A Project Report

On

Climate Change and its Impact on Health

Submitted in partial fulfilment of the requirement of

University of Mumbai

For the Degree of

Bachelor of Engineering

in

COMPUTER ENGINEERING

Submitted by

Atharva Ajit Agwekar
Prasun Alexander Philip
Aniket Sandhbor

Supervised by

Ms Rakhi Kalantri

&

Mr KV Subrahmanyam

Department of Computer Engineering

Fr. Conceicao Rodrigues Institute of Technology
Sector 9A, Vashi, Navi Mumbai - 400703

UNIVERSITY OF MUMBAI

2019-2020

APPROVAL SHEET

This is to certify that the project entitled

"Climate Change and its Impact on Health"

Submitted by

Atharva Ajit Agwekar	**101602**
Prasun Alexander Philip	**101645**
Aniket Sandhbor	**101674**

Supervisors : ____________________

Project Coordinator : ____________________

Examiners : 1. ____________________

2. ____________________

Head of Department : ____________________

Date :

Place :

Declaration

We declare that this written submission for B.E. Declaration entitled "**Climate Change and its Impact on Health**" represent our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declared that we have adhere to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any ideas / data / fact / source in our submission. We understand that any violation of the above will cause for disciplinary action by institute and also evoke penal action from the sources which have thus not been properly cited or from whom paper permission have not been taken when needed.

Project Group Members:

1. Atharva Ajit Agwekar, 101602

2. Prasun Alexander Philip, 101645

3. Aniket Sandhbor, 101674

Abstract

Weather patterns across different geographic locations have different seasons which results in different health issues for people of different age groups, genders, etc. for example, The winter season in Canada is much more severe than the winter experienced in India due to which the precautions to be taken and medication required will be different. Using this real-world evidence, the objective will be to analyze the data and predict the diseases or epidemic that might occur concerning a particular location and the associated weather patterns. This will help the individual user to make informed decisions and plan better. A system based on predictive analytics can be used to achieve such a task. The system will do this task using data previously collected as real-world evidence, combined with data from externally impacting forces. The proposed system will consist of an application where a user can create a login and upload health records. This health data will be juxtaposed along with input parameters provided by users like date, location, symptoms (if any) to generate a personalized output detailing the different probabilities of different diseases possible along with some recommendations for the same. As personal details of the user will also be factored, the results will always be idiosyncratic for each user.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

There is a myriad of different diseases that affect human health. Recording cases of such different diseases along with the weather parameters in which these diseases occur, it can be observed from analysis that some diseases can be directly correlated with the changes in weather patterns. Thus we can predict such diseases based on weather patterns. This can be done using a system based on predictive analytics. Predictive analytics is a type of data analytics which is capable of predicting the current trend or patterns, such as the inclination of a large number of customers to a particular product. Predictive analytics takes into account the past or historical information as well as the current information for forecasting any behavioral trends or for some cases like estimating the buying patterns of customers. This kind of analytics, for example, can help in generating revenue for companies since they would know what their customers are interested in buying. Also, these companies can offer discounts on such intriguing products, which would benefit the customers. Predictive analytics involve the use of different statistical or machine learning algorithms for estimating the probability or likelihood of an event[1]. Different variables are used by software applications performing predictive analytics that can be used to measure and analyze to predict the likely behavior of individuals, machinery or other entities. Another example would be of an insurance company that is likely to take into account potential driving safety variables consisting of age, gender, location, type of vehicle and driving record when pricing and issuing auto insurance policies.

1.2 Motivation

Some applications that already exist just gives few possible disease outbreak information[2]. Predicting the probable diseases based on some weather patterns will be greatly beneficial to all as it would help prevent many cases and prepare the user for the same. This project hopes to propose a system which has relatively high accuracy in predicting disease based on weather pattern as well as an application with a simple interface for users to interact. The system will also regularly update its database to improve its predictions. This would be done by collecting data from real-world evidence as well as some externally impacting forces and also by the predictions it has made. This data would be fed to a neural network which would be able to intelligently estimate the likelihood of occurrence of any disease[3]. The additional touch of symptoms experienced by the

Figure 1.1: Predictive Analytics

user would make recommendations personalized for every individual.

1.3 Aim and Objective

This project aims to build a system that will predict diseases based on weather patterns with relatively high accuracy based on real-world data and provide the user with an application that is simple to use with a user-friendly GUI. The GUI will hide the complex working of the system beneath it to provide an easy to operate application and a richer experience for the users. The major role of the application would be to collect the inputs provided by the user and feeding them to the engine which is based on neural networks since they tend to provide better accuracy[3]. The engine would do the processing task and return the results which would be then displayed in the application for the user. Also, the system will generate personalized output based on the user's medical history, as entered in the application by the user. The system would also provide a recommendation regarding remedies for the estimated output[4].

1.4 Report Outline

The project hopes to develop a system that will be able to predict disease based on weather patterns. To do so, it will use real world data along with externally impacting resources. It will also take into consideration the predictions it has estimated. The system will provide the user with an application with a simple to use interface to interact with the system where the user will get personalized results based on his/her input. The system will cover vast areas across the world and time up to 3 days in advance. It will use Neural Network models to make predictions.

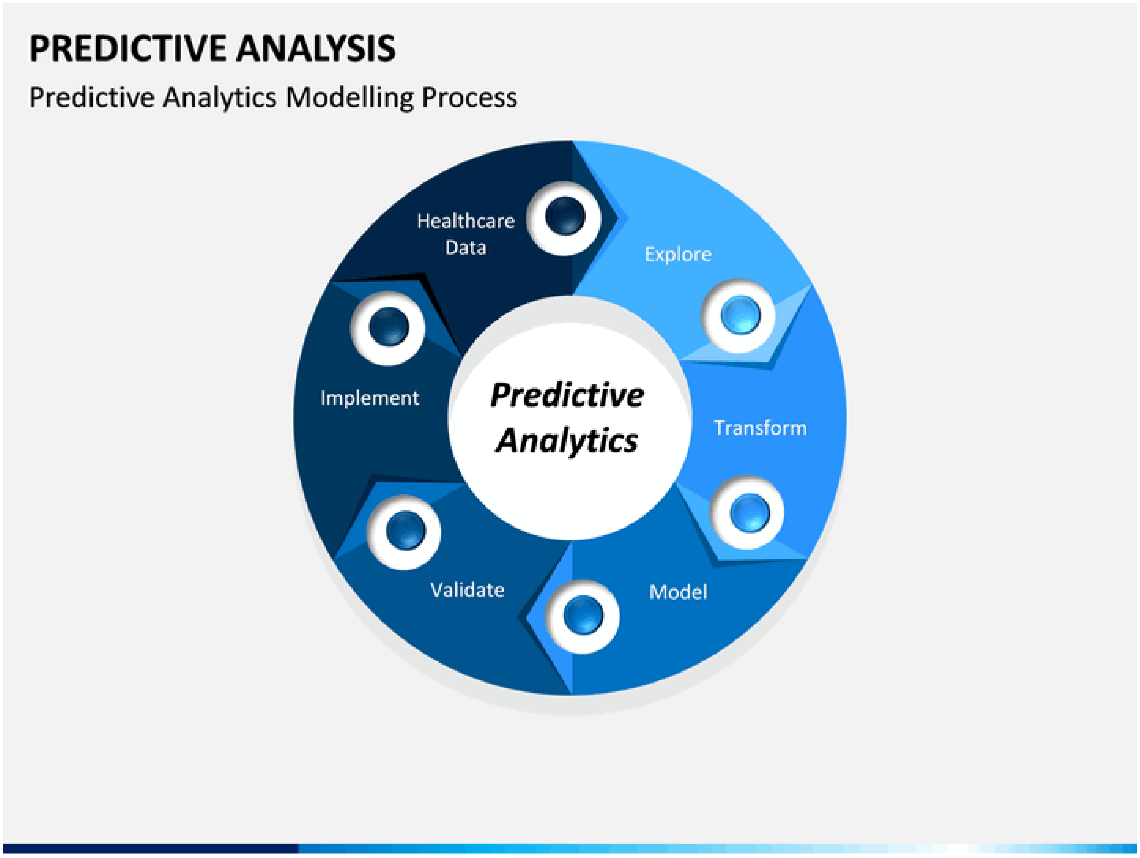

Figure 1.2: Predictive Analytics Process

Chapter 2

Study Of the Predictive Analytics System

2.1 Weather

2.1.1 Weather Patterns

It has been observed that the weather conditions often repeat themselves after a certain time interval. This repeating behavior of weather is referred to as weather patterns. The weather could stay the same for several days and then suddenly one day the weather changes and for example, sunny weather conditions for 2 weeks might suddenly become rainy for a week. This change in weather is referred to as weather pattern changes. The primary source that determines the weather pattern in the middle latitudes is the upper-level flow pattern. An examination of the jet stream winds shows regions likely with below normal temperatures and above normal temperatures. If a forecast area repeatedly stays within a trough the weather will tend to be cooler than normal. If a forecast area repeatedly stays within a ridge the weather will tend to be warmer than normal. Troughs and ridges can become locked in position for many days or weeks in a row. When this happens, some regions will have droughts while other areas have flooding. Some regions will continually have warm weather while other areas are much cooler weather. Such insights can be drawn by capturing information which would help in finding weather patterns.

2.1.2 Impact of Weather on Health

Weather patterns affect the health of individuals across the world[5]. Unexpected weather changes bring about allergies, flu's, viral infections etc. It can also be attributed to disease outbreaks in certain regions. For example, heavy rainfall can increase water-borne diseases like cholera, typhoid etc whereas a dry climate will result in the spread of airborne diseases like influenza, SAR etc. Abrupt changes in weather can occur at any time. Some people are affected more severely than others due to the change in weather patterns. Thus study of weather patterns will be greatly beneficial as it would help people better prepare themselves as well as take preventive steps for a dire situation. Given that the impacts of climate change are projected to increase over the next century, certain existing health threats will intensify and new health threats may emerge[5]. Connecting our understanding of how climate is changing with an understanding of how those changes may affect human health can inform decisions about mitigating (reducing) the amount of future climate change, suggest priorities for protecting public health, and help identify research needs.

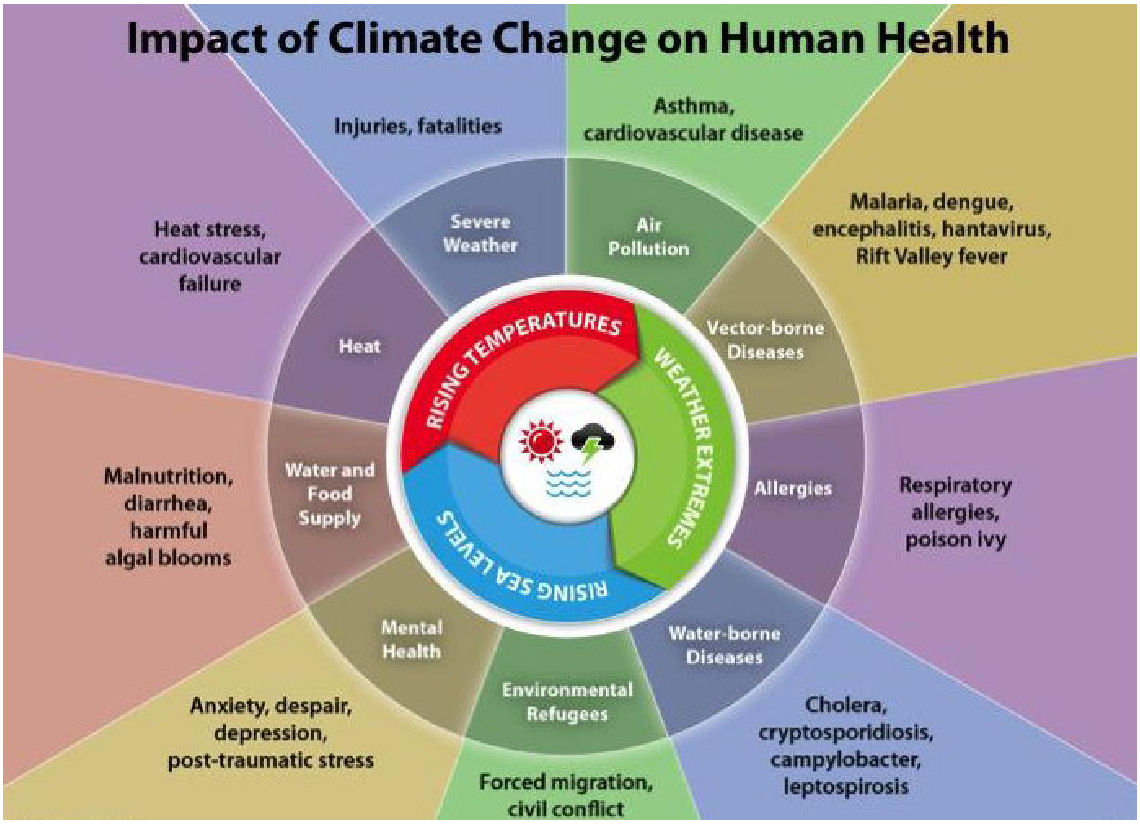

Figure 2.1: Climate Change & it's Health Impacts

2.2 Available Techniques

2.2.1 Types Of Predictive Analytics

Predictive analytics is a technique which is used to make predictions about future occurring based on the data that we presently have as well as data from the past i.e. historical data. Predictive analytics makes use of machine learning techniques or other statistical methods for predicting the future outcomes.Predictive Analytics can be applied to any field.

1. Predictive model:
Predictive modelling uses predictive models to analyze the relationship between the specific performance of a unit in a sample and one or more known attributes or features of the unit. The objective of the model is to assess the likelihood that a similar unit in a different sample will exhibit the specific performance. This category encompasses models in many areas, such as marketing, where they seek out subtle data patterns to answer questions about customer performance, or fraud detection models.

2. Descriptive model:
Descriptive models quantify relationships in data in a way that is often used to classify customers or prospects into groups. Unlike predictive models that focus on predicting a single customer behavior (such as credit risk), descriptive models identify many different relationships between customers or products. Descriptive models do not rank-order customers by their likelihood of taking a particular action the way predictive models do. Instead, descriptive models can be used, for example, to categorize customers by their product preferences and life stage.

3. Decision model:
Decision models describe the relationship between all the elements of a decision—the known data (including results of predictive models), the decision, and the forecast results of the decision—in order to predict the results of decisions involving many variables. These models can be used in optimization, maximizing certain outcomes while minimizing others. Decision models are generally used to develop decision logic or a set of business rules that will produce the desired action for every customer or circumstance.

2.2.2 Artificial Neural Networks

These are computing system inspired by neural networks present in the animal brain. These systems can learn from their past mistakes and try to improve their performance. An Artificial Neural network tries to reduce the error it is getting. There are three main layers in a Neural Network - input layer, hidden layer and output layer. Input layer deals with the input provided by the user to the Neural Network. Output layer deals with the output that is expected from the Neural Network. The hidden layer is where the actual computation occurs. There can be any number of layers in the hidden layer. More layers result in more precise outputs but are computationally heavy. They use different activation functions to activate a neuron-like Sigmoid function, ReLu etc. There are different types of Neural Networks developed to cater to different needs. For example Convolution Neural Network is ideal for recognising pictures, Recurrent Neural Networks for developing language models etc.

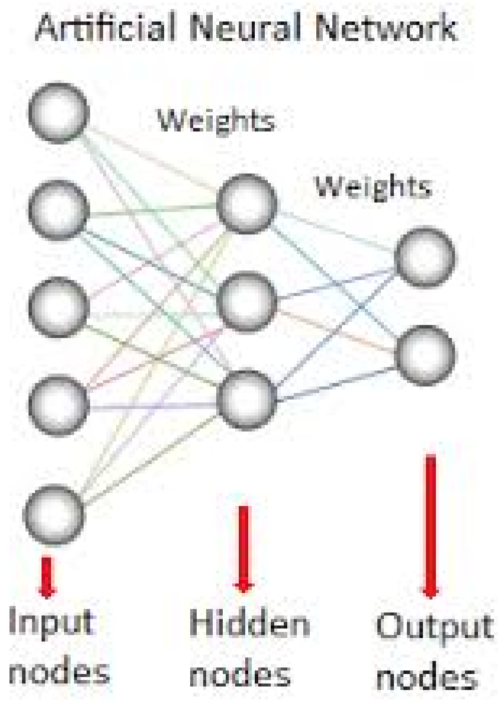

Figure 2.2: Artificial Neural Network

2.3 Related Works

Related work (expected at least 5 research papers brief description with comparison)

2.3.1 Predictive Analytics Using Data Mining Technique

Hina Gulati
Computer Science and Engineering Amity University, Noida, INDIA

The process mentioned in this paper consists of three steps namely Data Gathering, Pre-processing and Data Mining. Data Mining is the step where different mining algorithms are applied to predict student failure. This system has been implemented using WEKA tool which has various options for rule-based classification algorithms like Jrip, NNge, conjunctive rule, DTNB, PART, etc. Results obtained from such models can help teachers and management to identify the problem areas and reasons that affect dropout the most.

2.3.2 Predictive Big Data Analytics in Healthcare

A.Rishika Reddy
Computer Science and Engineering Kakatiya Institute of Technology & Science, Warangal, INDIA
This paper provides an admiring impact that would provide meaningful information about the approaches to adopt Big Data Analytics in the healthcare sector. This paper also discussed the tools that are used for handling Big Data and provides enough solutions to face the challenges in healthcare data with the help of Big Data Analytics. With rapid advances in applications of healthcare data, there is an acute need for Big Data Analytics.

2.3.3 Predictive Analytics in Data Science for Business Intelligence Solutions

Parth Wazurkar
Dept of CSE, Indian Institute of Information IIIT, Nagpur, Maharashtra, INDIA
Predictive Analytics in business can help to identify the needs of individual customers and also provide personalized recommendations. This helps a company to popularize its products among customers. This paper also talks about predictive analytics in healthcare industries wherein models based on individual health costs can be used to improve the quality of healthcare. Different algorithms can be used to determine the most predictive model among different models. The results of such a model can be visualized to draw different insights.

2.3.4 Weather and Health Symptoms

Mihye Lee
Graduate School of Public Health, Japan
This paper shows how Weather affects the daily lives of individuals. However, its health effects have not been fully elucidated. It may lead to physical symptoms and/or influence mental health. They used daily reports on health symptoms from 4548 individuals followed for one month in October of 2013, randomly sampled from the entirety of Japan. The logistic mixed-effects model with a random intercept for each individual was applied to evaluate the effect of temperature and humidity on physical symptoms. Stratified analyses were conducted to compare weather effects by sex and age group. The weather was associated with various physical symptoms.

Women seemed to be more sensitive to weather conditions associated with physical symptoms, especially higher humidity and lower temperature.

2.3.5 Health symptoms in relation to temperature, humidity, and self-reported perceptions of climate in New York City residential environments

Ashlinn Quinn
Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY 10032, USA.
In this paper, monitoring has been conducted on temperature and humidity inside homes even though these conditions may be relevant to health outcomes. Here, they investigate associations between measured temperature and humidity, perceptions of indoor environmental conditions, and health symptoms in a sample of New York City apartments. After monitoring they measured the temperature and humidity in 40 New York City apartments during summer and winter seasons, and collected survey data from the households' residents. Health outcomes of interest were: sleep quality, symptoms of heat illness (summer season), symptoms of respiratory viral infection (winter season). Perceptions of indoor temperature were significantly associated with measured temperature in both the summer and winter, with a stronger association in the summer season. Sleep quality was inversely related to measured and perceived indoor temperature in the summer season only.

2.3.6 Big Data Analytics in Medicine and Healthcare

Blagoj Ristevski
St. Kliment Ohridski University – Bitola, Faculty of Information and Communication Technologies, ul. Partizanska bb, 7000 Bitola, Republic of Macedonia
Ming Chen
Department of Bioinformatics, College of Life Sciences, Zhejiang University Zijingang Campus, Hangzhou, P.R. China
This paper mentions the use of Big Data Analytics in Healthcare along with the challenges, privacy and security of data in the healthcare field. This paper speaks about the integration and analysis of complex data in large amounts such as various omics data (genomics, epigenomics, etc) for analysis in medicine and healthcare. It also speaks about the use of models based on Electronic Health Records (EHR) for providing personalized and

preventive medicine. This would prove beneficial for the patients, clinicians and health policymakers. This paper also speaks about using Data Mining techniques on EHRs which could help identify the different association rules in EHRs for disease monitoring and health-based trends.

2.3.7 Predictive Analytics to Prevent and Control Chronic Diseases

Kumari Deepika
M.Tech Student, Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore-560054
Dr. S. Seema
Associate Professor, Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore-560054
This paper presents effective mechanisms have been used for chronic disease prediction by mining the data containing historical health records. The experiment was conducted using two separate datasets from the UCI machine learning repository for the diagnosis of heart disease and diabetes. Different classifiers such as Naïve Bayes, SVM, Decision Tree, and Artificial Neural Networks were used. This experiment found that SVM gives the highest accuracy rate of 95.556% in case of heart disease and case of diabetes Naïve Bayes classifier gives the highest accuracy of 73.588%. The accuracy was determined based on four cases True positive, True Negative, False Positive, False Negative. Fitting predictive models like those utilized as a part of this paper can be utilized to create individual/clinical decision support systems, to keep up wellness or enhance the administration of chronic diseases, for example, heart disease and diabetes.

2.3.8 An Interactive Predictive System for Weather Forecasting

Ayham Omary
Engineering and Architecture faculty, Umm Al-Qura University, Mecca, KSA
This paper presents a view for building a weather prediction model based on the weather of Jordan. Such models are usually complex and will take significant time and resources to accomplish. This model will be based on HIRLAM and ALADIN models that will be used for weather prediction models focused on Jordan geographic area. The ALADIN model, developed in an international cooperation led by Météo France, is operationally used for weather prediction. The grid step of the model is 12 km. The integration domain covers a major part of Europe. High-Resolution Limited

Area Model (HIRLAM) model is primarily a numerical weather prediction model with parameterizations that are aimed at the short-range weather forecast only. This paper mentions the use of an information retrieval robot application that can collect all-weather and precipitation related information about Jordan for years of no less than 5 years. Using this information, data mining and AI algorithms can be applied to build a future prediction and forecasting for precipitation based on studying the historical data.

Chapter 3

Proposed System

3.1 Problem Statement

Use the Climate data to forecast what impact it will create on the health of people across the globe. Weather changes bring unexpected allergies, flu's, viral infections, and many other health issues. People in different age groups either homemakers, senior citizens, people traveling for business or holidays will benefit from this. When this climate data is juxtaposed with personal health information, it can foretell what changes to expect and hence preventive steps can be taken.

3.2 Scope

The goal of this project is to relate multiple diseases across multiple regions at any time with personalized results. This project aims to achieve the same through a mobile/web application wherein an individual can see personalized predictions based on the current location of the user or any location he/she wishes and symptoms the user might be experiencing which can be specified by typing the symptoms and selecting symptoms from the drop-down menu that appears. The application can also get information from uploaded medical reports. It will be able to predict the possible impact on health up to one week in advance and suggest suitable remedies for the same.

3.3 Proposed System

This project aims to build a system that will predict diseases based on weather patterns with relatively high accuracy based on real-world data and provide the user with an application that is simple to use with a user-friendly GUI. The GUI will hide the complex working of the system beneath it to provide an easy to operate application and a richer experience for the users. The major role of the application would be to collect the inputs provided by the user and feeding them to the engine which is based on neural networks since they tend to provide better accuracy[3]. The engine would do the processing task and return the results which would be then displayed in the application for the user. Also, the system will generate personalized output based on the user's medical history, as entered in the application by the user. The system would also provide a recommendation regarding remedies for the estimated output[4].

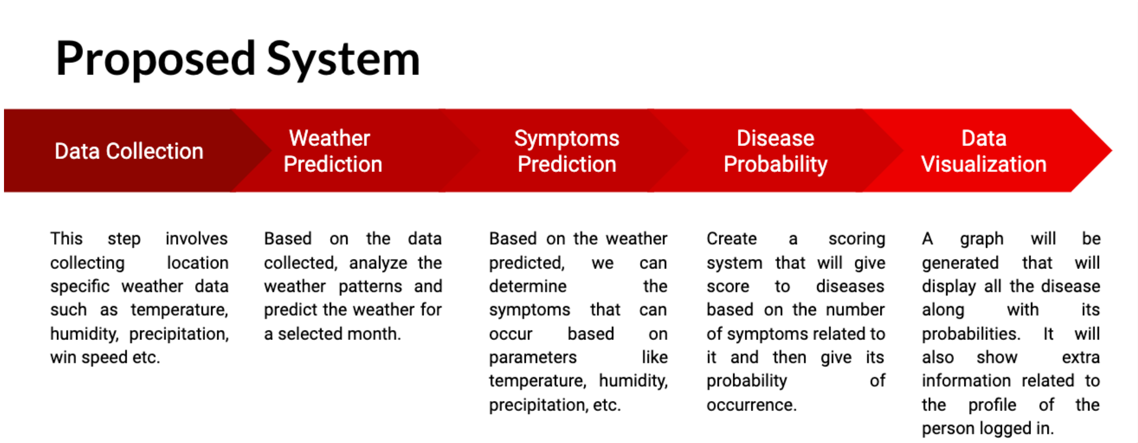

Figure 3.1: Proposed System

Chapter 4

Design Of the System

4.1 Requirement Engineering

4.1.1 Requirement Elicitation

The different actors of this project are the users, who will be using an applications, an expert panel on the admin side and a database manager. Scenarios where this application can be used include travel where users can determine if their destination is safe from to travel, Government Agencies to keep medical supply stock updated, Medical Institution who can better prepare for any potential outbreak etc. The different methods that can be used to achieve such application include interviews with the client to know about their requirements, brainstorming session with developers etc.

4.1.2 Software life cycle model

The Software life cycle model that was used for this project is Agile SDLC Model. The Agile model of SDLC divides the complete project into a number of small builds. These builds are worked upon by the developers and provided to the client in iterations. In every iteration, the development team undergoes through planning, analysing the customer requirements, designing, coding and testing. This model is basically a mix of the iterative and incremental SDLC model since each new iteration is incremental to the previous as the new build adds more features to already built features in the prior iteration. In this way, the last iteration will provide all the features that are expected by a client.

We divided the project into different parts - the Neural Network models for prediction of diseases, the server for running the model and connecting to the android application for returning the results from the server and finally developing the android application. The android application was further split up into the user interface and admin interface. We started out first with the prediction model, for which we had to collect data about weather parameters and their linkage to different diseases. The requirement of the model was that given a set of values for the different weather parameters, the model should predict the diseases which are possible based on these values.We initially started prediction with simpler models like Linear Regression Model, K-Nearest Neighbour(KNN), etc. After brainstorming the different models we finally concluded that a Neural Network model best served our needs. Similarly, we looked out for the requirements of each part that we had divided the project into and accordingly planned and designed those parts. After each part was completed it was thoroughly analyzed,

Figure 4.1: Agile Model

tested and improvements were made. Also there were regular discussions to better understand client requirements. Thus after many discussions and upgrades the the android application was brought to its completion.

4.1.3 Requirement Analysis

4.1.3.1 Use Case Diagram

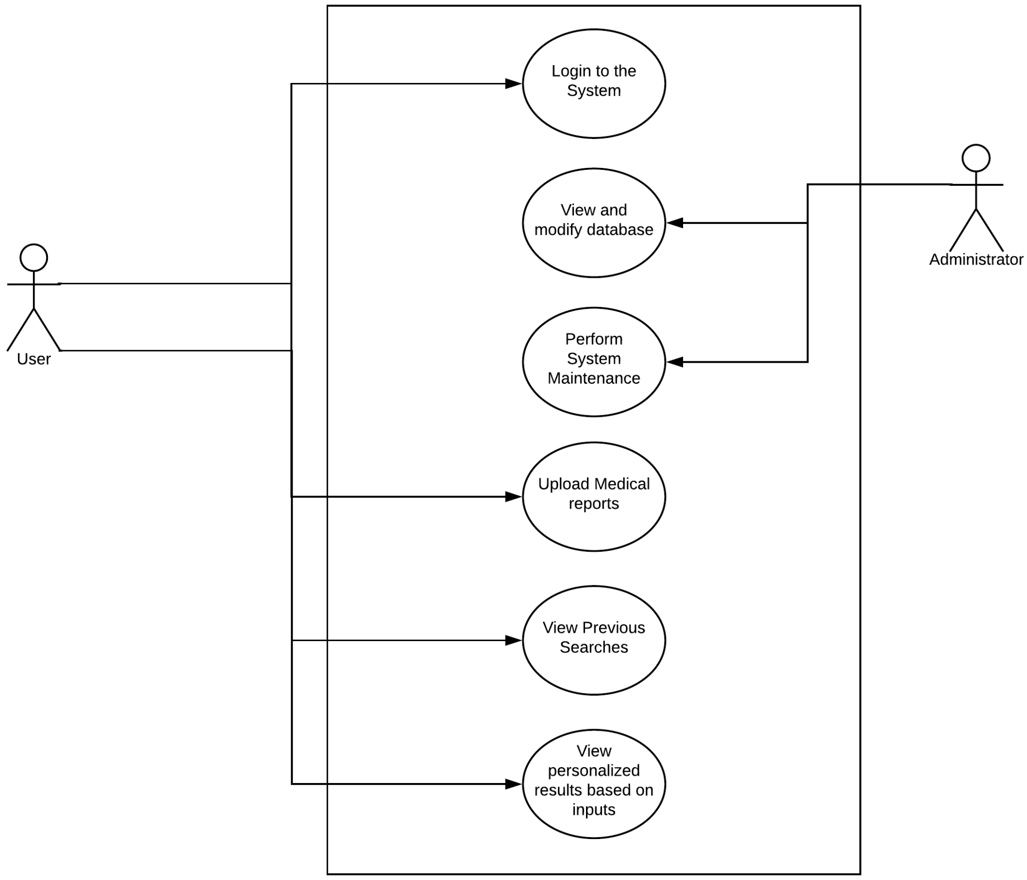

Figure 4.2: Use Case

Figure 4.2 depicts how different actors interact with the system to achieve a goal using a list of actions. In this diagram, we have an user who can perform a variety of functions by interacting with the application. An administrator monitors the system and performs overall system maintenance.

4.1.3.2 Activity Diagram

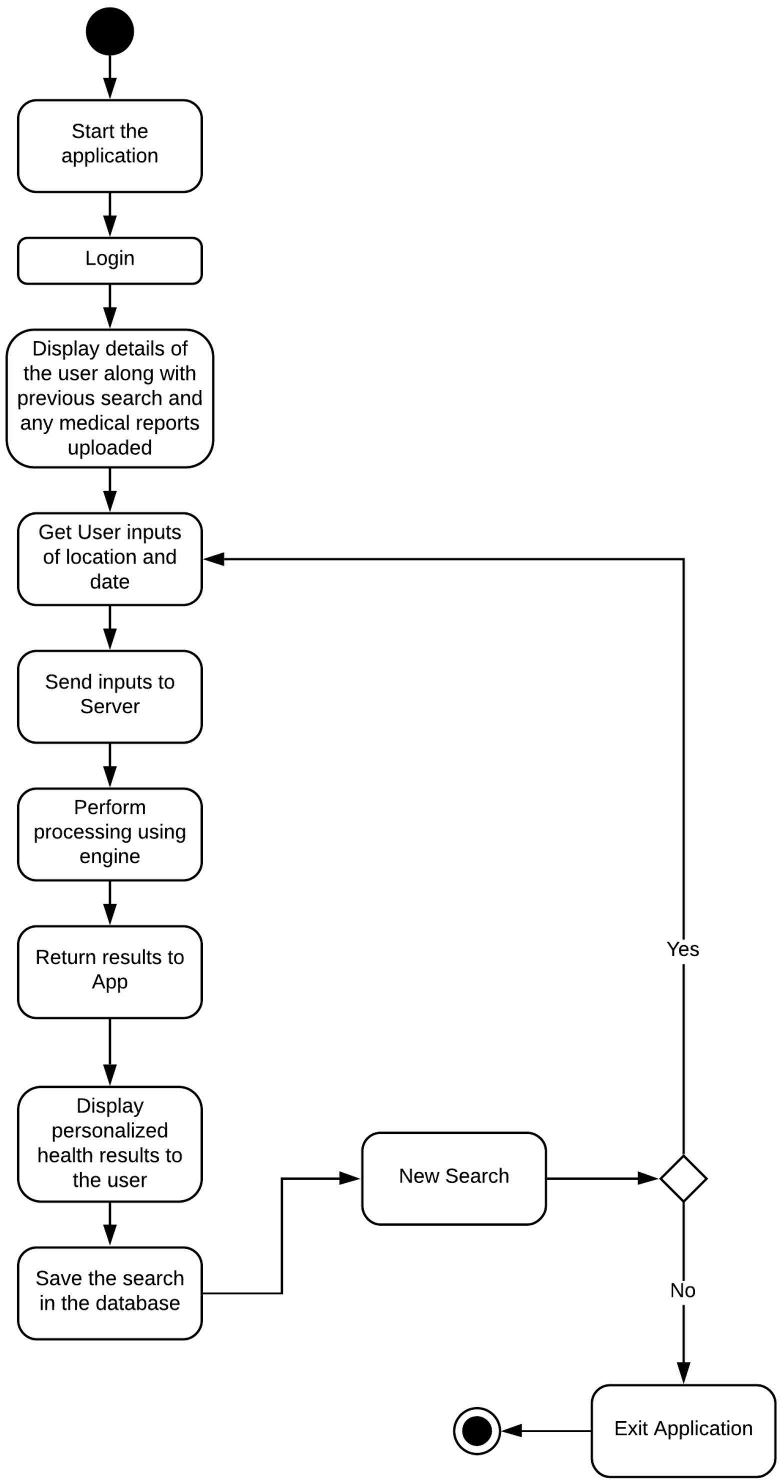

Figure 4.3: Activity Diagram

Figure 4.3 shows the various functions of the system. This is basically a graphical representation of workflow of the system showing how the different activities of the system are linked or carried out. Activities begin when the user starts the application and end when the application is closed.

4.1.3.3 Sequence Diagram

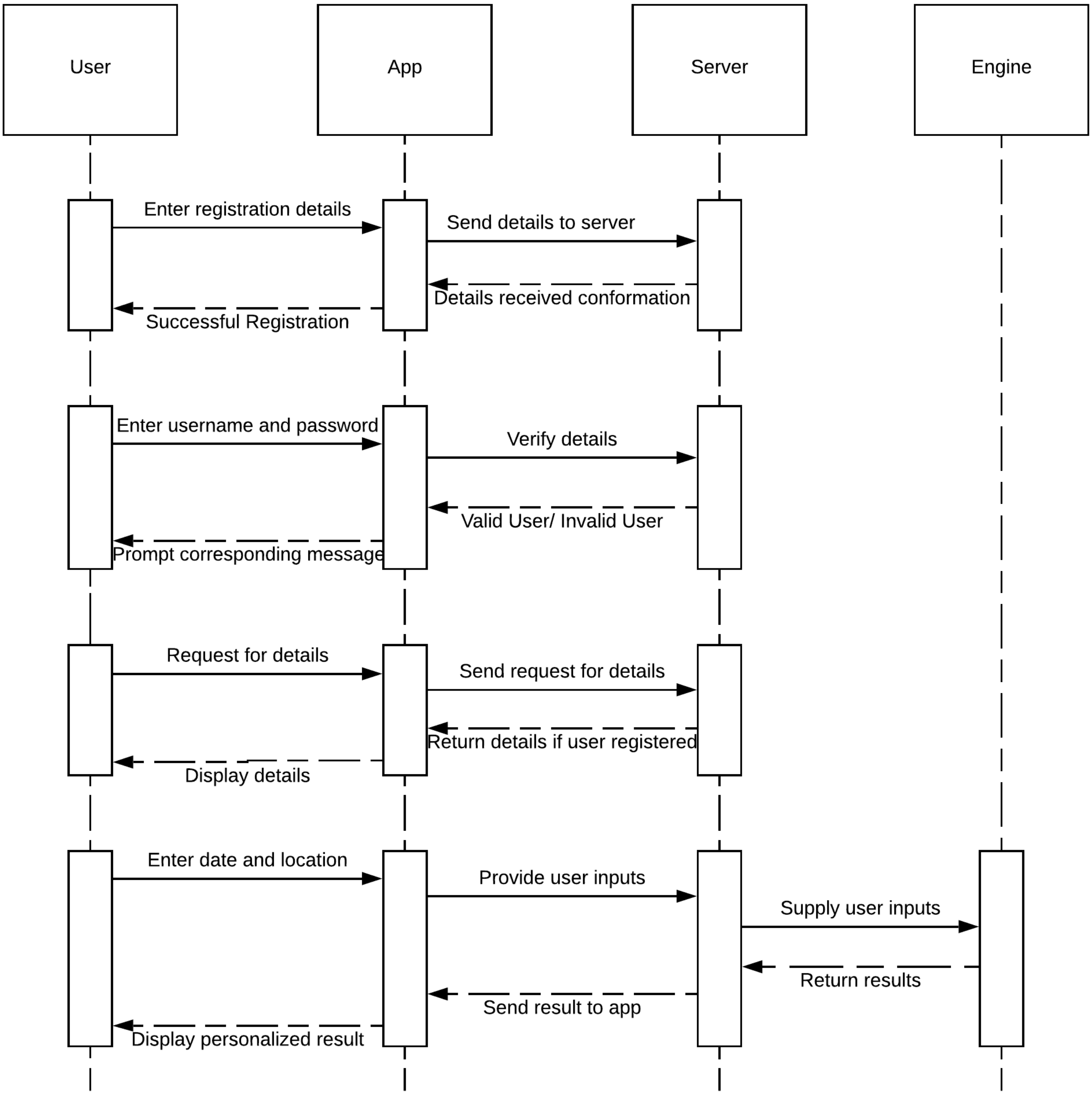

Figure 4.4: Sequence Diagram

Figure 4.4 shows the sequence of the different activities that are being carried out within the system. This diagram depicts how different objects of the system interact with each other in sequential order i.e. the order in which these interactions take place. It starts with an user communicating with an application which further communicates with the server for processing data using engine.

4.1.3.4 Deployment Diagram

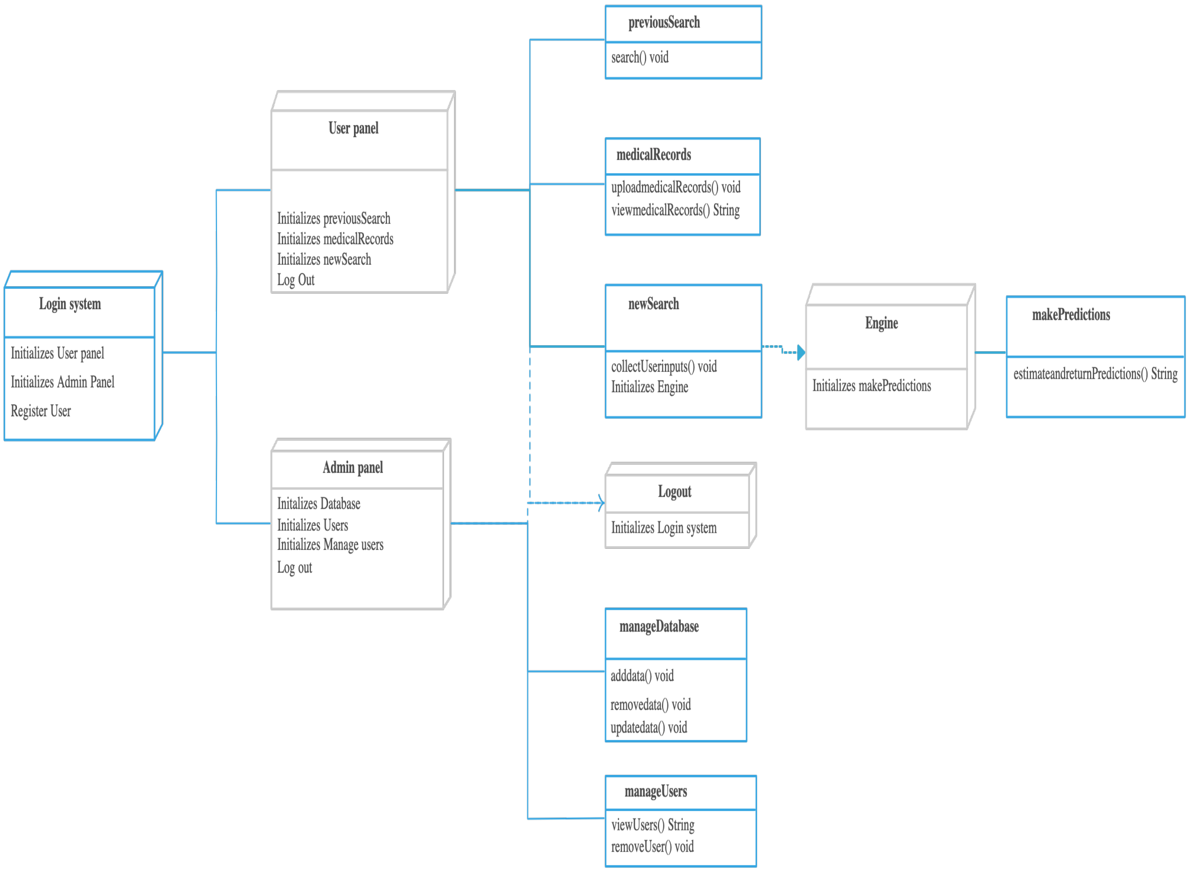

Figure 4.5: Deployment Diagram

Figure 4.5 basically represents what all modules are required for deploying the application in the market. It represents functions inside these modules and how different modules communicate with each other to achieve a specific task.

4.1.3.5 Cost Analysis

	Sub Category	Cost per unit/hour	Number of units/hours	Subtotal
Project Management	Project Managers	500/1	3/200	300000
	Developers	300/1	2/300	180000
	Software Testing	150/1	2/200	60000
	Documentation	100/1	2/50	10000
Systems	Workstations	2500/2	2	50000
	Servers	4000/2	1	40000
Office Space	Rented rooms	35000/Month	1/6 months	210000
Software	Licensed Software	2000/Month	1/6 months	12000
Total				862000

Table 4.1: Cost Analysis

Note: All costs are considered in INR

4.1.3.6 Hardware and software requirements

Hardware Requirements

- CPU with minimum 1.5 GHz frequency
- Minimum 8 GB RAM
- A minimum of 20 GB of available disk space
- Graphic Card of 4 GB or more VRAM

Software Requirements

- Python 3.6 or above
- Java 1.8 or above
- Android Studio 3.5 or above
- Python Libraries
  - TensorFlow
  - Keras
  - Numpy
  - Requests
  - sklearn

4.2 System architecture

4.2.1 UI/UX diagram

UX Diagram

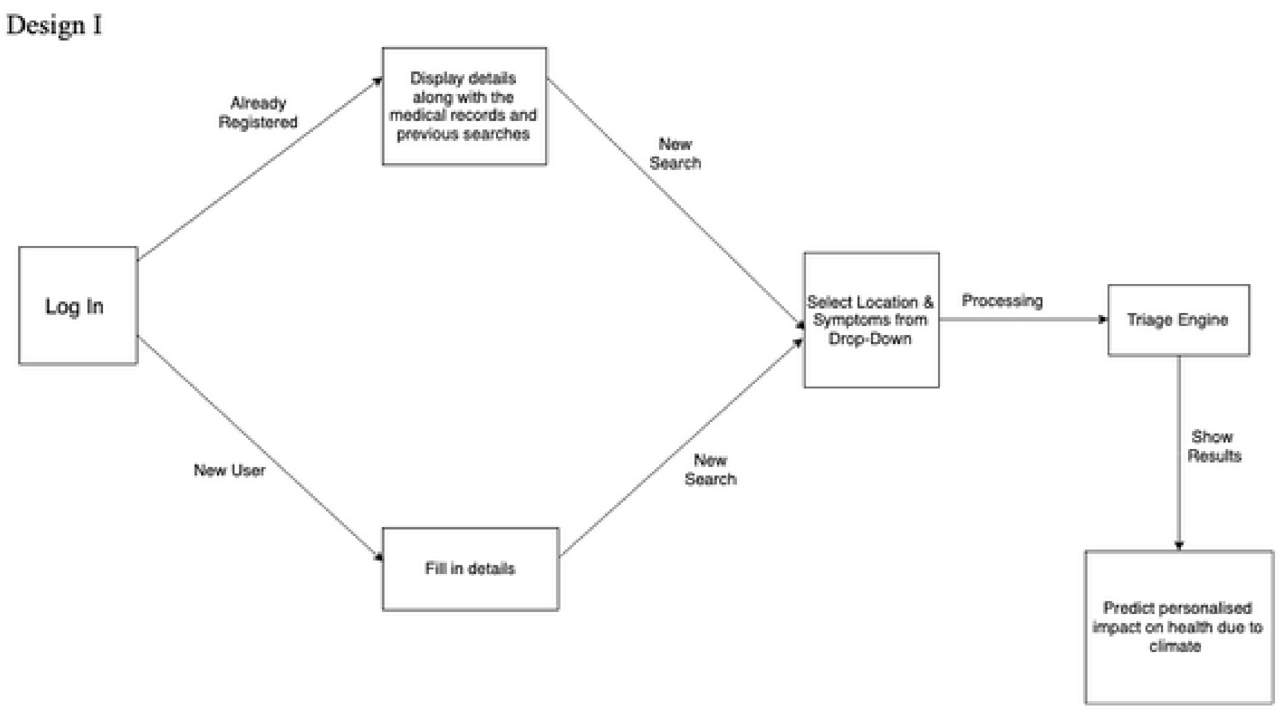

Figure 4.6: User flow

UI diagrams

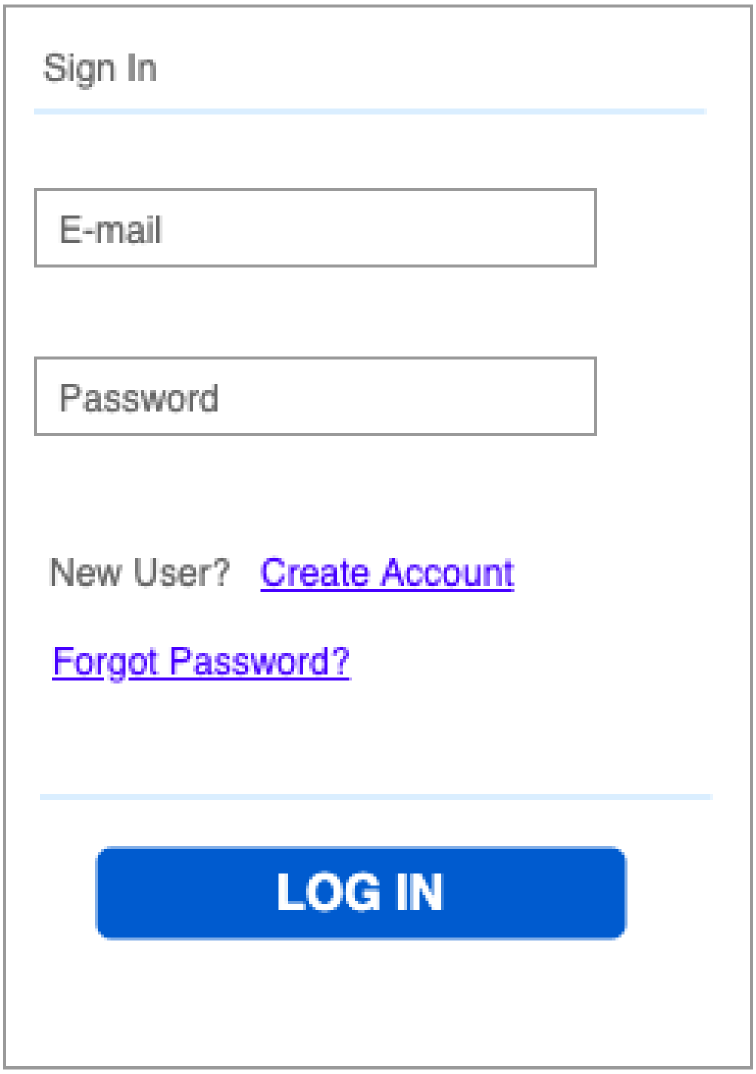

Figure 4.7: Sign-in

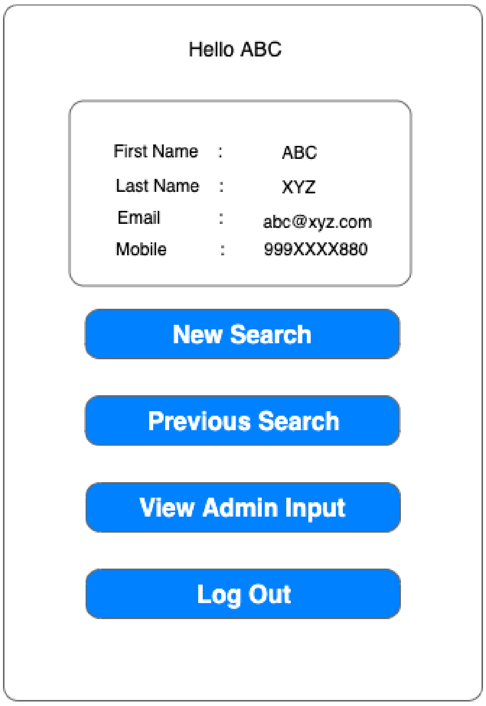

Figure 4.8: Home Page

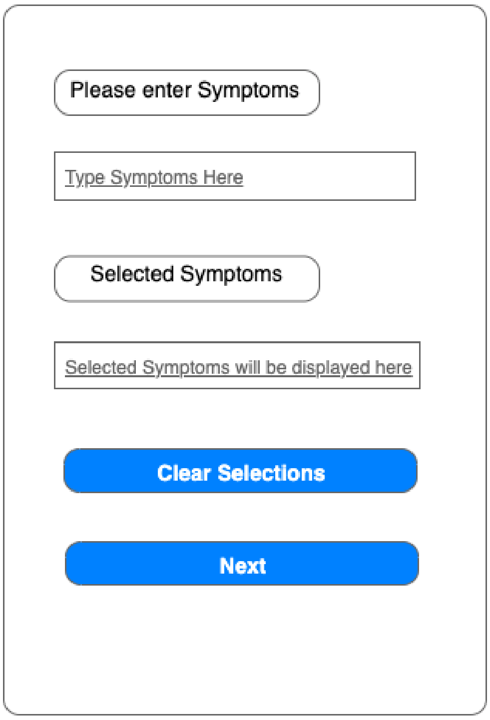

Figure 4.9: Selecting symptoms

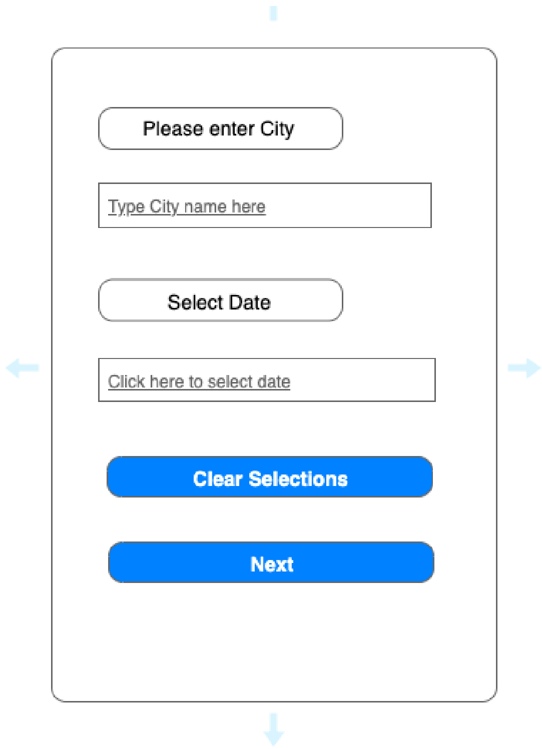

Figure 4.10: Selecting location and date

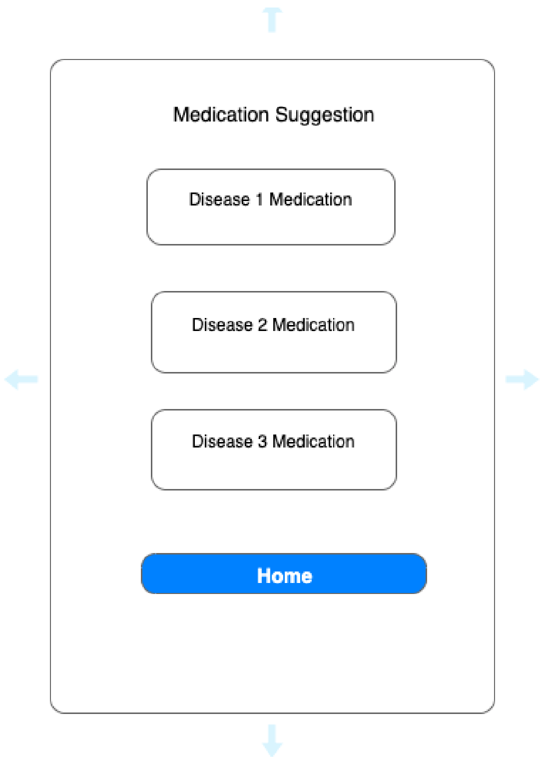

Figure 4.11: Medication Suggestion

4.2.2 Block Diagram

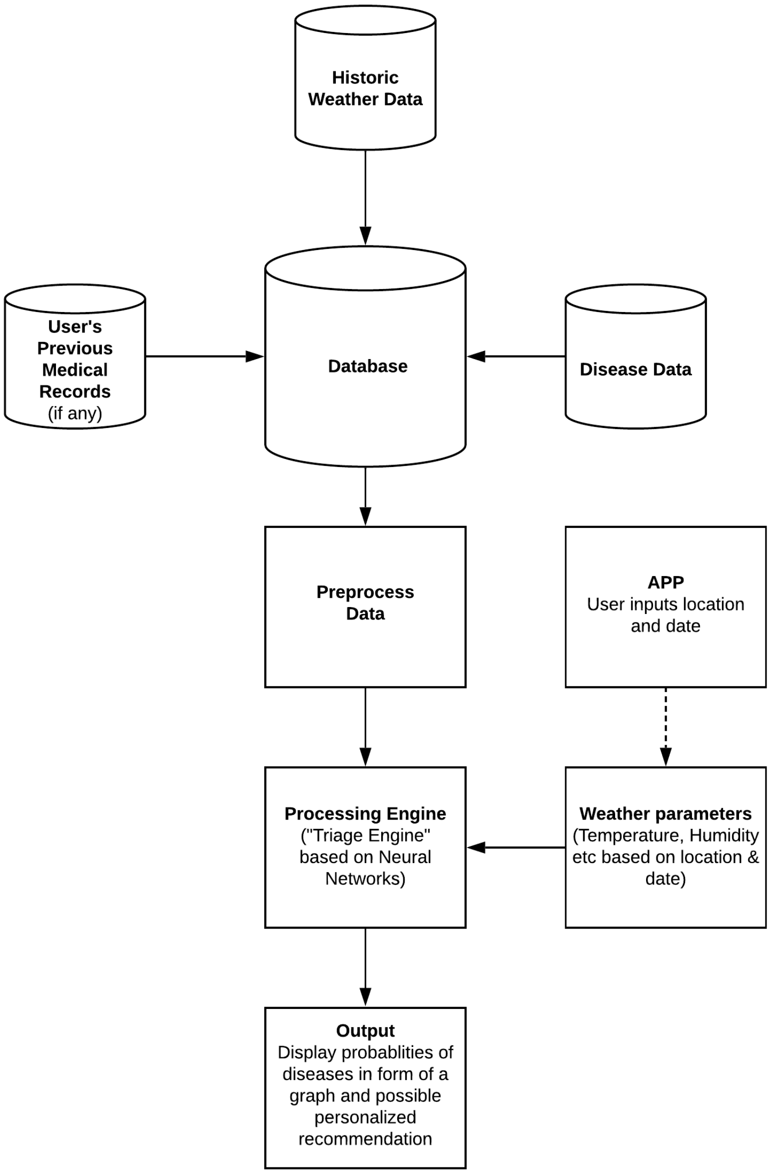

Figure 4.12: Block Diagram

Figure 4.12 depicts different important parts of the system as blocks which are connected by lines that show the relationship between the blocks. It has different blocks such as for data pre-processing data, engine for processing, which refer to the database for various information required.

Chapter 5

Result and Discussion

5.1 Screenshots of the System

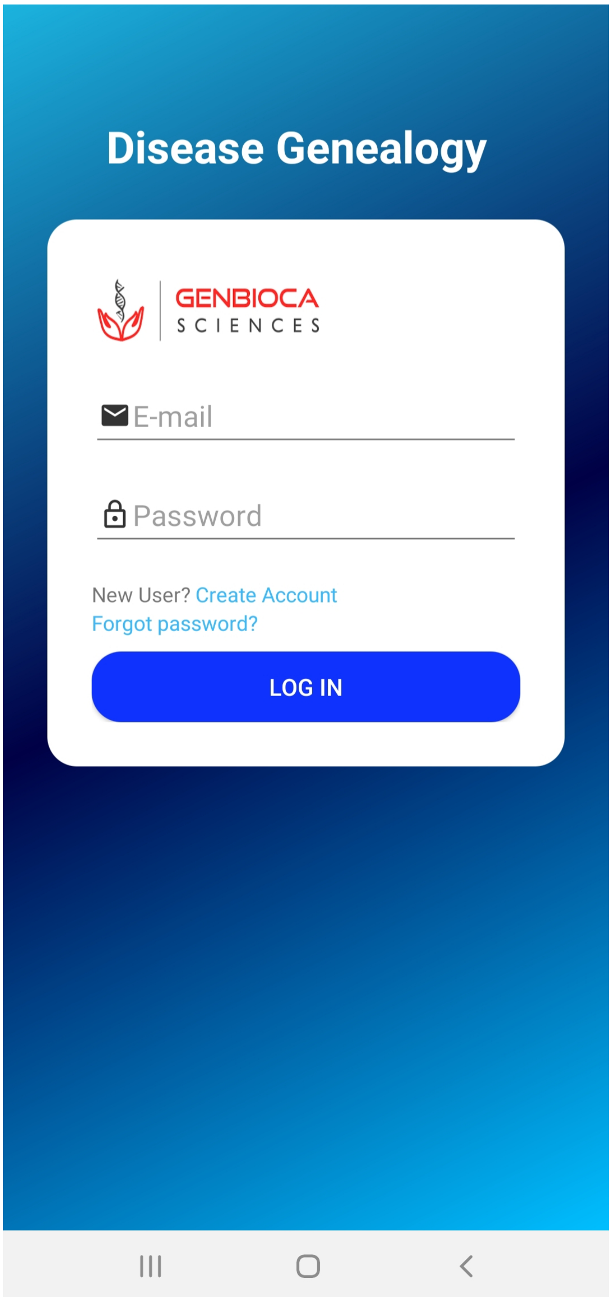

Figure 5.1: Login

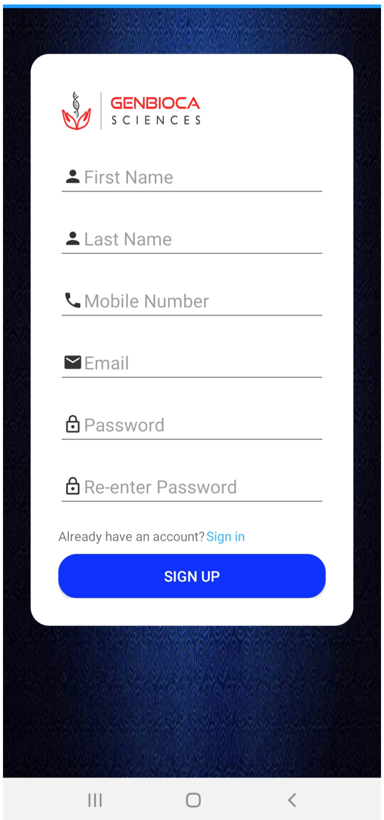

Figure 5.2: Sign-up

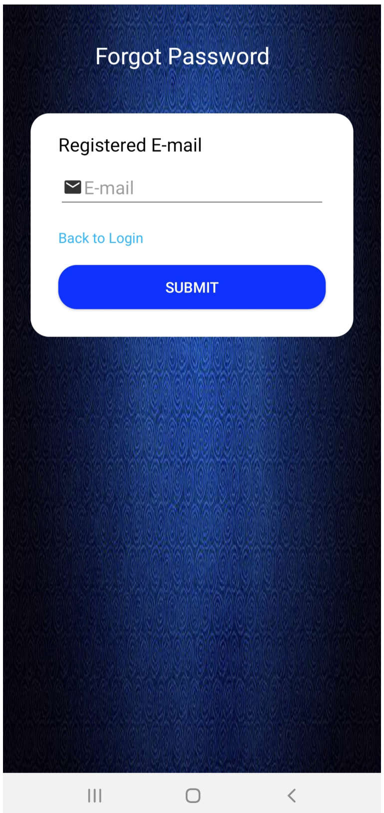

Figure 5.3: Forgot password

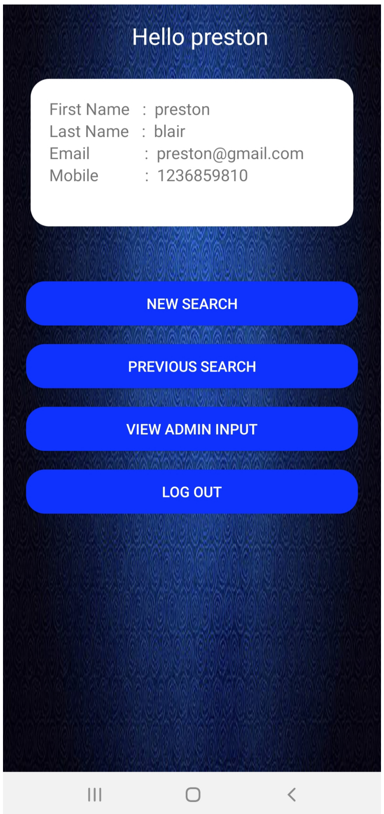

Figure 5.4: Home Page

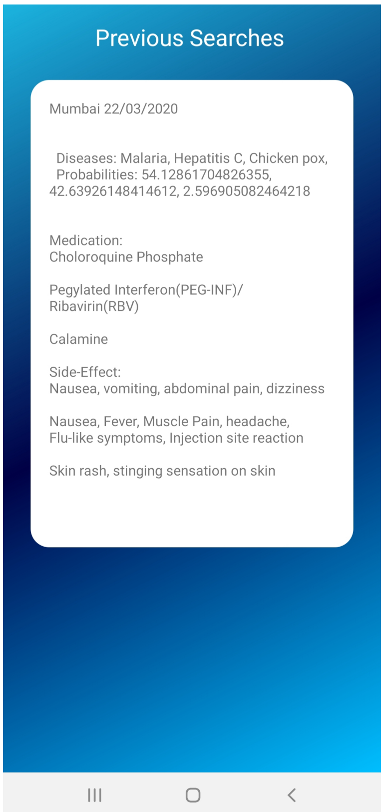

Figure 5.5: Previous Searches of User

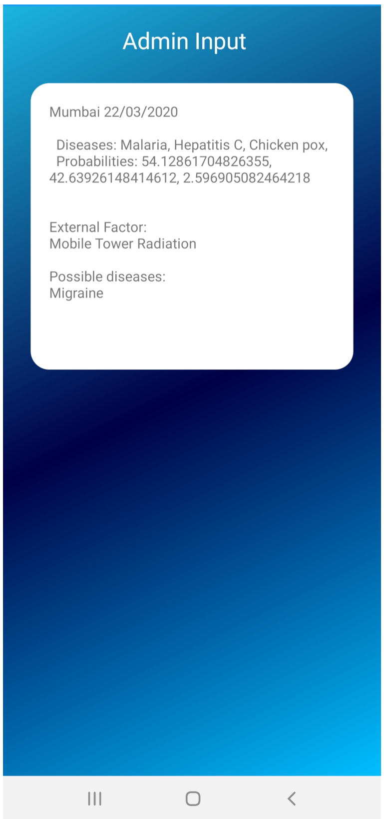

Figure 5.6: Admin Input

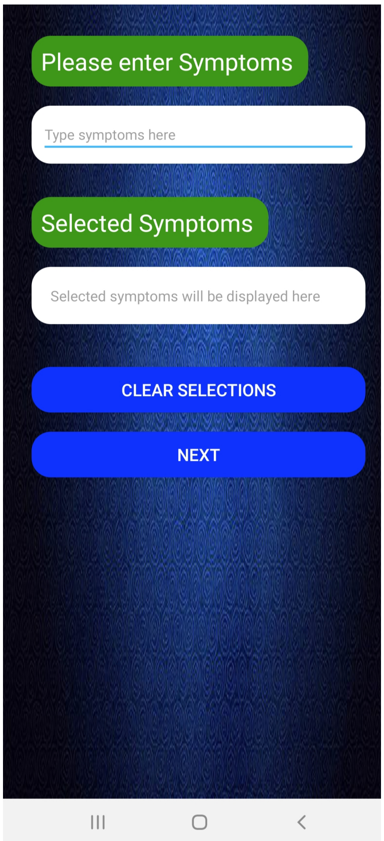

Figure 5.7: New Search

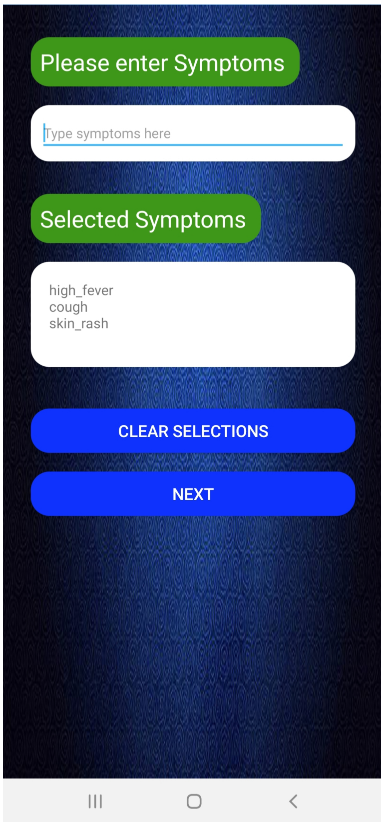

Figure 5.8: Entering Symptoms

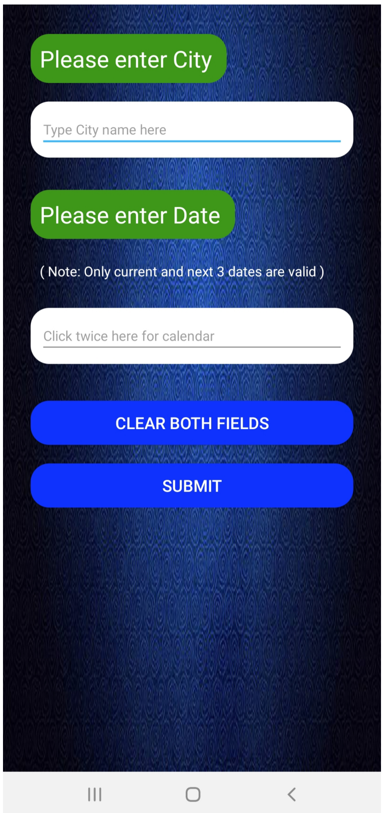

Figure 5.9: City input

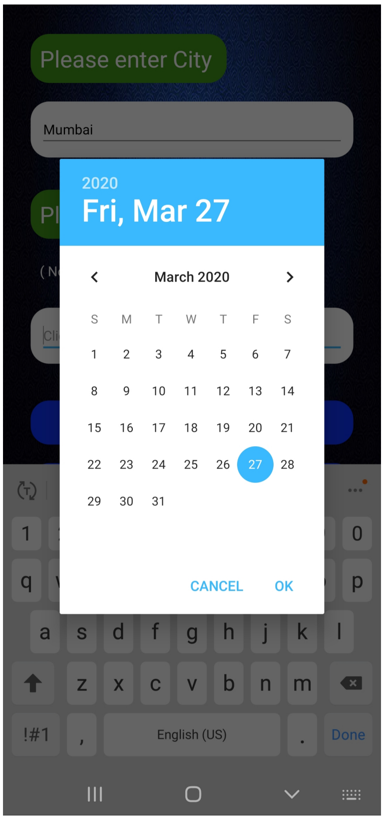

Figure 5.10: Selecting date

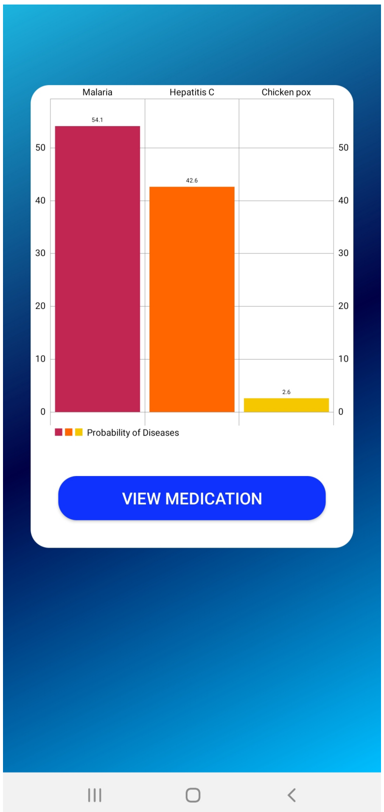

Figure 5.11: Disease Probability Graph

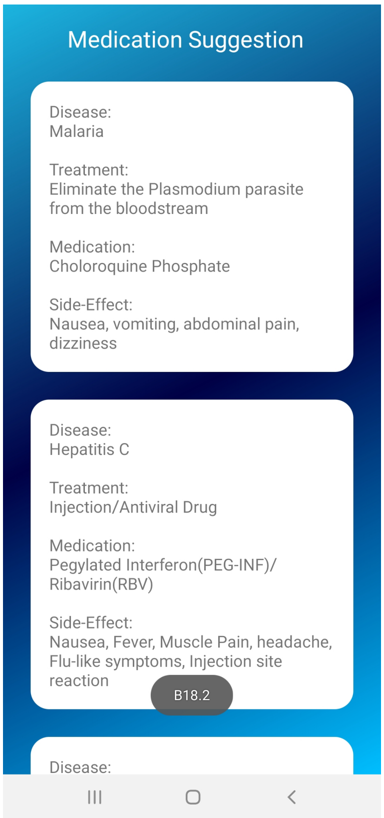

Figure 5.12: Medication Suggestions

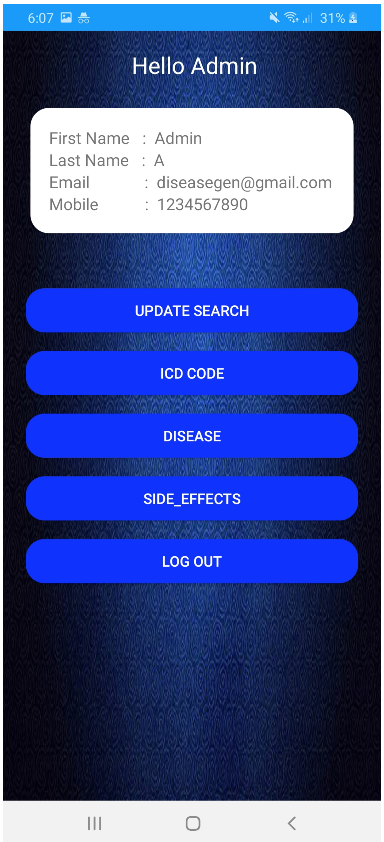

Figure 5.13: Admin Page

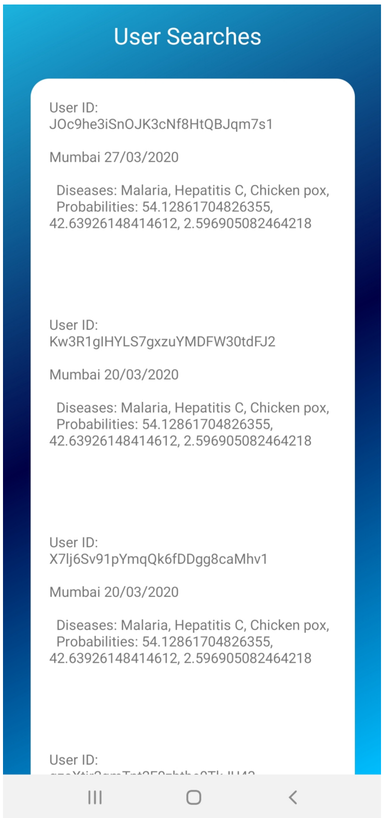

Figure 5.14: User Searches

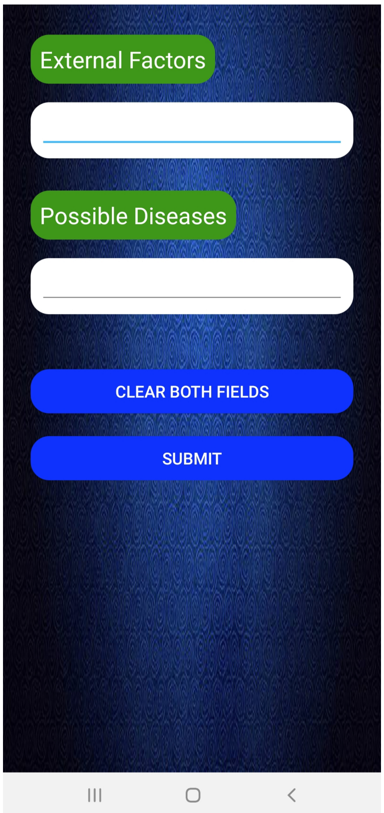

Figure 5.15: External Factors and Possible Diseases

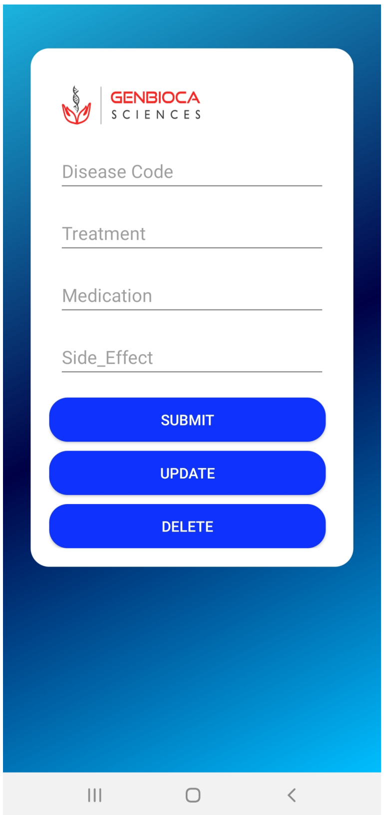

Figure 5.16: Entering Side-effect for a medication

5.2 Sample Code

5.2.1 Code for fetching weather parameters based on location

```
from flask import Flask, request, jsonify
import requests
import traceback
import numpy as np
import pandas as pd

import matplotlib.pyplot as plt
from keras.models import Sequential as seq
from keras.layers import MaxPooling2D as MaxP2D
from keras.layers import Dense, Flatten, Dropout
from keras.optimizers import SGD
from keras.layers import LSTM
import os
from sklearn.preprocessing import MinMaxScaler
from sklearn.preprocessing import LabelEncoder
from sklearn.metrics import accuracy_score, mean_absolute_error
from sklearn.ensemble import RandomForestClassifier as RFC
from sklearn.model_selection import train_test_split
from collections import OrderedDict


app = Flask(__name__)

@app.route('/predict', methods = ['POST','GET'])
def weather():

    mintemp = []
    maxtemp = []
    press = []
    humid = []
    windsp = []
    prec = []
    t = request.get_json()
    print(t)
    c = t['city']
    dt = t['date']
    symptoms = t['symptoms']
    print(symptoms[0])
    city = ''
    for item in c:
        city += item
    print(city)
    de = ''
    for item in dt:
        de += item
    print(de)
    api_key = 'eb14201873011xxxxxxxxx3e15ee267'
    api_call = 'https://api.openweathermap.org/data/2.5/
```

```
    forecast?appid='+api_key+'&units=metric'
    api_call += '&q=' + city
    json_data = requests.get(api_call).json()
    location_data = {
        'city': json_data['city']['name'],
        'country': json_data['city']['country']
    }

    print('\n{city}, {country}'.format(**location_data))

    count=temp_min=temp_max=wind_speed=pressure=humidity
    =prept=0
    date1 = json_data['list'][0]['dt_txt'].split(' ')[0]
    date_count=0


    for item in json_data['list']:
        time = item['dt_txt']
        next_date = time.split(' ')[0]

        if date1==next_date:
            count+=1
            if 'rain' in item and '3h' in item['rain']:
                prept+=item['rain']['3h']
            temp_min+=item['main']['temp_min']
            temp_max+=item['main']['temp_max']
            wind_speed+=item['wind']['speed']
            pressure+=item['main']['pressure']
            humidity+=item['main']['humidity']

        if date1 != next_date:
            date_count+=1
            year, month, day = date1.split('-')
            date = {'y': year, 'm': month, 'd': day}
            d = day + '/' + month + '/' + year
            print(d)
            print('\n{m}/{d}/{y}'.format(**date))
            if de == d:
                print("Min temp: {0:.2f}C".
                    format(temp_min/count))
                a = round(temp_min/count,2)
                mintemp.append(a)
                print("Max temp: {0:.2f}C".
                    format(temp_max/count))
                b = round(temp_max/count,2)
                maxtemp.append(b)
                print("Wind speed: {0:.2f} m/s".
                    format(wind_speed/count))
                c = round(wind_speed/count,2)
                windsp.append(c)
                print("Pressure: {0:.2f} hpa".
                    format(pressure/count))
```

```
                d = round(pressure/count,2)
                press.append(d)
                print("Humidity: {0:.2f} %".
                    format(humidity/count))
                z = humidity/100
                e = round(z/count,2)
                humid.append(e)
                print("Precipitation: {0:.2f} mm".
                    format(prept/count))
                f = round(prept/count,2)
                prec.append(f)
                count = 1
                temp_min=temp_max=wind_speed=pressure=humidity
                =prept=0
                if 'rain' in item and '3h' in item['rain']:
                    prept+=item['rain']['3h']
                temp_min +=item['main']['temp_min']
                temp_max +=item['main']['temp_max']
                wind_speed+=item['wind']['speed']
                pressure+=item['main']['pressure']
                humidity+=item['main']['humidity']
                break

        date1 = time.split(' ')[0]

    uv = 0

    sym_list = weather_to_symptoms(maxtemp[0], mintemp[0],
        humid[0], prec[0], windsp[0], uv, press[0])

    string = ''
    for i in symptoms[0]:
        if(i == '[' or i == ' '):
            continue
        else:
            if i == ',' or i == ']':
                if string not in sym_list:
                    sym_list.append(string)
                string=''
            else:
                string += i


    sym_list = list(OrderedDict.fromkeys(sym_list))

    print(sym_list)
    final_list,v = symptoms_to_disease(sym_list)

    return jsonify({'Diseases': str(final_list),
        'Probabilities': str(v)})
```

5.3 Testing

5.3.1 Unit Testing

In this type of testing individual components are tested to see how they perform. These components could be just few lines of code giving a certain output. Unit testing was a integral part of the project making. After each part that was developed unit testing was performed. For example: After the code for the login page was developed, it was immediately tested on a phone to see the layout of the different fields, how the input was taken etc. Similarly after a basic server was set up, it was tested to see how it interacts with the app by sending simple messages and checking reply after which we moved on to more complex one.

5.3.2 Integration Testing

In this kind of testing different components are clubbed for testing to see how they perform together. There were many instances of integration testing during our project. When we created the two models for prediction, the output of one model was to be input to the other model. Thus we tested it together to see the performance. Then we attached these two models with API that retrieves location weather parameters which were the input to the first model and tested the performance. Finally we created the server which housed all these and made a connection to the android application which the user would interact with. The android application had to send the parameters over to the server and the server had to return back the output. All the components need to work in synchronization.

5.3.3 Black-box Testing

In black-box testing the application is tested by a person who does not know how the internal coding is done. This test is done to see how well the application would perform in hands of the user who is not concerned with internal working. To do so we gave the application to our friends and colleagues, and asked to use the application with out guiding them. This test helped to fix certain issues like ambiguity of what to do on certain screen due to lack of instructions, how to improve the layout, changes to the aesthetics of the application etc. All these were done after getting feedback from the test.

Chapter 6

Conclusion & Future Scope

Currently, we were able to develop a system that can predict few different diseases like dengue, malaria, jaundice etc based on certain weather parameters and some symptoms if experienced by user with 45-50% accuracy. Also, we have completed the android application which is user-friendly and developed programs to extract weather data based on location and date, connect the android application to a server using Flask; which is a framework in python and, where the API and models based on Neural Networks are executed and results are returned to the android application to be displayed to user.

We are going to increase the scope of our system to multiple diseases which would enable the application to be used by government agencies or relief teams to better prepare for any disease that may occur. Also, it could help doctors as they can expect a high number of cases related to a particular disease. Tourists or business people can also plan their journey accordingly as the results given by the app are personalized. It will also be beneficial to people who suffer from a certain disease to prepare themselves accordingly.

References

[1] D. B. Parth Wazurkar, Robin Singh Bhadoria, "Predictive analytics in data science for business intelligence solutions," *IEEE, 7th International Conference on Communication Systems and Network Technologies*, pp. 368–370, 2017.

[2] A. Z. A. B. Ayham Omary, Ahmad Wedyan and I. Alsmadi, "An interactive predictive system for weather forecasting," *IEEE, International Conference on Computer, Information and Telecommunication Systems*, 2012.

[3] A. Quinn and J. Shaman, "Health symptoms in relation to temperature, humidity, and self-reported perceptions of climate in new york city residential environments," *International Journal Of Biometeorology*, vol. 61, pp. 1209–1220, 2017.

[4] P. S. K. A. Rishika Reddy, "Predictive big data analytics in healthcare," *IEEE, Second International Conference on Computational Intelligence & Communication Technology*, pp. 623–626, 2016.

[5] M. C. Blagoj Ristevski, "Big data analytics in medicine and healthcare," *Journal of Integrative Bioinformatics*, 2018.

[6] H. Gulati, "Predictive analytics using data mining technique," *IEEE, Second International Conference on Computing for Sustainable Global Development*, pp. 713–716, 2015.

[7] K. Y. U. O. T. Mihye Lee, Sachiko Ohde and T. Fukui, "Weather and health symptoms," *International Journal Of Environment Research And Public Health*, vol. 15(8), p. 1670, 2017.

[8] K. Deepika and D. S. Seema, "Predictive analytics to prevent and control chronic diseases," *2nd International Conference on Applied and Theoretical Computing and Communication Technology*, 2016.

Acknowledgement

Success of a project like this involving high technical expertise, patience and massive support of guides, is possible when team members work together. We take this opportunity to express our gratitude to those who have been instrumental in the successful completion of this project. We would like to show our appreciation to **Mrs. Rakhi Kalantri** for their tremendous support and help, without them this project would have reached nowhere. We would also like to thank our project coordinator **Mrs. Rakhi Kalantri** for providing us with regular inputs about documentation and project timeline. A big thanks to our HOD **Dr. Lata Ragha** for all the encouragement given to our team. We would also like to thank our principal, **Dr. S. M. Khot**, and our college, **Fr. C. Rodrigues Institute of Technology, Vashi**, for giving us the opportunity and the environment to learn and grow.

Project Group Members:

1. Atharva Ajit Agwekar, 101602

2. Prasun Alexander Philip, 101645

3. Aniket Sandhbor, 101674

Appendix A : Timeline Chart

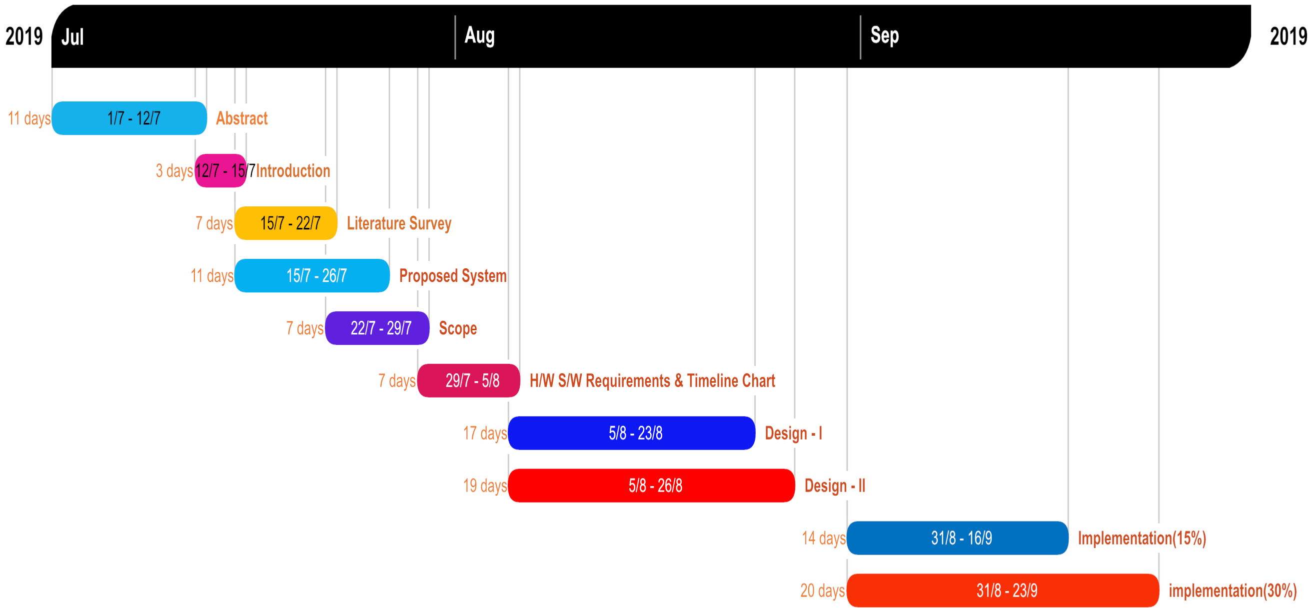

Figure 6.1: Timeline Chart - Sem 7

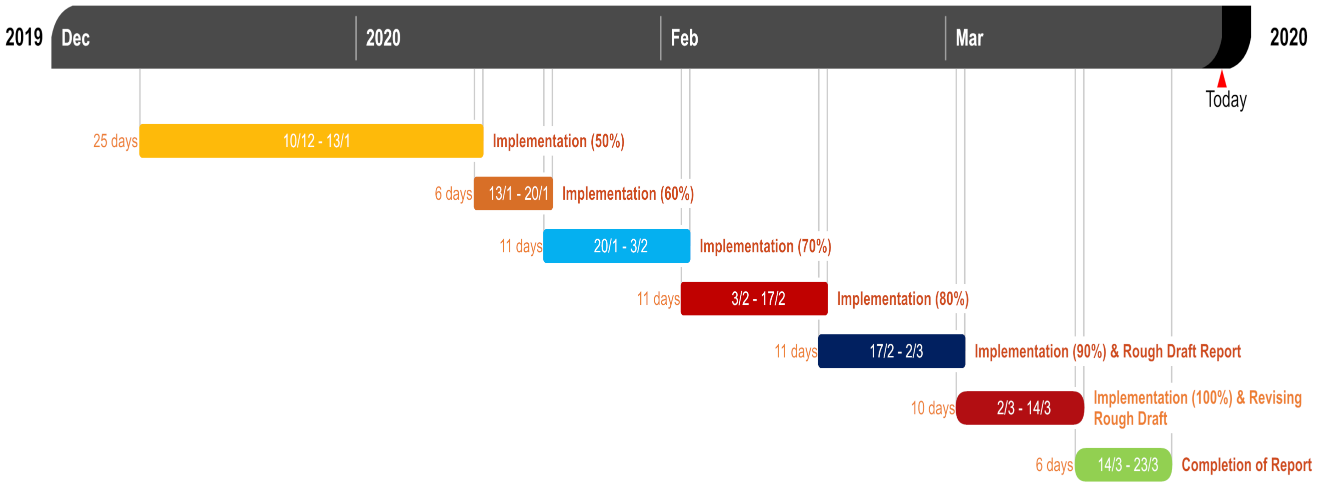

Figure 6.2: Timeline Chart - Sem 8