Final Project Submission

By

Group 4

Dara Venkata Sai Sandeep-5
Khandelwal Shuchita-10
Podilli Venkata Krishna-23
Sawant Anuja-25

Introduction:

RoboCare – A human friendly application, one's own personal doctor. As its name suggests, this app takes good care of its user right from diagnosing the disease to suggesting the user the nearest practitioner who can prescribe medicines for his or her illness. All that it needs is few inputs from the user and it does all the work on its own. RoboCare with the help of the Smart watch and the user's Smart Phone provides its user excellent functionalities. Thus, RoboCare can be regarded as a guide to make a human's life illness free.

Architecture & Application:

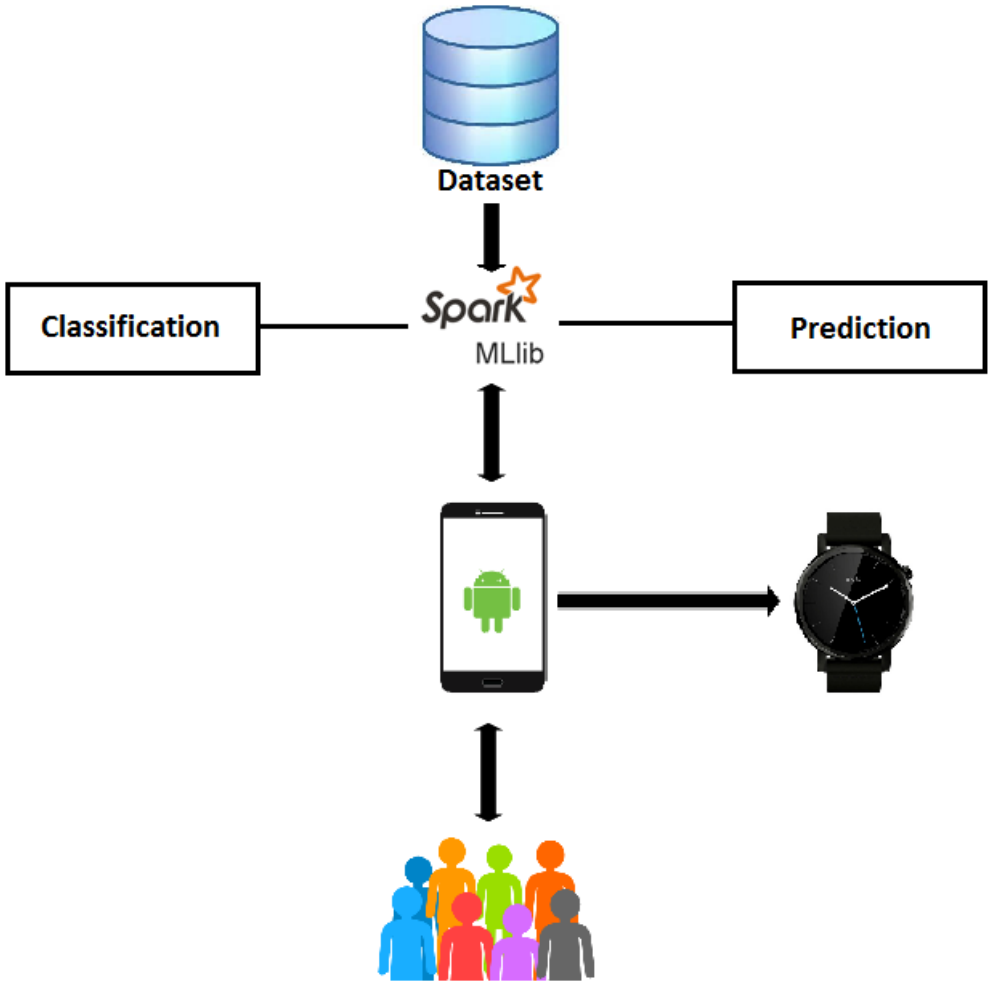

Figure 1.Architecture of RoboCare

The above diagram portrays the architecture of the overall application. The main components of the application are:

- Users: Provide symptoms of disease as input to the system in various formats and views the results given by the application.
- Android Phone: Acts as the front end of the application. Takes input from the user, passes it to the Spark System and provides results back to the user. Also sends notifications to the user's smart watch.
- Smart Watch: Displays notifications sent by the phone to the User.
- Spark Machine learning library: Takes input from the user via the Smart Phone and performs Naïve Bayes Prediction on the speech and text inputs and Random Forest Classification on the image inputs with the help of the training data provided to the system.
- Dataset: It consists of information about large number of diseases along with a list of symptoms of the diseases. Also consists of images of a few skin related diseases.

The workflow of the application goes as follows: The user registers himself by providing his or her necessary details. Then with the help of the credentials, the user logs into his or her account and then can make use of the multiple features provided by the system.

The user can provide symptoms to the 'Know your disease' function of the application via speech or text or also by taking images of the skin diseases and then the application predicts the nearest possibility of the disease. Then the user can also know about the medicines that can help to cure that.

The application also provides an Emergency button which helps the user to immediately call a person in urgency whose contact details he mentions during registration.

The application also provides the user with information about the doctors in the nearby areas who can help them cure that disease.

By making the most of all these functions the user logs out of the application.

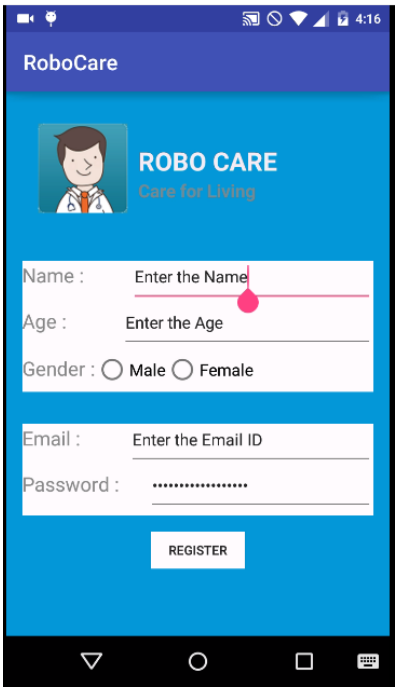

Figure 4.Registration Page

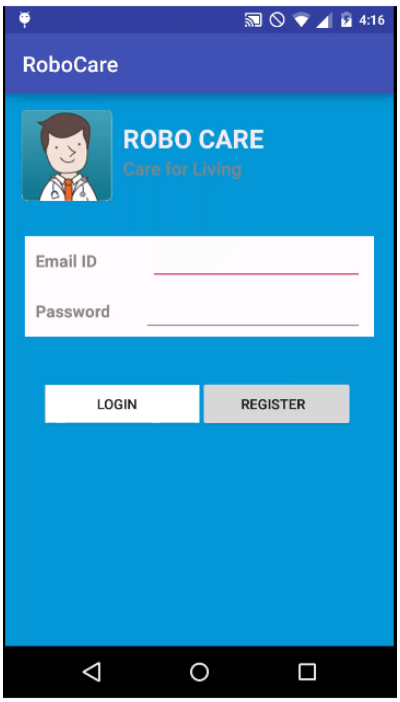

Figure 2.Login Page

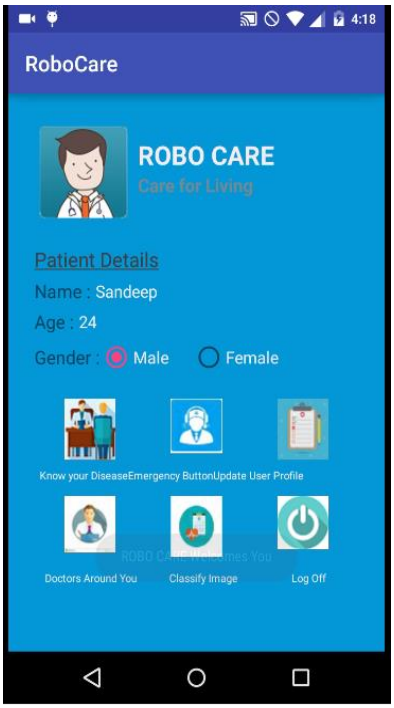

Figure 3.Home Page

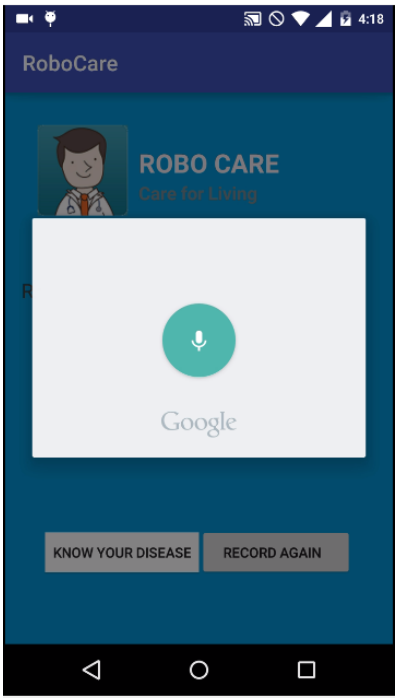

Figure 7.Speech Input

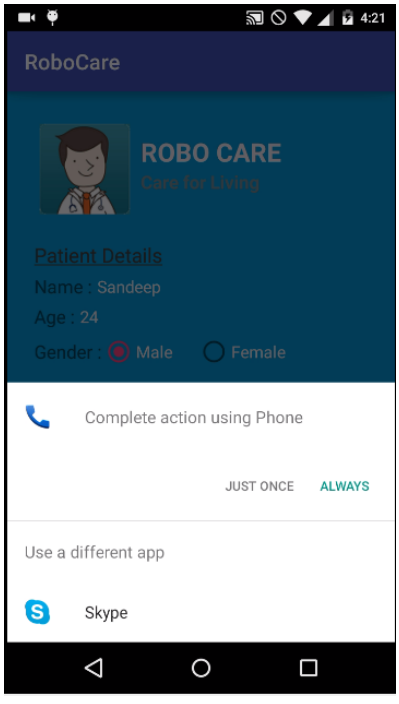

Figure 6.Emergency Call

Figure 5.Image input

Machine Learning:

Disease Prediction: Used Naïve Bayes Algorithm for the disease prediction in the application as it comes up with a rapid, highly scalable model development and recording. It makes use of parallelization as well. In spite of a small training set of data, Naïve Bayes algorithm provides efficient result. The accuracy attained is about 60%.

Image Classification: Random Forest is a collective classifier which contains many decision trees and generates the output class which is the general approach of the class's output by distinct trees. It provides a highly accurate classifier for the datasets. It is one of the effective method for estimating missing data and also maintains accuracy. The accuracy rate achieved is about 73%.

Data:

The dataset consists of comma separated value(.csv) files which entails information of various diseases and their symptoms. This information is gathered from a number of health related websites manually. Also, the image data used for image classification is of different skin diseases such as Eczema, Psoriasis etc. and has been collected from google images.

Evaluation:

The development of the entire application has been done by all the team members with equal dedication and co-operation amongst each other. The entire distribution of work is as follows:

Name: Dara Venkata Sai Sandeep
Class ID: 5
Participation: 31%
Tasks Accomplished:

- Project plot decision
- Worked on the front end of the application
- Developed a disease prediction system using Naïve Bayes Algorithm
- Yelp API for searching doctors
- M lib for data storage in cloud

Name: Khandelwal Shuchita
Class ID: 10
Participation: 23%
Tasks Accomplished:

- Development of the disease dataset
- Disease information functionality
- User authentication
- X code for Romo

Name: Podili Venkata Krishna
Class ID: 23
Participation: 23%
Tasks Accomplished:

- Worked on the front end of the application
- Emergency button functionality
- Speech to text functionality

Name: Sawant Anuja Ajay
Class ID: 25
Participation: 23%
Tasks Accomplished:

- Development of the disease dataset
- Image input using mobile camera
- Image classification using Random Forest Algorithm

Accuracy:

Precision:
The disease prediction provided using the Naïve Bayes Algorithm is about 60% accurate. The accuracy is achieved as a result of efficient functioning of the Naïve Bayes algorithm and the distinct dataset provided to it. An accuracy of 73% for image classification is attained by using the Random Forest Algorithm for the same. The amount of such accuracy is obtained by accurate decision tree generation by the algorithm and discrete training data of images.

Related Work (references):

- http://spark.apache.org/docs/latest/mllib-naive-bayes.html
- http://www.yelp.com/kansas-city-ks-us
- https://www.nih.gov/
- http://archive.ics.uci.edu/ml/
- https://www.google.com/imghp?hl=en&tab=wi&ei=ougvV4u3OYSEmQGMppCICg&ved=0EKouCBYoAQ
- www.medicinenet.com/symptoms_and_signs/article.htm#introView
- http://www.mayoclinic.org/diseases-conditions/index?letter=A

Future Work:

The application which has been developed efficiently predicts the disease that the user might be suffering from with the help of the symptoms provided by the users. But the application works on the basis of different datasets for the symptoms provided in text and image formats. This application can be upgraded to a better version wherein both the inputs rely on same dataset for classification and then prediction of disease.