Dr. ROBO (A DIGITAL GENRAL PHYSICIAN)

KHWAJA FAREED
UEIT
RAHIM YAR KHAN

Faculty of
Computer Sciences &
Information Technology

Submitted by

Fareed Yar

Cs181106

&

Imran Hussain

Cs181068

Supervised by:

Mr. Shahzad Hussain

Project Submitted to

Supervisor

Mr. Shahzad Hussain

In partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science by,

Sr. no	Student Name	Reg. no
01	Fareed Yar	Cs181106
02	Imran Hussain	Cs181068

Final Year Project
Session 2018-2022

Department of Computer Science
Khawaja Fareed University of Engineering & Information Technology
Rahim Yar Khan

DECLARATION

We hereby declare that the project work entitled “Dr. Robo (a digital general physician)” that we’re going to submit to ICT in Khawaja Fareed University of Engineering and Information Technology, is a record of an original work done by us (Fareed Yar and Imran Hussain) carried out during the course of our study under the guidance or supervision of Respected Lecturer Mr. Shahzad Hussain, Department of Computer Science in KFUEIT and this project work is submitted in the partial fulfillment of the requirements for the award of the degree of Bachelor in Computer Science. The results/work in this document have not been submitted to any other University or Institute for the award of any degree or diploma. We’ve followed the guidelines provided by the university in writing report.

Member 1:

Name: Fareed Yar

Reg No: CS181106

Signature: _____

Member 2:

Name: Imran Husaain

Reg No: CS181106

Signature: _____

APPROVAL FOR SUBMISSION

It is certified that this project report entitled **"Dr. Robo (a digital general physician)"** was prepared by **Fareed Yar** and **Imran Hussain** has met the required standard for submission in partial fulfilment of the requirements for the award of Bachelor of Computer Science at Khawaja Fareed University of Engineering & Information Technology.

Approved by,

Signature: _____

Supervisor: Mr. Shahzad Hussain

Date: _____

“ACKNOWLEDGEMENT”

We would like to thank everyone who had contributed to the successful completion of this project. We would like to express our gratitude to our Project supervisor, Mr. Shahzad Hussain for his invaluable advice, guidance and his enormous patience throughout the development of the project. In addition, we would also like to express our gratitude to our loving parents who had helped and given us encouragement.

ABSTRACT

"Dr. Robo (a digital general physician)"is basically disease prediction system based on predictive modeling predicts the disease of the user on the basis of the symptoms that user provides as an input to the system. The system analyzes the symptoms provided by the user as input and gives the probability of the disease as an output Disease Prediction is done by implementing the General Machine learning.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

1.1 Introduction

At present, when one suffers from particular disease, then the person has to visit to doctor which is time consuming and costly too as well as in Pakistan the doctors recommend unnecessary lab tests. Also if the user is out of reach of doctor and hospitals it may be difficult for the user as the disease cannot be identified. So, if the above process can be completed using an automated program which can save time as well as money, it could be easier to the patient which can make the process easier. There are other Heart related Disease Prediction System using data mining techniques that analyzes the risk level of the patient.

Dr. Robo (a digital general physician) is a hybrid application that predicts the disease of the user with respect to the symptoms given by the user. Dr. Robo has data sets collected from different health related sites. With the help of Dr. Robo the user will be able to know the probability of the disease with the given symptoms and recommend medicines for minor diseases and done further evaluation for chronic disease like covid-19, heart disease, kidneys diseases. It suggests necessary test when user or patient has symptoms of chronic disease like ultrasound for kidney diseases ECG for heart diseases and X-Rays for detecting covid-19

As the use of internet is growing every day, people are always curious to know different new things. People always try to refer to the internet if any problem arises. People have access to internet than hospitals and doctors. People do not have immediate option when they suffer with particular disease. So, this system can be helpful to the people as they have access to internet 24 hours.

1.2 Problem Statement

There are many tools related to disease prediction. But particularly heart related diseases have been analyzed and risk level is generated. But generally there are no such tools that are used for prediction of general diseases. So Dr. Robo helps for the prediction of the general diseases as well as recommend medicines for minor diseases and suggest necessary lab tests for chronic diseases for further evaluation also provide detailed information about disease.

1.3 Objective

1.3.1 General Objective:

-To implement Naïve Bayes Classifier that classifies the disease as per the input of the user.

-To Implement CNN (convolution neural network) That classifies ECG, ultrasound, x-ray for further evaluation of chronic diseases

1.3.2 Specific Objective:

-To develop cross platform based system for the prediction of the disease.

1.4 Scope and Limitations

1.4.1 Scope:

This project aims to provide a web platform to predict the occurrences of disease on the basis of various symptoms. The user can select various symptoms and can find the diseases with their probabilistic figures.

1.4.2 Limitations:

The limitations of this project are:

a. Dr. Robo does not recommend medications of the chronic disease.

b. Past history of the disease has not been considered

c. Works at all platform like web, android, iOS, desktop but internet connection is necessary

CHAPTER 2 REQUIREMENT ANALYSIS AND FEASIBILITY ANALYSIS

2.1 Literature Review

K.M. Al-Aidaroos, A.A. Bakar and Z. Othman have conducted the research for the best medical diagnosis mining technique. For this authors compared Naïve Baeyes with five other classifiers i.e. Logistic Regression (LR), KStar (K*), Decision Tree (DT), Neural Network (NN) and a simple rule-based algorithm (ZeroR). For this, 15 real-world medical problems from the UCI machine learning repository (Asuncion and Newman, 2007) were selected for evaluating the performance of all algorithms. In the experiment it was found that NB outperforms the other algorithms in 8 out of 15 data sets so it was concluded that the predictive accuracy results in Naïve Baeyes is better than other techniques.

Table 1- Predictive Accuracy of Bayes and other Technique

Medical Problems	NB	LR	K*	DT	NN	ZeroR
Breast Cancer wise	97.3	92.98	95.72	94.57	95.57	65.52
Breast Cancer	72.7	67.77	73.73	74.28	66.95	70.3
Dermatology	97.43	96.89	94.51	94.1	96.45	30.6
Echoeardiogram	95.77	94.59	89.38	96.41	93.64	67.86
Liver Disorders	54.89	68.72	66.82	65.84	68.73	57.98
Pima Diabetes	75.75	77.47	70.19	74.49	74.75	65.11
Haeberman	75.36	74.41	73.73	72.16	70.32	73.53
Heart-c	83.34	83.7	75.18	77.13	80.99	54.45
Heart-statlog	84.85	84.04	73.89	75.59	81.78	55.56
Heart-b	83.95	84.23	77.83	80.22	80.07	63.95
Hepatitis	83.81	83.89	80.17	79.22	80.78	79.38
Lung Cancer	53.25	47.25	41.67	40.83	44.08	40
Lymphpgraphy	84.97	78.45	83.18	78.21	81.81	54.76
Primary tumor	49.71	41.62	38.02	41.39	40.38	24.78
Postooerative Patient	68.11	61.11	61.67	69.78	58.54	71.11
Wins	8\15	5\15	0\15	2\15	1\15	1\15

(Al-Aidaroos, Bakar, & Othman, 2012)

2.2 Requirement Analysis

2.2.1 Functional requirements:

a. Predict disease with the given symptoms.

b. Compare the given symptoms with the input datasets

c. Check status whether is it chronic or minor disease

d. If minor then suggest the medicines, provide details of disease through web scrapping

e. If chronic disease the run further classifier like cnn to evaluate chronic disease

2.2.2 Non-functional requirements:

a. Display the list of symptoms where user can select the symptoms.

b. Naïve Bayes Classifier is used to classify the data sets.

c. CNN used to analyze images of reports

d. Web scrapping for medicines suggestion and information extraction

2.3 Feasibility Analysis

2.3.1 Technical feasibility:

The project is technically feasible as it can be built using the existing available technologies. It is a web based applications that uses python's web framework Django using restful api to treat all platform like android, iOS, desktop defiantly web. The technology required by Disease Predictor is available and hence it is technically feasible.

2.3.2 Economic feasibility:

The project is economically feasible as the cost of the project is involved only in the hosting of the project. As the data samples increases, which consume more time and processing power. In that case better processor might be needed.

2.3.3 Operational feasibility:

The project is operationally feasible as the user having basic knowledge about computer and Internet. Disease Predictor is based on client-server architecture where client is users and server is the machine where datasets are stored.

CHAPTER 3 SYSTEM DESIGN

3.1 Methodology

Disease Prediction has been already implemented using different techniques like Neural Network, decision tree and Naïve Byes algorithm. Particularly heart related disease is mostly analyzed. From the analysis it was found that Naïve Bayes is more accurate than other techniques. So, Disease Predictor also uses Naïve Bayes for the prediction of different diseases.

3.1.1 Data collection:

Data collection has been done from the internet to identify the disease here the real symptoms of the disease are collected i.e. no dummy values are entered. The symptoms of the disease are collected from different health related websites

3.1.2 Algorithm implemented:

The algorithm implemented in this project are Naïve Bayes Classifier and CNN for reports images analysis to evaluate chronic diseases.

Naïve Bayes classifier depends on Bayes Theorem

3.2.1 State diagram

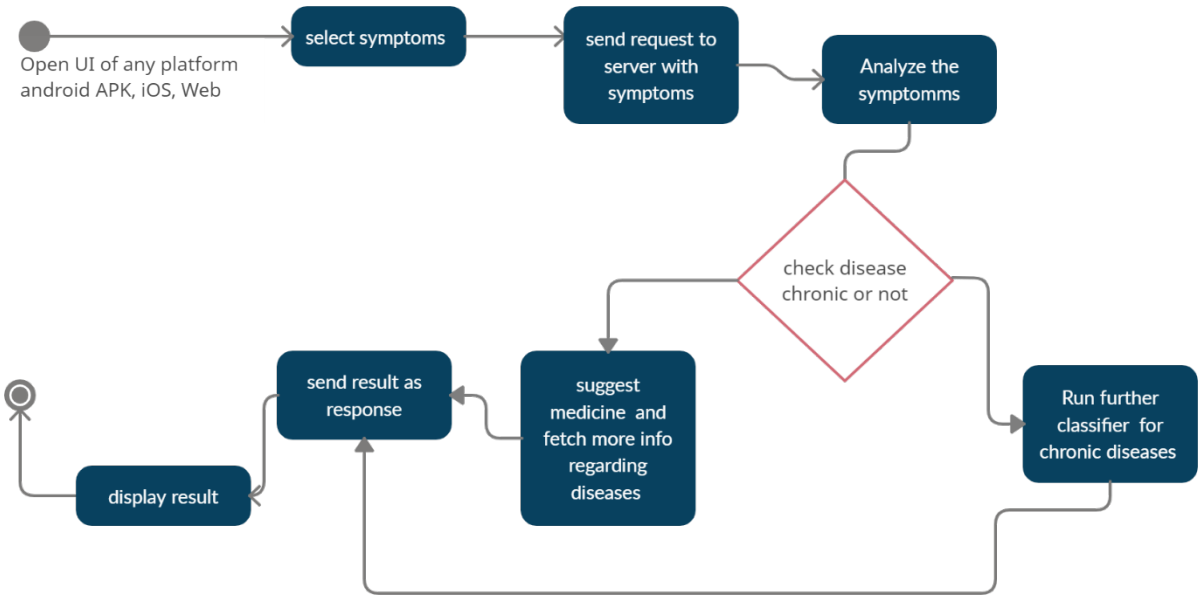

Figure 1- State Diagram

It explains different state of the system. First the user opens any ui of platform. The user selects the symptoms. When finished selecting symptoms the user submits the symptoms as response disease predicted on base of symptoms further check it if it is chronic like covid-19, heart-attack, kidney problem etc we run further models to evaluate if disease are minors then we done web scrapping and suggest medicines and more details and finally send result as response and display at ui end.

3.2.3 Sequence diagram

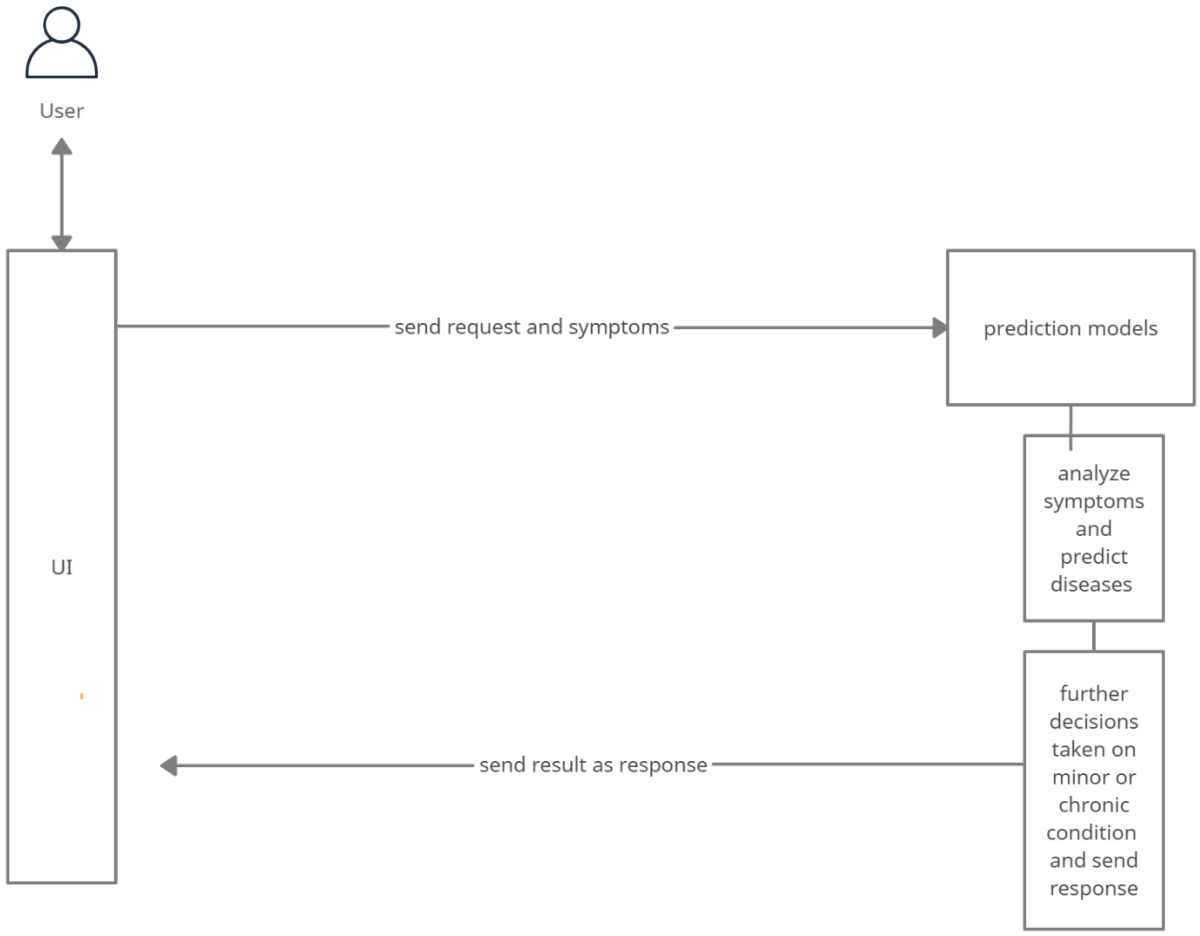

Figure 3- Sequence Diagram

It explains the sequence of the project. Initially system shows the symptoms to be selected at ui end. The user selects the symptoms and submits to the system as request. The system process fisrt predict the disease based on symptoms once disease is predicted after it we decide whether is it chronic or minor after we decide further processing and send result as response

CHAPTER 4 IMPLEMENTATION AND TESTING

4.1 Implementation

Disease Predictor is the ability to predict the disease that has been provided to the system. For disease prediction, we need to implement the naïve Byes Classifier.

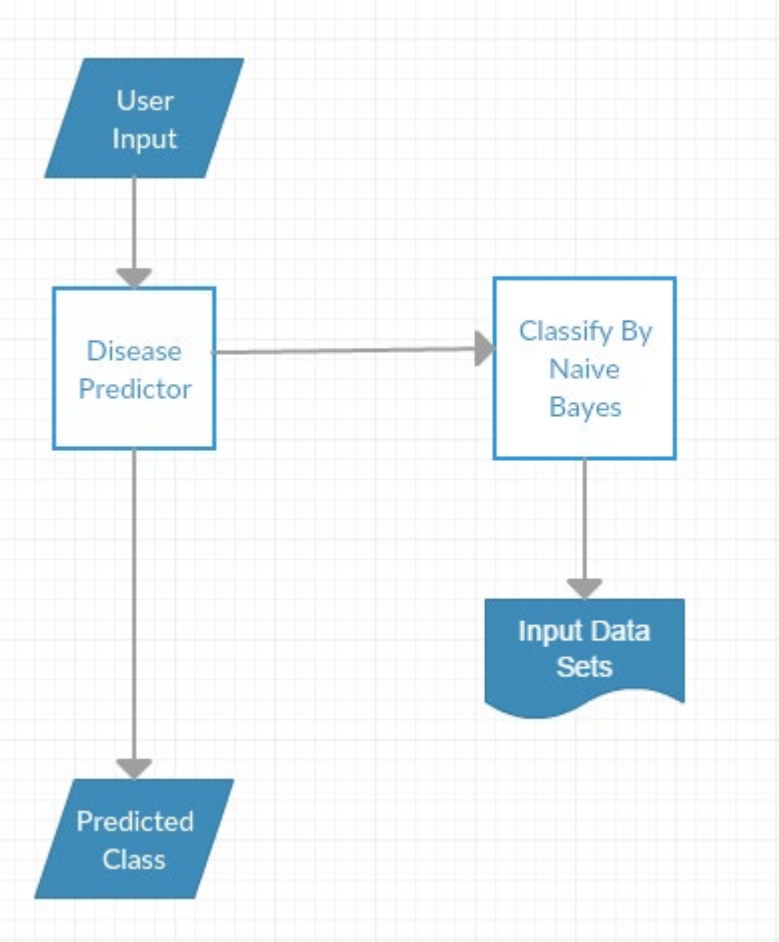

Figure 4- Workflow

As shown in the figure the input data sets are classified using Naïve Bayes classifier. The sample input data sets is shown below

Table 2- Sample Data Sets

Symptoms	Disease
Runny nose ,Sore throat ,Cough ,Congestion, body aches, headache ,Sneezing , fever	Common cold
Fever ,profuse sweating ,headache ,nausea ,vomiting ,diarrhea ,anemia ,muscle pain ,convulsions ,coma bloody stools ,shaking chill	Malaria
poor appetite ,abdominal pain ,headaches ,generalized aches and pains ,fever ,lethargy ,intestinal bleeding or perforation ,diarrhea , constipation	Typhoid

Naïve Bayes classifier uses the following rule to classify the datasets:

$$\hat{Y} = \text{ARG MAX } P(\text{Disease}) \prod_{i=1}^{n} P(\text{symptom}_i \mid \text{Disease})$$

User gives input to the system. The input consists of symptoms. The user marks the symptoms due to which the user is feeling unwell.

1. ☐ Fever
2. ☐ Cough
3. ☐ Vomiting

The “Dr Robo” predicts the disease according to the input data sets and calculates the probability of the disease.

The sample output is given as:

Table 3- Sample Output

Disease Name	Probability
Typhoid	0.5%
Malaria	0.3%
Flu	0.333%

4.1.1 Tools Used

1. KERAS based on Tenser flow framework python coding to implement ML,DL algorithms
2. Goggle Colab to build, train and save models
3. Django a python web framework to implement saved train model to predict
4. Django Restful API to enable cross platform working
5. UI depends upon platform used like android,web,iOS
6. Web scrapping for medicine suggestion and info extraction
7. Creatly.com to design diagrams

4.2 Testing

The test case designed for the project is discussed below:

Test Case- I: Submit the symptoms from the list
Precondition: The application is open.
Assumptions: The symptoms for the disease are available
Test steps: 1. Select the checkbox from the list 2. Select submit
Expected Result: The symptoms selected should be submitted and further analyzed to calculate the probability of the disease.

References

A.Davis, D., V.Chawla, N., Blumm, N., Christakis, N., & Barbasi, A. L. (2008). Predicting Individual Disease Risk Based On Medical History.

Adam, S., & Parveen, A. (2012). Prediction System For Heart Disease Using Naive Bayes.

Al-Aidaroos, K., Bakar, A., & Othman, Z. (2012). Medical Data Classification With Naive Bayes Approach. *Information Technology Journal* .

Darcy A. Davis, N. V.-L. (2008). Predicting Individual Disease Risk Based On Medical History.

JyotiSoni, Ansari, U., Sharma, D., & Soni, S. (2011). Predictive Data Mining for Medical Diagnosis: An Overview Of Heart Disease Prediction.

K.M. Al-Aidaroos, A. B. (n.d.).

K.M. Al-Aidaroos, A. B. (n.d.). 2012. *Medical Data Classification With Naive Bayes Approach* .

NisharBanu, MA; Gomathy, B;. (2013). Disease Predicting System Using Data Mining Techniques.