi Shah (160280107105) and Mr. Jay Khatri (160280107033) under my guidance in fulfillment of the Degree of Bachelor of Engineering in Computer Engineering - 8th. Semester of Gujarat Technological University, Ahmedabad during the academic year 2018- 2019.

**Name of Guide Head of Department**

**Mr. Zishan Noorani** Assistant Professor Computer Engineering

**Dr. S.M. Shah** Associate Professor Computer Engineering

**CANDIDATE’S DECLARATION**

We have finished our project report entitled “**SEPSIS ANALYSIS & PREDICTION”** and submitted to our respected guides. We are in 7th semester and we have tried our best. We have done our work honestly and in a good way.

First candidate’s name : **ADARSH SHAH**

Branch : **CE**

Enrollment No. : **160280107097**

Second candidate’s name : **RUSHI SHAH**

Branch : **CE**

Enrollment No. : **160280107105**

Third candidate’s name : **Jay Khatri**

Branch : **CE**

Enrollment No. : **160280107033**

Submitted to:

### L.D. COLLEGE OF ENGINEERING, AHMEDABAD

Affiliated by:

### GUJARAT TECHNOLOGICAL UNIVERSITY

ACKNOWLEDGEMENTS

We wish to express our heartfelt appreciation to all those who have contributed to this project, both explicitly and implicitly, without the co-operation of whom, it would not have been possible to complete this project.

We would like to thank our guide Prof. Zishan Noorani for constantly guiding and showing us the correct path to reach towards our desired goal. Also we thank them for sharing their experience, knowledge and valuable time with me and showing their concern in our project to make it better.

We would like to extend our gratitude to Dr. S.M. Shah, Head of Computer Engineering Department, L.D. College of Engineering, Ahmedabad, for his continuous encouragement and motivation

We would also like to thank all the other faculties who directly or indirectly supported us in making our project successful by sharing their ideas and knowledge.

At last, we would like to thank our parents, friends and seniors who have directly or indirectly helped us in making our project work successful.

NAME:

Adarsh Shah (160280107097) Rushi Shah (160280107105)

Jay Khatri (160280107033)

**ABSTRACT**

***“DSP(Deep Sepsis Predictor)”*** is a unique sepsis prediction mechanism that uses Deep Neural Network in order to predict  whether the patient has sepsis infection or not. DSP is built using the Keras framework written in python language. Sepsis is a severe disease which occurs when chemicals released in the bloodstream to fight an infection trigger inflammation throughout the whole body. This can cause many changes that damage multiple organ systems, leading them to fail, sometimes even resulting in death. Symptoms include fever, difficulty breathing, low blood pressure, fast heart rate and mental confusion. DSP is constructed using two tier architecture, one makes up the model which does the prediction and the other layer consists of Web Portal which makes the user interface easier.  Our model uses 96 parameters for its input and all of the variables can be obtained from the patient's ICU data. Our model shows a precision score of 95.83. In order to train our model we have used Mimic-III dataset which we have obtained from MIMIC-III Clinical Database v1.4. MIMIC-III is a restricted-access database comprising of de identified health-related data associated with over forty thousand patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012. Hence we have used state of the art data in order to produce such accurate results.

Index

|  |  |  |
| --- | --- | --- |
| Number | Title | Page Number |
|  |  |  |
|  | Title Page |  |
|  | GTU Completion Certificates |  |
|  | Declaration |  |
|  | Acknowledgements |  |
|  | Abstract |  |
|  | List of Figures |  |
|  | List of Tables |  |
|  |  |  |
| 1 | Project Details | 1 |
| 1.1 | Definition | 1 |
| 1.2 | Basic Objective | 1 |
| 1.3 | Working Principle | 1 |
|  |  |  |
| 2 | Introduction | 2 |
| 2.1 | Background Details | 2 |
| 2.2 | Overall Description | 2 |
| 2.3 | Tools and Technology | 2 |
| 2.4 | Environmental Characteristics | 2 |
| 2.5 | Literature Review | 3 |
|  |  |  |
| 3 | Requirement Analysis | 5 |
| 3.1 | Functional Requirements | 5 |
| 3.2 | Non-Functional Requirements | 5 |
| 3.3 | Behavioural Description | 6 |
|  |  |  |
| 4 | Scheduling | 7 |
|  |  |  |
| 5 | Data Flow Diagrams | 8 |
|  |  |  |
| 6 | ER Diagram | 10 |
|  |  |  |
| 7 | Steps to deploy the project | 11 |
|  |  |  |
| 8 | UML Diagrams | 12 |
| 8.1 | Use Case Diagram | 12 |
| 8.2 | Class Diagram | 14 |
| 8.3 | Activity Diagram | 16 |
|  |  |  |
| 9 | User Interface Design | 18 |
| 9.1 | Use Case Diagram | 18 |

|  |  |  |
| --- | --- | --- |
| 9.2 | Data Dictionary | 19 |
| 9.3 | Interface Description | 20 |
|  |  |  |
| 10 | Conclusion | 21 |
|  |  |  |
| 11 | References | 22 |
|  |  |  |
|  | Appendix | 23 |
|  | Design Engineering Canvas | 23 |
|  | Periodic Progress Reports | 27 |
|  | Patent Search and Analysis Reports | 35 |
|  |  |  |
| 12 | Plagiarism Certification | 55 |

List of Figures

|  |  |  |
| --- | --- | --- |
| Number | Title | Page Number |
|  |  |  |
| 4.1 | Iterative model | 7 |
| 5.1 | Data Flow Diagram Level 0 | 8 |
| 5.2 | Data Flow Diagram | 9 |
| 6.1 | ER Diagram | 10 |
| 8.1.1 | Use Case Diagram (BMI) | 12 |
| 8.1.2 | Use Case Diagram (Blood Connect) | 12 |
| 8.1.3 | Use case Diagram (Doctor) | 13 |
| 8.1.4 | Use case Diagram (Disease) | 13 |
| 8.2.1 | Class Diagram (BMI) | 14 |
| 8.2.2 | Class Diagram( Blood Connect) | 14 |
| 8.2.3 | Class Diagram (Doctor) | 15 |
| 8.2.4 | Class Diagram (Disease) | 15 |
| 8.3.1 | Activity Diagram (BMI) | 16 |
| 8.3.2 | Activity Diagram (Blood Connect) | 16 |
| 8.3.3 | Activity Diagram (Doctor) | 17 |
| 8.3.4 | Activity Diagram (Disease) | 17 |
| 9.1 | User Interface Diagram | 18 |
| A.1 | AEIOU Summary | 23 |
| A.2 | Empathy Map | 24 |
| A.3 | Ideation Canvas | 25 |
| A.4 | Product Development Canvas | 26 |

List of Tables

|  |  |  |
| --- | --- | --- |
| Number | Title | Page Number |
|  |  |  |
| 9.2.1 | Login/signup Table | 19 |
| 9.2.2 | BMI Table | 19 |
| 9.2.3 | Blood Donation & Receiver Table | 19 |
| 9.2.4 | Find me a Doctor Table | 19 |
| 9.2.5 | Disease Prediction Table | 20 |

# Project Details

#### Definition

The core aim of Sepsis Analysis and Prediction is to cure patients by early sepsis prediction through the use of our trained neural network model. It has been suggested by many doctors that the main reason why so many patients die due to sepsis infection is because the disease is discovered in the body when it has reached a severe stage. Hence in order to discover the disease in an early stage we have constructed a network that uses statistical reasoning and deep networks for predicting the sepsis positive patients. In order to achieve our goal we have created a dense neural network containing 45,561 trainable parameters in a total of 6 layers. Also the accuracy of the model significantly increases as the time of sepsis infection passes, but we are trying to give accurate results from the beginning.

#### Basic Objective

Sepsis is a life threatening disease and every year 1 million cases are registered only in India regarding Sepsis infection. Hence the main objective of our final year project is to try to decrease the number of cases by discovering most of the cases in early stages so that they can be cured with ease. In order to accomplish the task, we have constructed a Neural Network using data obtained from the MIMIC-III dataset available online. We have also provided an online user interface to freely interact with our model. Mimic-III database is a very powerful and diverse database which is available online. It is a trusted database  and contains data of over forty thousand patients. Also our model takes 96 different medical data variables from patients out of which all of them come from patient’s ICU information. After processing the data, our model suggests the patient using past information and Artificial Intelligence whether the patient has Sepsis Infection or not. The patient can check the result directly on the website.

#### Working Principle

We have created the model using keras framework on the structure of tensorflow library. Keras is a multipurpose framework which is very robust in creating various types of neural networks.In order to efficiently train and validate the data , we have divided the data set in two parts , one for training which constitutes the training phase and test data on which we have validated our model.

In order to obtain the required variables for the training of our model we have used the PostGres Sql queries to efficiently dig the data that is related to the Sepsis disease. We have created a complex query to obtain the 96 variables needed for the training model. We decided the required variables by suggesting doctors and doing more research on the disease.

We have trained our model at 50 epochs, 60 epochs, and 40 epochs. By Try and Error, we found the best result at 50 epochs, hence we have used the model trained at 50 epochs. Also, we have used 45561 total parameters out of which all are trained in the training phase. This has helped the model to become acquainted with the data of the patient very well and hence gives almost accurate results. The output of the model is quite simple, it gives 0 if the patient is diagnosed negative with sepsis else 1. The output is then sent to the website and displayed to the user whose data has been submitted.

# Introduction

#### Background Details

Sepsis is a life threatening situation in which the inflammation occurs in the body causes a series of changes which severely damage the bloodstreams and internal organs as a result of which the patient likely dies. The Sepsis disease has a total of three differentiated stages: 1)Sepsis 2)Severe Sepsis 3)Septic Shock. The likelihood of patient’s survivability decreases gradually in severe sepsis and the chances of death are very low in the first stage of sepsis while the condition becomes life threatening in the third stage of sepsis, septic shock. In the advent of the Big Data Age, many people have tried to automate the discovery of sepsis in the patient using different approaches  like K-Nearest Neighbour Algorithm, Logistic Regression and so on. We have tried a different approach to tackle the disease. We have tried to give results using Neural Networks. Neural Networks are computer Systems inspired by our own brain and are very efficient at learning from data without using any hardcoded rules.

We have obtained the data from the <https://physionet.org/content/mimiciii/1.4/>. Mimic-III is the third version of mimic database, the first two being mimic-1 and mimic-11. Mimic-III is the home to data of over forty thousands patients at Beth Israel Deaconess Medical Center between 2001 and 2012. The data contains more than three hundred different ICU values of a single patient. Hence the data alone gave us a wide range of spectrum to discover the dependent variables of Septic Shock. Of all the different data variables available to us, we decided on the 96 ones that were connected to sepsis in one way or another.

#### Overall Description

Our project is made up of two very important layers. The core of our project is the Deep Neural Network Model trained with 45561 parameters at 6 layers for 50 epochs taking 96 inputs. It is meticulously created by try and error on by trying various different parameters and hyper parameters like epoch size, learning rate, number of parameters, etc. One other reason for our high precision rate is the huge amount of data and large number of dependent variables available at our disposal. As stated above the data is obtained from the MIMIC-III database.

#### Tools and Technology

The technologies used to implement this project are:

* Python language  version 3.6.0 or above
* Tensorflow framework 1.1.0 or above
* Postgresql database server
* GPU if available to increase the computation speed

#### Environmental Characteristics

#### Hardware and Peripherals

* Processor: > Intel Pentium 5/Dual Core
* Free Space: 200 GB minimum
* RAM : 4 GB minimum

#### Users

* Patients
* Doctors
* Pathologist
* Admin

#### Literature Review

There are many mobile applications that provide facilities to calculate BMI , find a specialist Doctor and Finding a blood donor but in different packages.

1. BMI Calculator : BMI Calculator is free app that allows you to monitor BMI and percentage of fat in your body. App is designed for people of different ages and supports for both metric and imperial but is unable to give solutions.
2. Find a Doctor: It allows quick search for doctors by name, workplace or speciality near your current location, also allows booking appointments with several registered doctors but is unable to search doctors by City area.
3. Blood Donation App- India : This is free blood donation mobile application which helps needful patients to search for blood donors in their city or area and request for blood donation. Helps in saving lives of those who are in immediate need of blood . However, it does not support OTP based Authentication.
4. Symptomate – Symptom checker: Symptomate is the advanced symptom checker that uses AI to evaluate your symptoms .It analyzes the symptoms and find out what could be causing them. However, it does not find or predict disease on the basis of symptoms entered.

In conclusion, there are many applications that supports some of the facilities such as merely calculating the BMI , finding a doctor , donating blood and also checking symptoms . Hence, our project aims at building solution to this different problems in one application only.

The material that has been referred by us in order to get insight in the work that has been done in the field of Health care technology can be found in the appendix section of this report.

List of Algorithms studied:

1. Obesity Algorithm
2. Grand Rounds Algorithm
3. CART Algorithm
4. ID3 Algorithm

# Requirements Analysis

#### Functional Requirements

The main functional requirements of our project are as under:

1. User interface to interact with the system.
   1. Facility to single or multiple inputs.
   2. Facility to start and stop searching process.
   3. Facility to view the output results.
2. Getting inputs in proper format.
   1. Getting inputs in proper format for BMI calculation.
   2. Getting exact inputs for proper prediction .
   3. Getting inputs in either metric or imperial format.
3. Searching on basis of Symptoms.
   1. Going through every symptoms to predict exact result.
4. Combining entered inputs to get an enhanced and filtered data set.
   1. Using height and weight to calculate BMI.
   2. Examining every symptoms for proper prediction.
   3. Combining inputs like Blood group and location for faster search.
   4. Combining inputs like city area and field to get a proper specialist .

#### Non-Functional Requirements

1. The system should respond quickly. The user interface should respond within 1s of an event by the user.
2. The Search must be performed in a reasonable time. A single search must be searched within a minute.
3. The results should be searched with sufficient accuracy. The system must have an accuracy of 80% or better.

#### Behavioural Description

* + 1. System States

The system states are as under:

1. The system is accepting inputs from the user.
2. The system is searching the results in the databases.
3. The system outputs the results.
4. The search is paused.
5. The search was stopped before it produced the output.

#### Events and Actions

The events and actions performed by the user are as under:

1. The user inputs a Blood Group.
2. The user inputs multiple inputs like city area , field and diseases.
3. The user initiates searching.
4. The user pauses searching.
5. The user resumes searching.
6. The user stops searching.
7. The user views the output generated by the system.

# Scheduling

**SDLC** - **Prototype Model**. : For our Sepsis Predicting model, we have used the Prototype Software approach. We have successfully created our project on the basis of two previous prototypes. We have used the prototype model because it was very efficient for us to create new neural network models on the basis of past models and using that information in order to increase various performance metrics.

In our first prototype model, we tried to preprocess the data that was available to us by the mimic-iii database. It contained a huge amount of information about myriad diseases. Thus it was necessary to prune the data and obtain only the necessary data variables required for our study. For this we had used Postgresql Server in order to create complex queries that could dig out the important information from the huge database. In the end we were able to obtain data of around ten thousands patients regarding sepsis and each one having 96 dependent variables for sepsis infection.

For the Model, in our first attempt we were able to create a model with 5 layers. Each layer had neurons in the form 2 raised to n, like 1024,512,etc.

But the first model was not practical as it had a low precision score of 70%.In the first attempt we were training in the ratio of 95:5 and we found out that it was overfitting the data and had very few validation samples. Thus for our new model we changed the ration to 90:10. Also we tried different approaches like changing learning rate, epochs and even tried to add more layers to it. In the end we were able to find the sweet spot of our project at 6 layers,45561 trainable parameters and 50 epochs. We also tried to change the layer structure by changing 1024,512,etc to 1000,500 and so on. Hence after a few more iterations of train and test methods, we were able to achieve a precision of 96%. Now our model was finally completed.

In the final iteration of our Prototyping process, we started creating a user interface in order to show the results and take inputs from various ICU units. We started creating the website.

# Data Flow Diagram

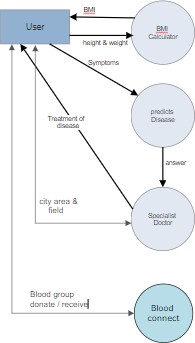
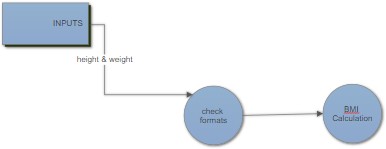
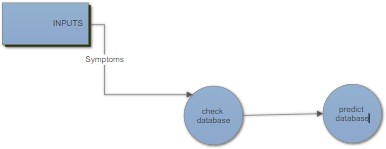
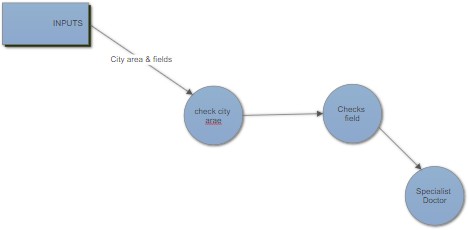


Figure 5.1 - Data Flow Diagram Level 0







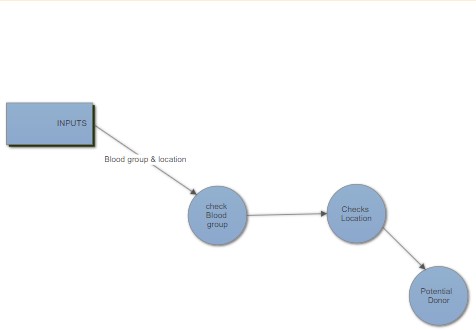


Figure 5.3 - Data Flow Diagram

# ER Diagram

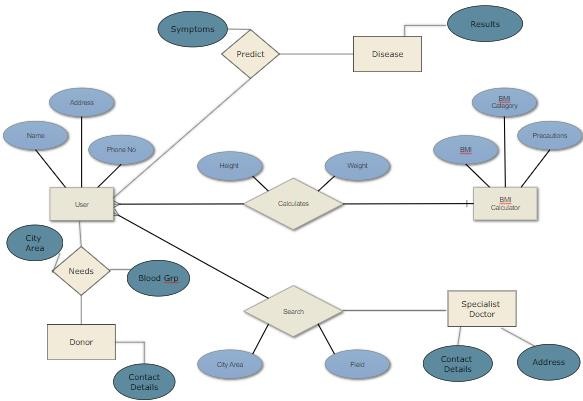


Figure 6.1 – ER Diagram

# Steps to Deploy the Project

So far we have done the following:

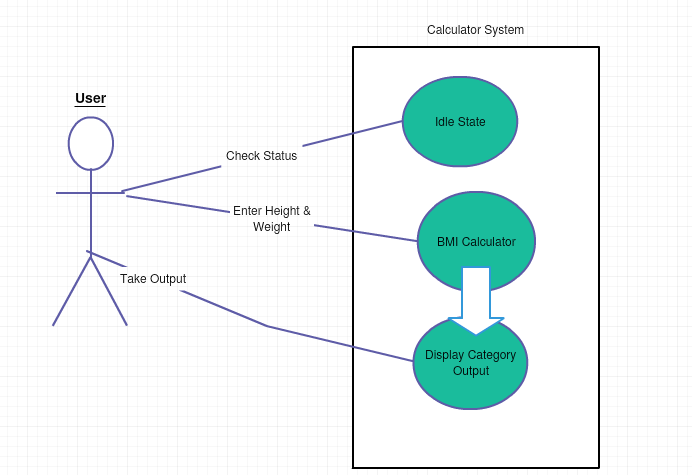
* 1. First we analysed the problem thoroughly and understood each and every aspect of the problem.
  2. After the analysis was done we designed a solution of the problem.

The end product of analysis and design was a specification and UML diagrams. In the 8th semester we will carry forward from the work we did in this semester and code as well as test the project. We plan on doing the following in the 8th semester:

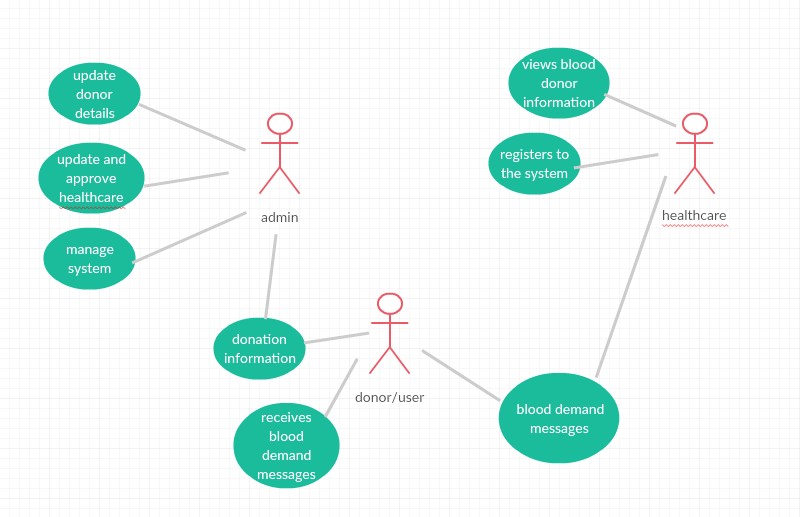
1. Implement the module to Search Exact Disease from the entered Symptoms.
2. Test the above module and optimize it.
3. Implement the module to Find the nearest and expert doctor in a particular field.
4. Test the above module and make necessary changes to the module mentioned in 1, if necessary.
5. Implement a module to differentiate between Two different Diseases by the module in 1.
6. Implement a module to separate verified donors from unverified donors.
7. Test modules implemented in 1 and 6 together.
8. Implement module to combine the inputs for the best quality of results.
9. Test all modules together and make necessary changes.

# UML Diagrams

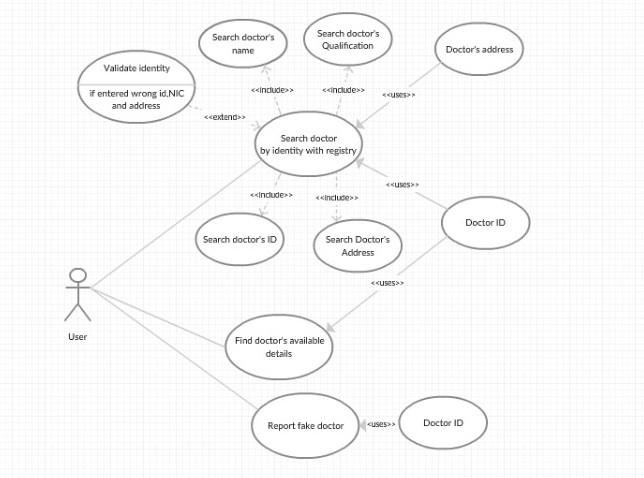
#### Use Case Diagram



##### BMI CALCULATOR



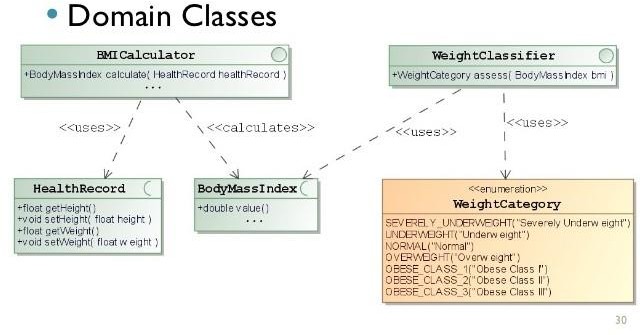
* + 1. **BLOOD DONATION & RECEIVER**



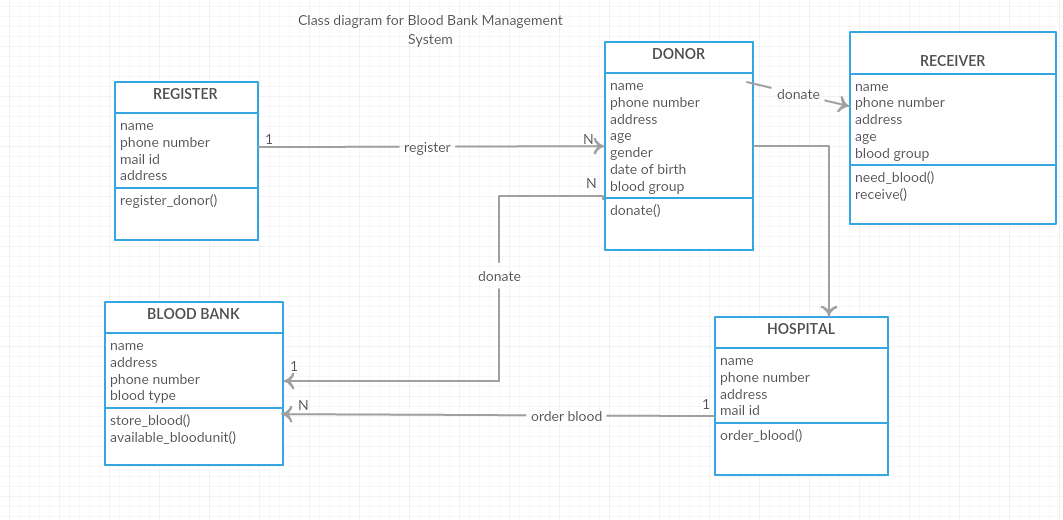
* + 1. **FIND ME A DOCTOR**



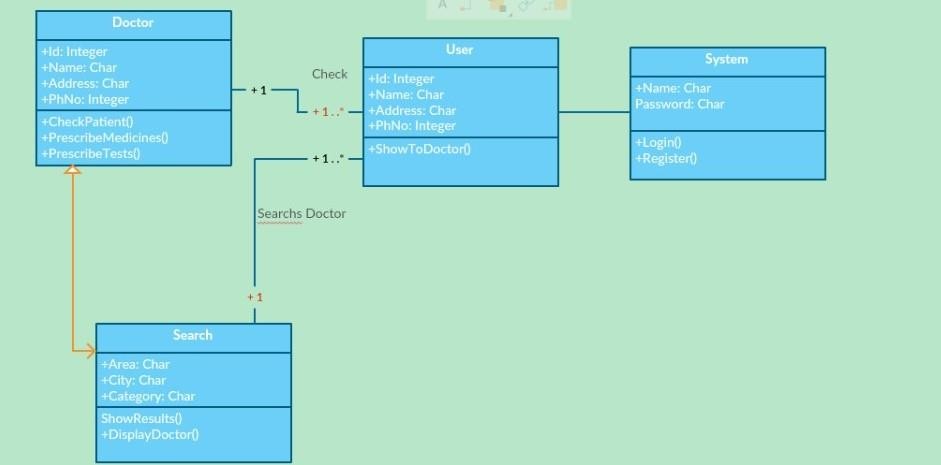
* + 1. **DISEASE PREDICTION**
  1. Class Diagram



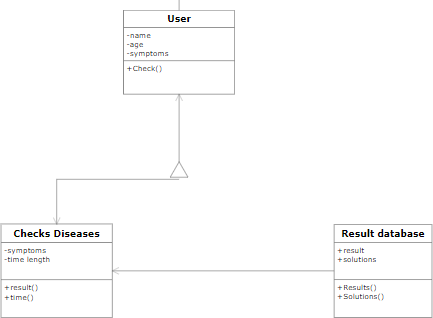
* + 1. **BMI CALCULATOR**



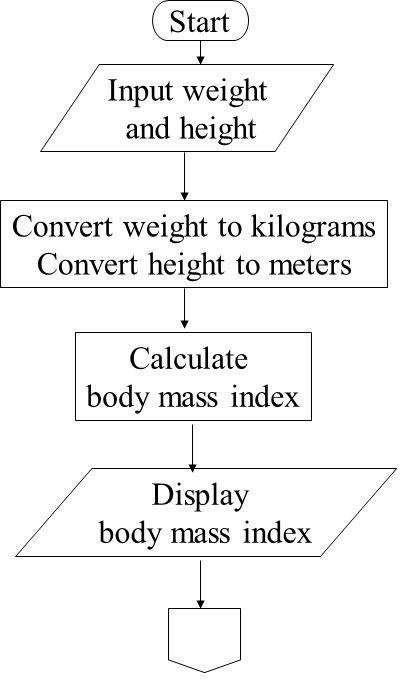
* + 1. **BLOOD DONATION & RECEIVER**



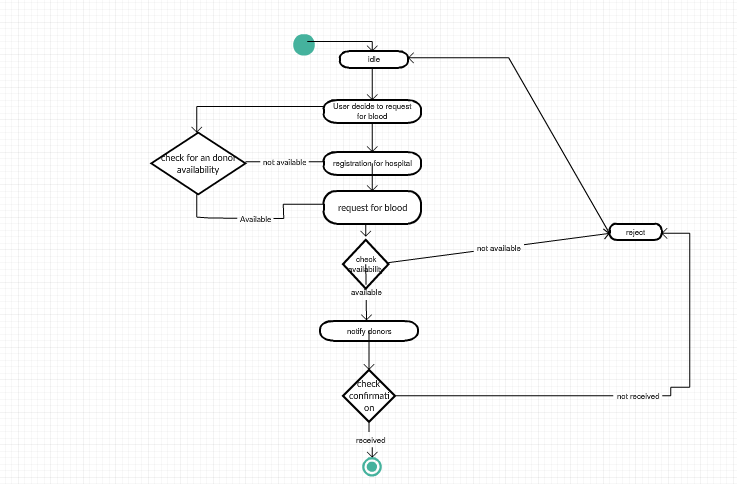
##### FIND ME A DOCTOR



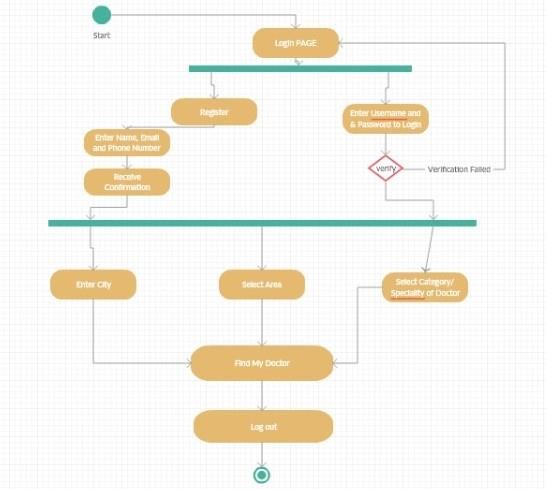
* + 1. **DISEASE PREDICTION**
  1. Activity Diagram



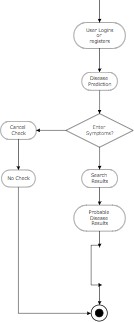
* + 1. **BMI CALCULATOR**



* + 1. **BLOOD DONATION & RECEIVER**



* + 1. **FIND ME A DOCTOR**



* + 1. **DISEASE PREDICTION**

# User interface design

#### Use Case Diagram

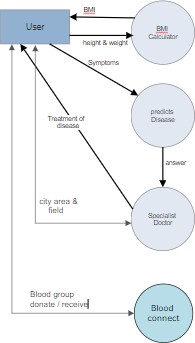


Figure 9.1.1 – Use Case Diagram

#### Data Dictionary

Login/signup table:

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Datatype | Constraint | Description |
| UserId | Int | Primary key | Id of the user |
| Password | varchar | Not null | Password of the user in the system |
| Phone no | varchar | Not Null | Phone number of the user |

Table 9.2.1 – Login/signup Table

BMI table:

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Datatype | Constraint | Description |
| UserId | Int | Primary key | Id of the User |
| Height | Int | Not null | height of the user |
| Weight | Int | Not Null | weight of the user |

Table 9.2.1 – BMI Table

Blood Donation & Receiver table:

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Datatype | Constraint | Description |
| UserId | Int | Primary key | Id of the user |
| Location | varchar | Not null | Location of the Donor |
| Blood Group | varchar | Not Null | Blood group of the Blood needed. |

Table 9.2.3 – Blood Donation & Receiver Table

Find a Doctor table:

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Datatype | Constraint | Description |
| UserId | Int | Primary key | Id of the user |
| City Area | varchar | Not null | City area Of the user |
| Category | varchar | Not Null | Category/field of the doctor needed should  be mentioned. |
| Contact details | varchar | Not Null | Contact Information of the user. |

Table 9.2.4 – Find a Doctor Table

Disease Prediction table:

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Datatype | Constraint | Description |
| Symptoms | varchar | Not Null | Symptoms of disease |
| Result | varchar |  | Disease result |

Table 9.2.5 Disease prediction Table

* 1. Interface Details

The user will be provided with a graphical user interface where they can do the following:

1. The user inputs a single input.
2. The user inputs multiple inputs that are to be the part of the same module.
3. The user initiates search or calculations.
4. The user pauses search or calculation.
5. The user resumes search or calculation.
6. The user stops search or calculation.
7. The user views the output generated by the system.

# Conclusion

Despite of multiple number of similar services available in the world, we didn’t find them enough capable of solving the multiple problems related to health in the society. Even though we cannot ensure perfect health and wellbeing of a person , we have provided them a platform to solve that problem. Nobody can fully be sure about the perfection of a certain software , but we can assure that we will be trying our level best to match with the user’s expectation and never will be depressed with a bad review instead we will be taking it as a stepping stone for improvement.

# References

## [www.1000projects.com](http://www.1000projects.com/)

* + [**www.creatly.com**](http://www.creatly.com/)
  + [**www.smartdraw.com**](http://www.smartdraw.com/)
  + [**www.w3schools.com**](http://www.w3schools.com/)
  + [**www.googlepatents.com**](http://www.googlepatents.com/)
  + [**https://www.phekb.org/sites/phenotype/files/ObesityAlgorithm\_c omplete\_v04.pdf**](https://www.phekb.org/sites/phenotype/files/ObesityAlgorithm_complete_v04.pdf)
  + [**https://ai.stackexchange.com/questions/1705/selecting-the-right- technique-to-predict-disease-from-symptoms**](https://ai.stackexchange.com/questions/1705/selecting-the-right-technique-to-predict-disease-from-symptoms)
  + [**http://thescipub.com/pdf/10.3844/jcssp.2010.548.552**](http://thescipub.com/pdf/10.3844/jcssp.2010.548.552)
  + [**https://www.healthitoutcomes.com/doc/algorithms-connect- patients-to-the-right-doctors-0001**](https://www.healthitoutcomes.com/doc/algorithms-connect-patients-to-the-right-doctors-0001)

Appendix

#### Design Engineering Canvas

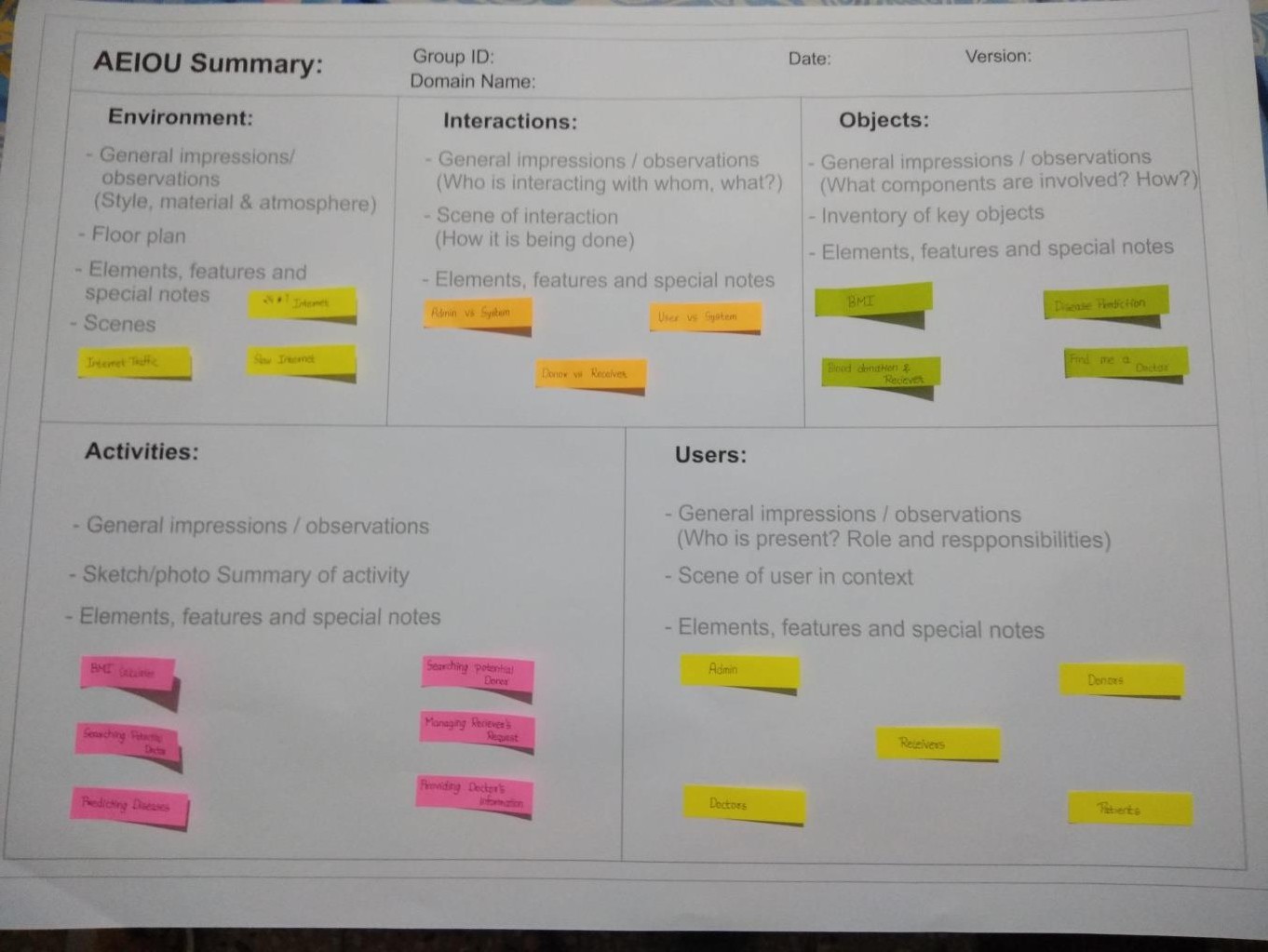


Figure A.1 – AEIOU Summary

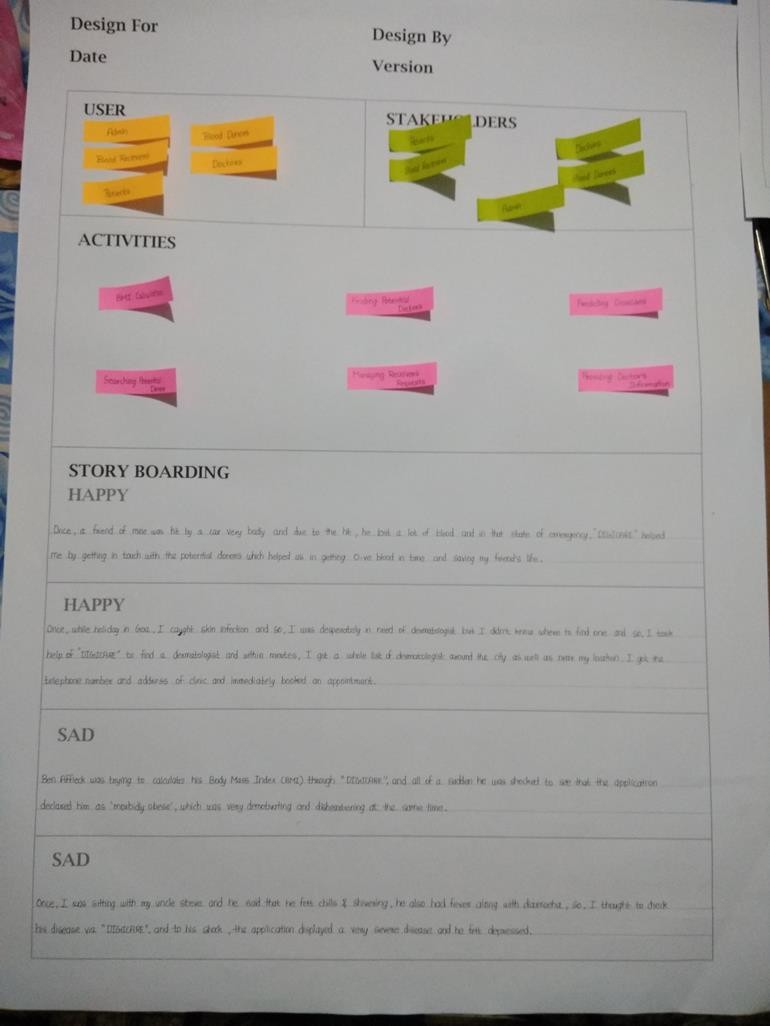


Figure A.2 – Empathy Map

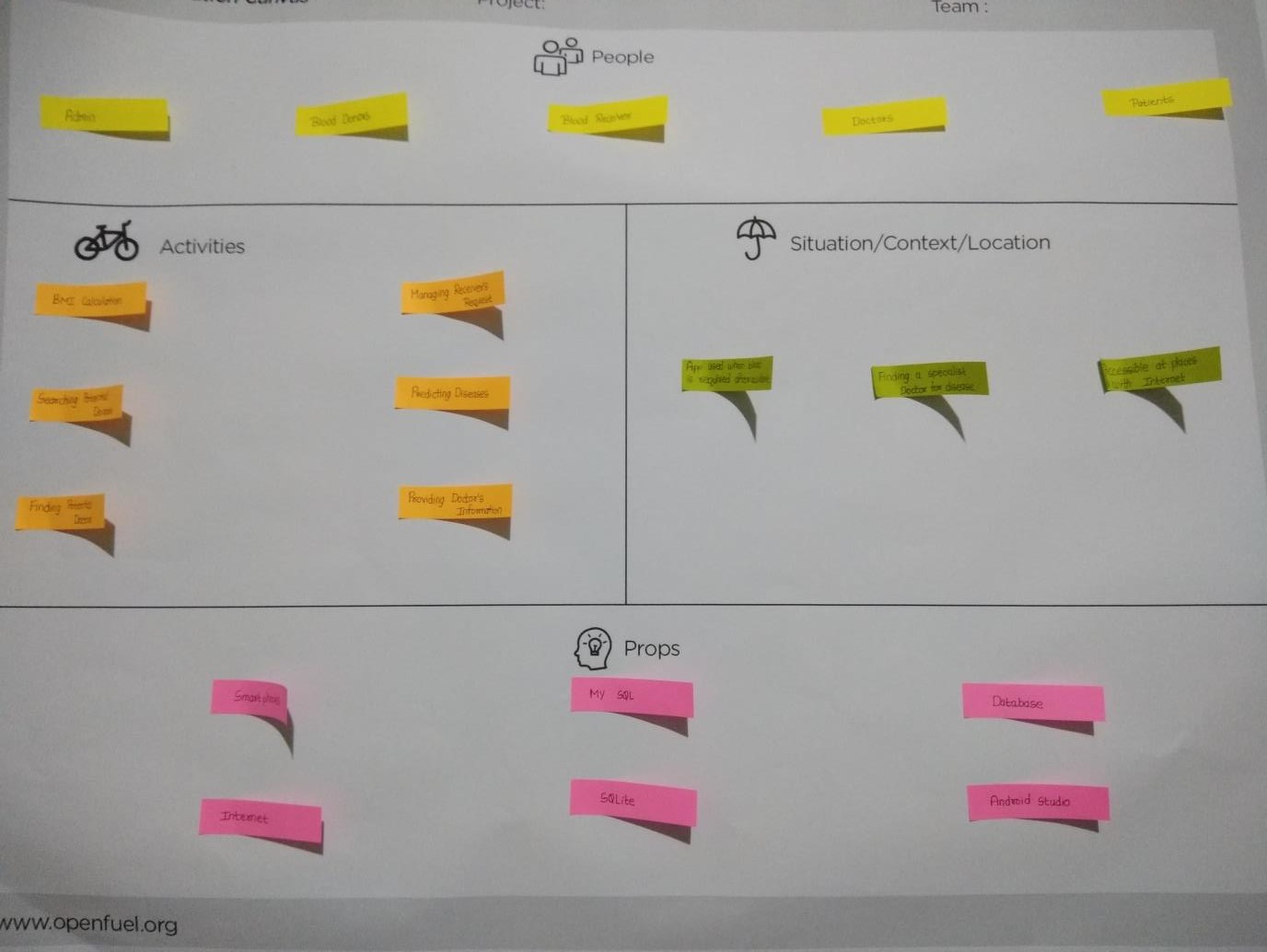


Figure A.3 – Ideation Canvas

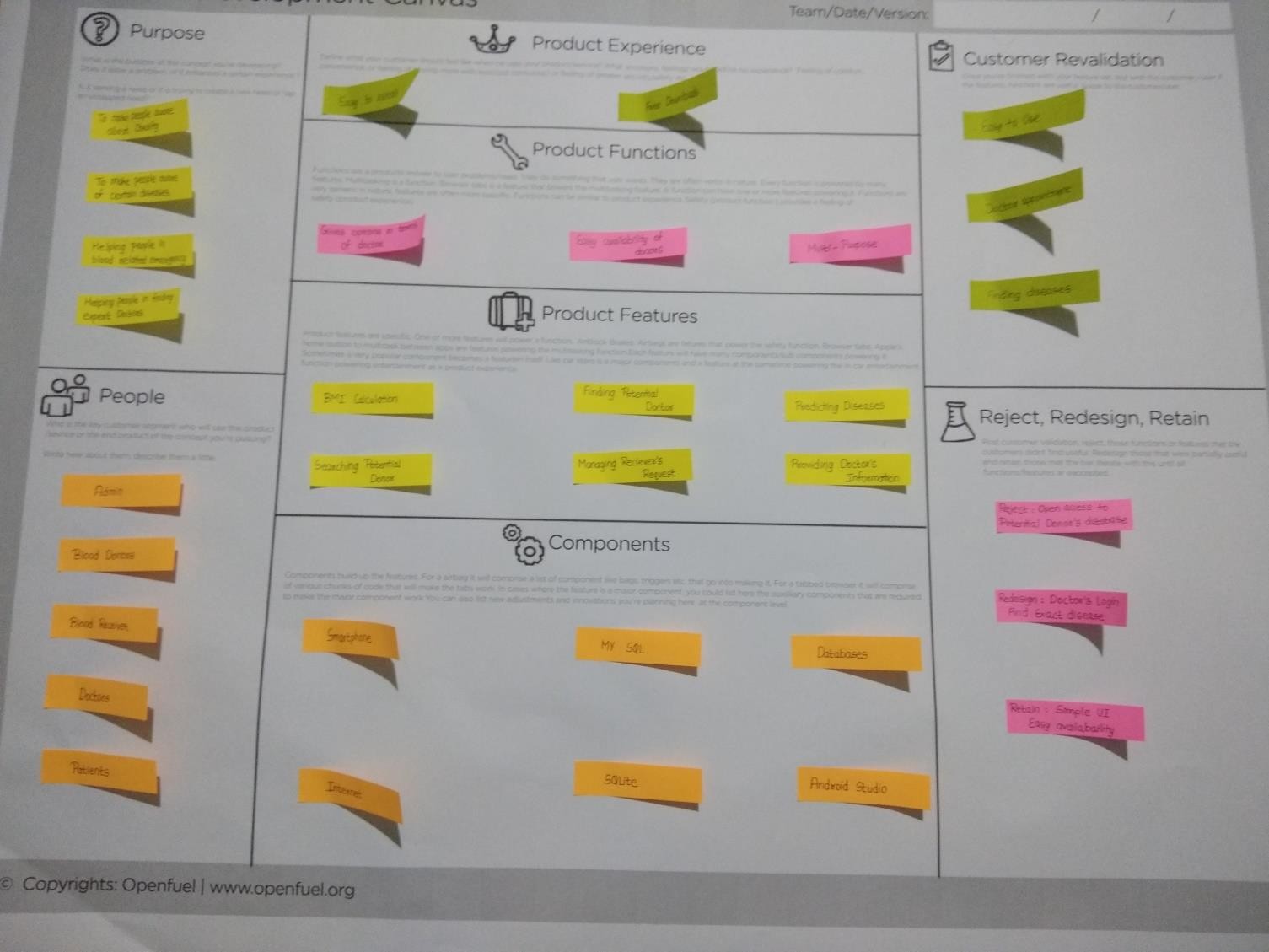
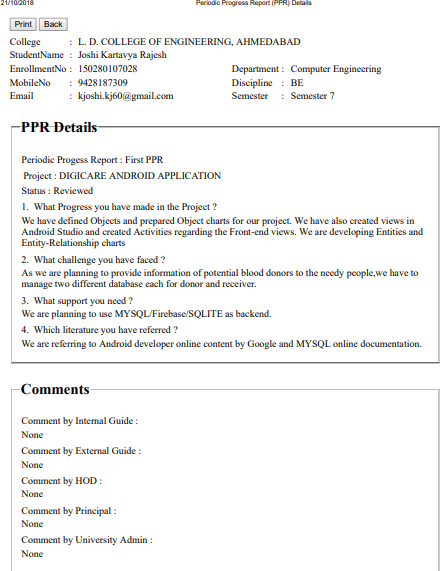
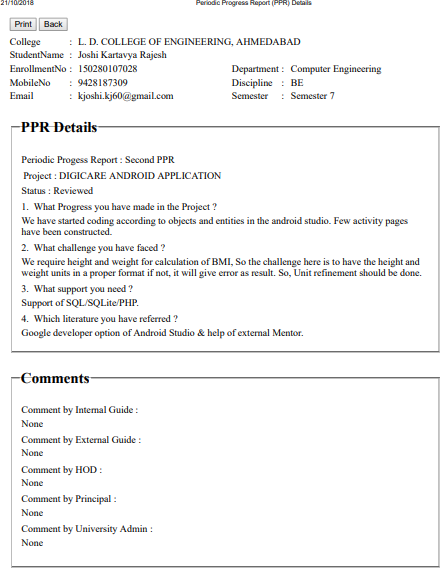


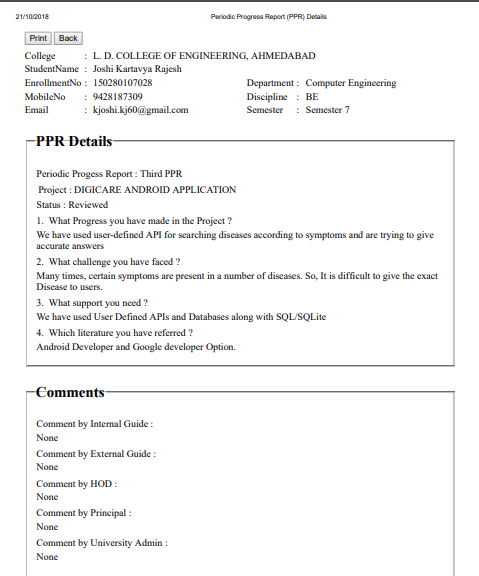
Figure A.4 – Product Development Canvas

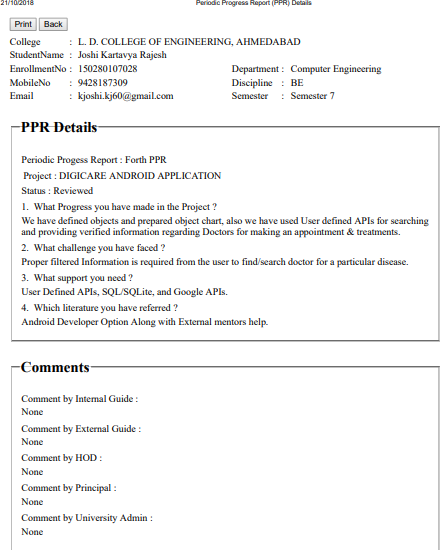
#### Periodic Progress Reports:

Kartavya Joshi PPRs

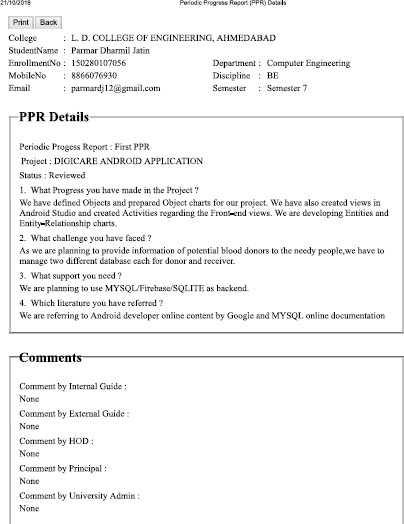


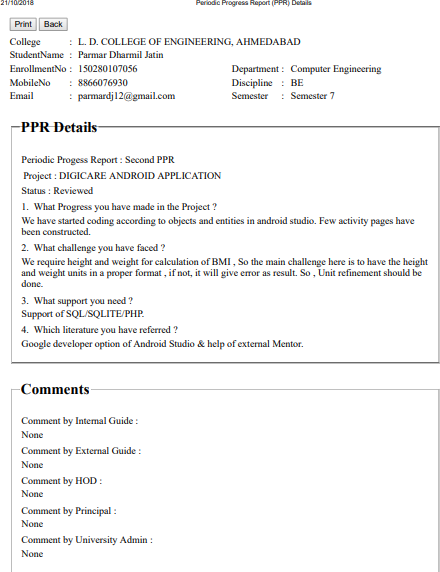


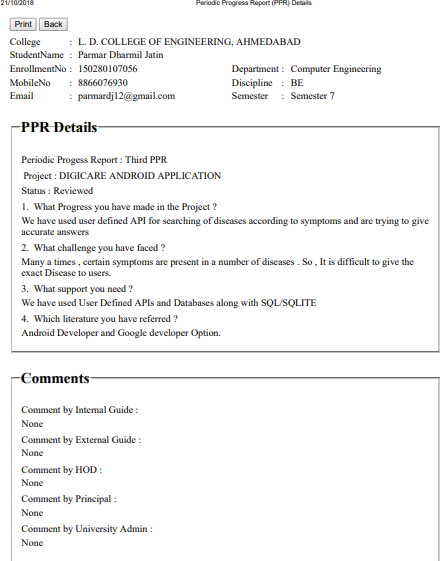


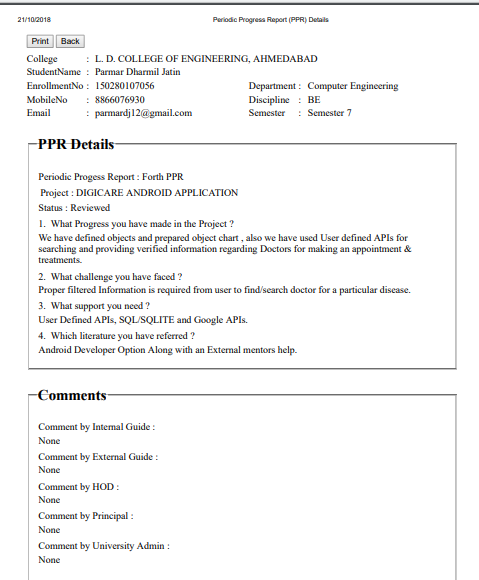


Dharmil Parmar PPRs



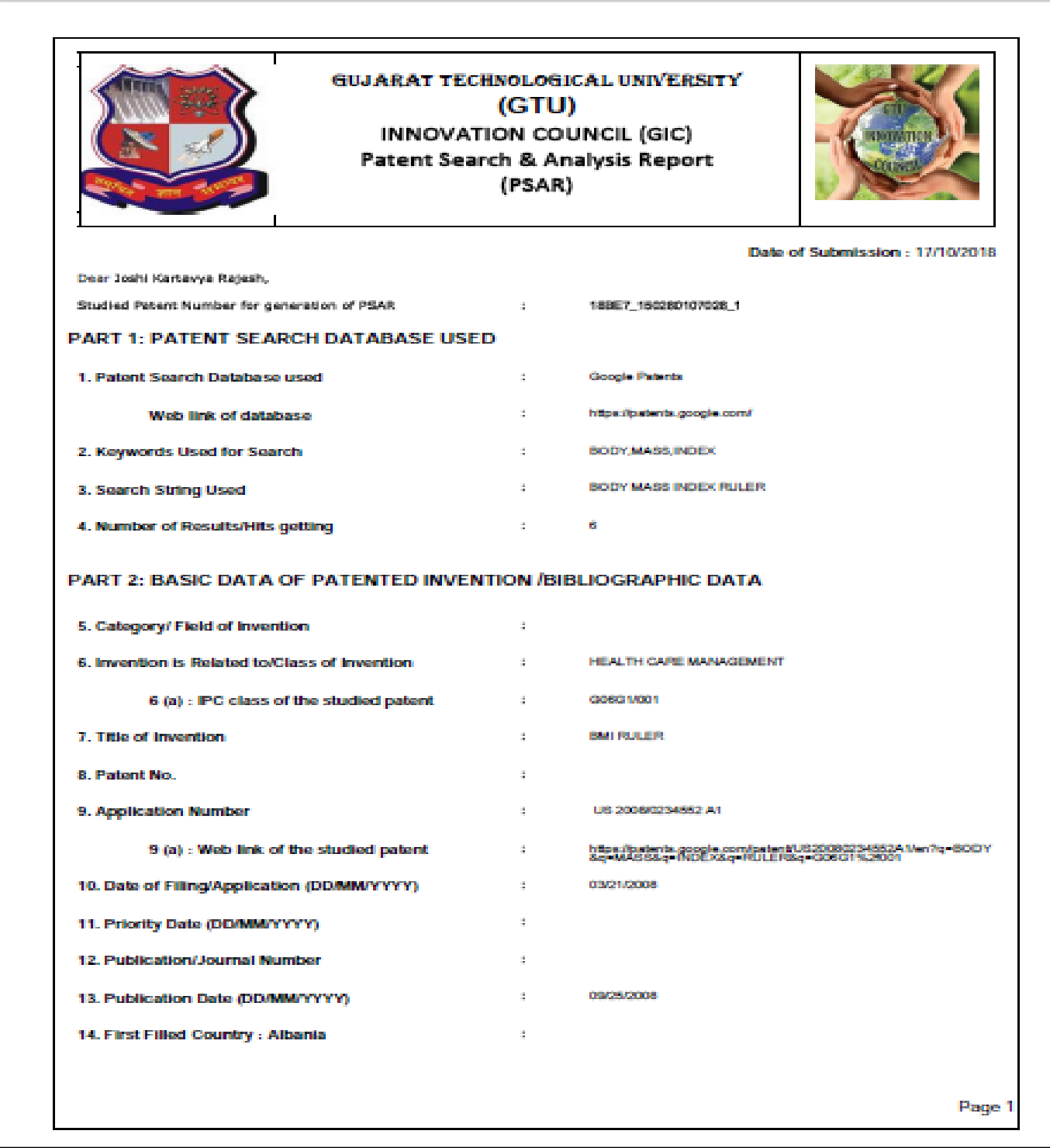


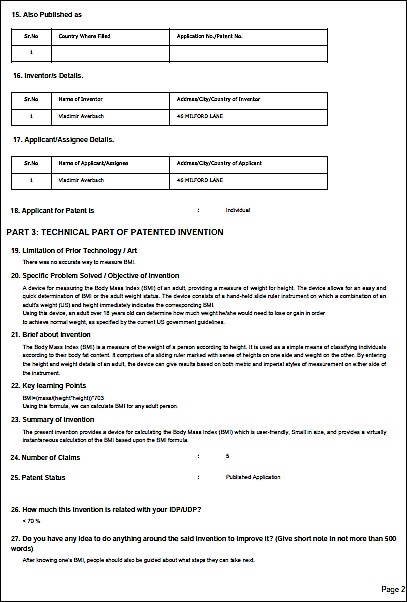


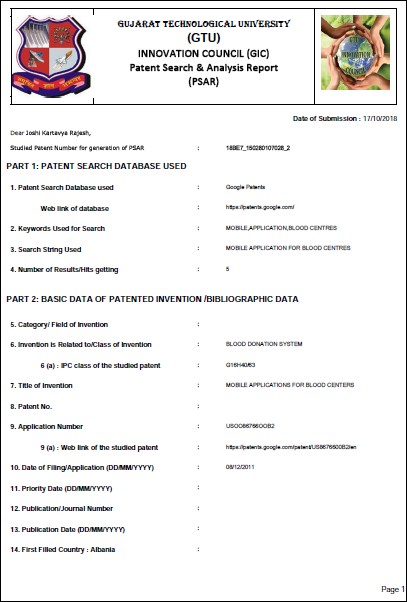


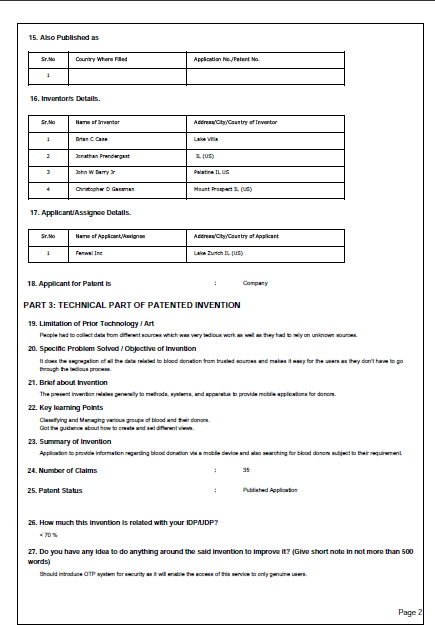
**PATENT SEARCH AND ANALYSIS REPORT**

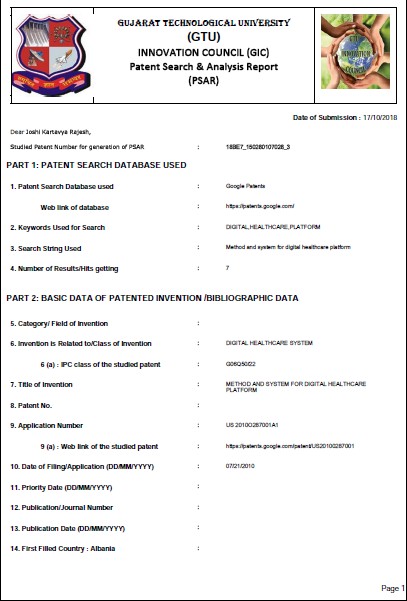
Kartavya Joshi -PSARs

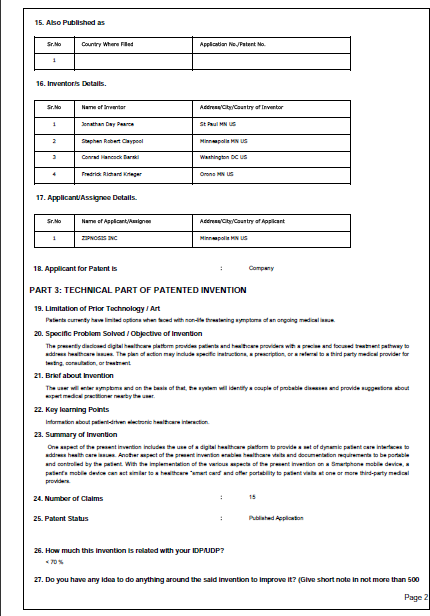


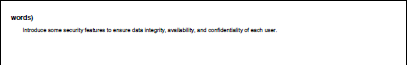


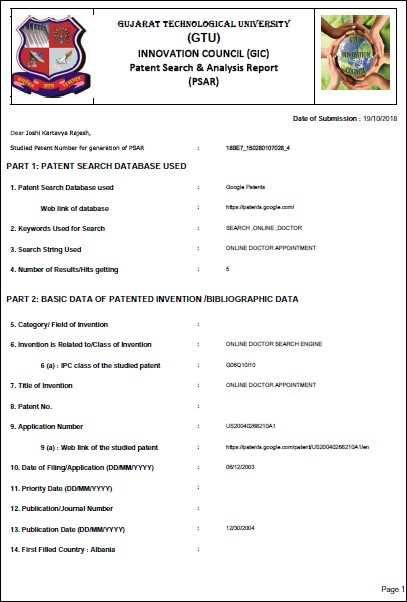


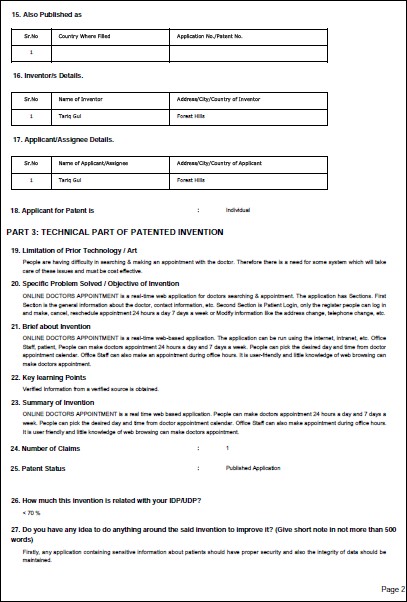


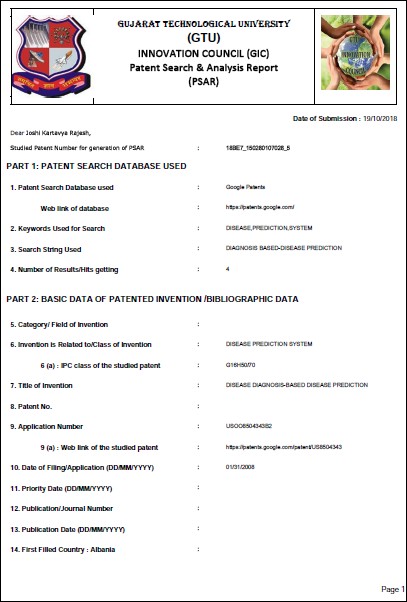


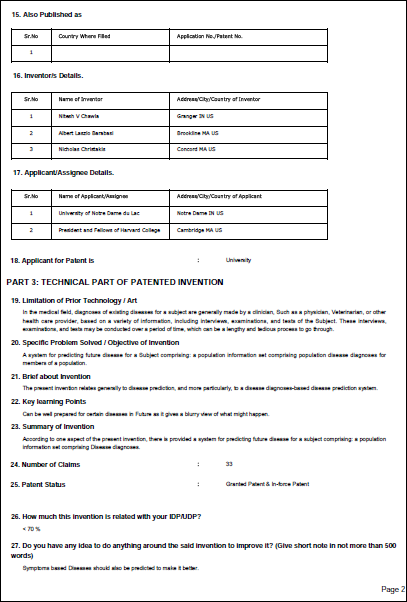




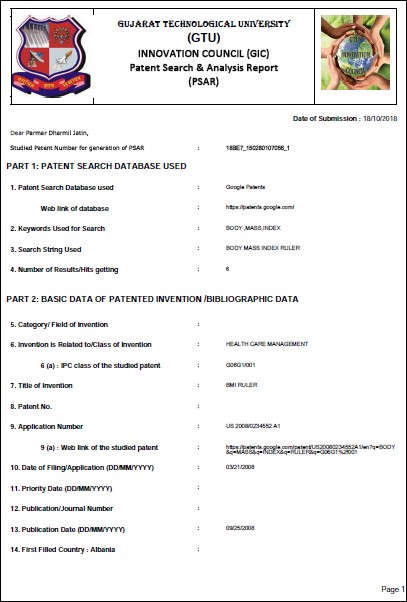


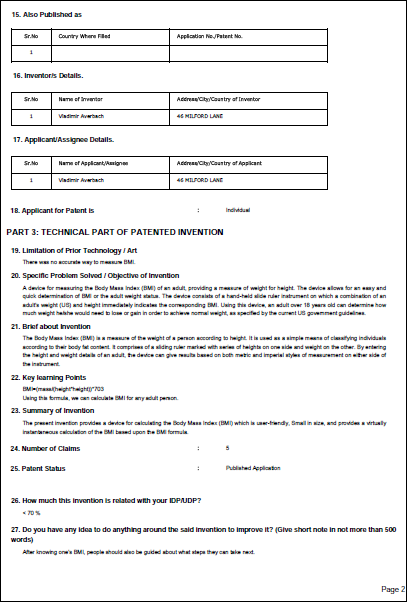


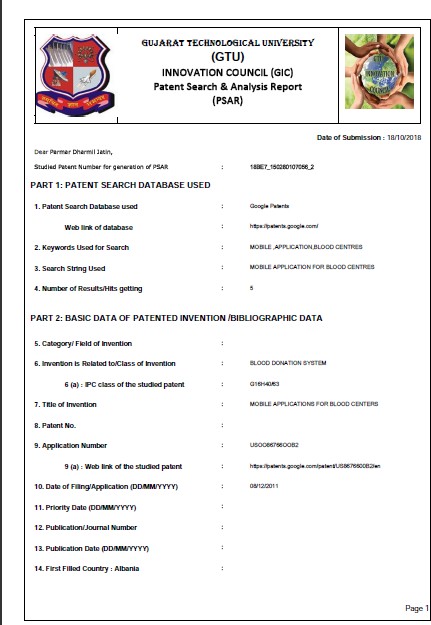


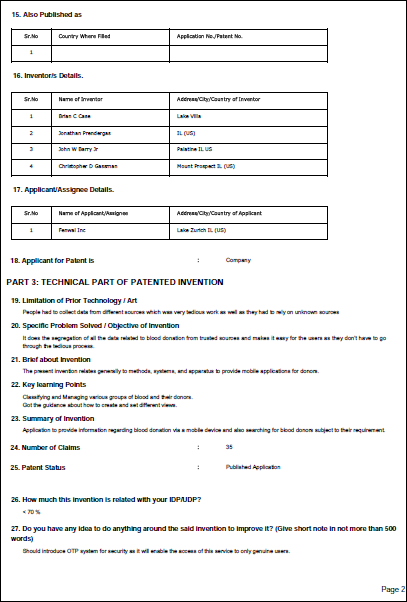


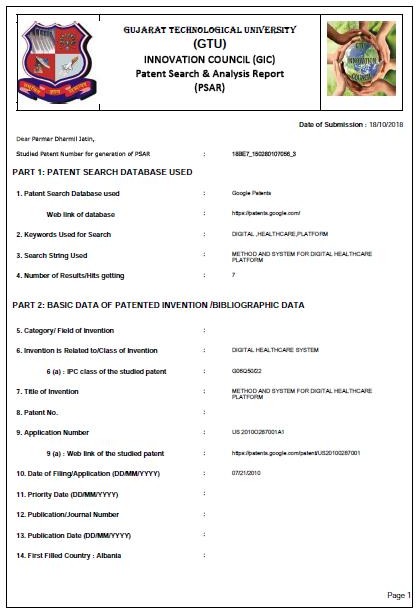
Dharmil Parmar - PSARs

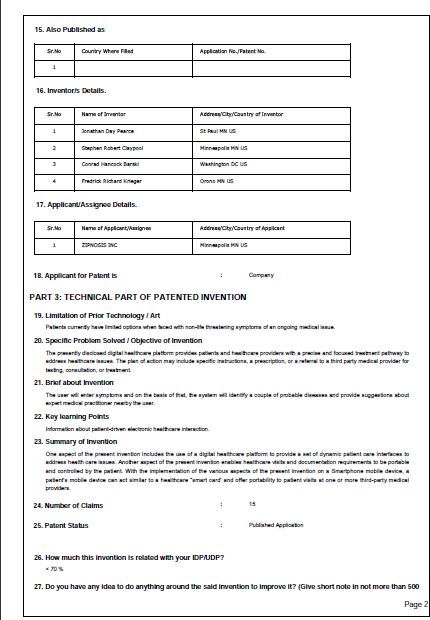


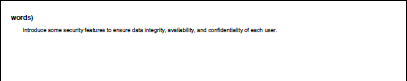


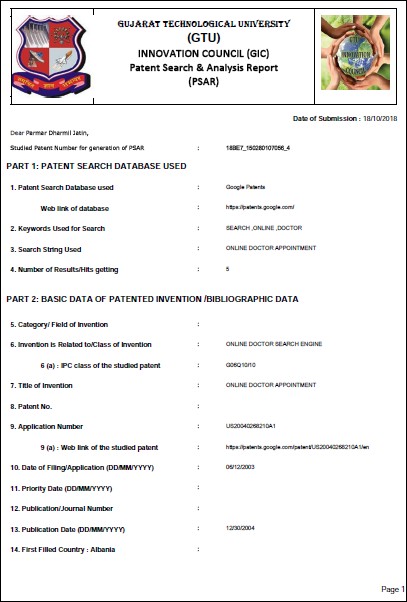


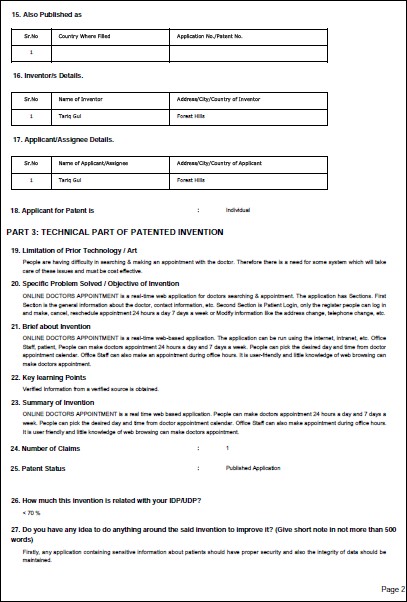


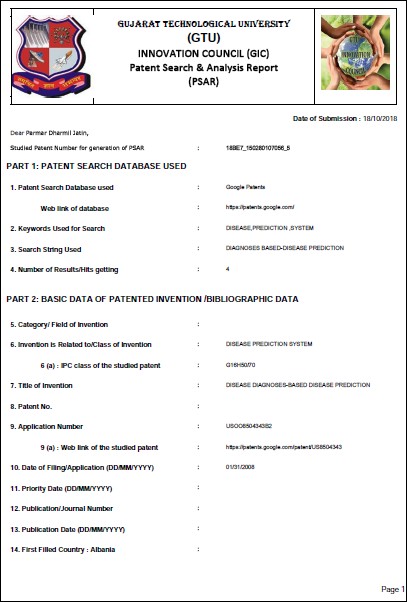


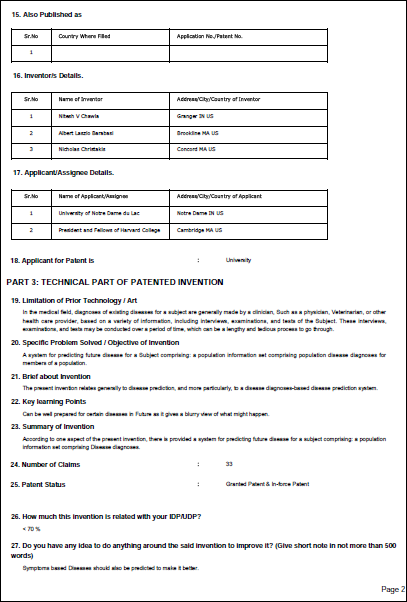












### PLAGIARISM CERTIFICATION

