

## Smart Health Guide App

Alison Almeida<sup>1</sup>, Tanisha Sangodkar<sup>2</sup>, Sahil Naik<sup>3</sup>, Yash Sangodkar<sup>4</sup>, Shreedatta Sawant<sup>5</sup>,  
Diksha Prabhu Khorjuvenkar<sup>6</sup>

<sup>1,2,3,4</sup>Department of Computer Engineering Agnel Institute of Technology and Design, Assagao, Goa, India

<sup>5,6</sup>Assistant Professor, Department of Computer Engineering Agnel Institute of Technology and Design, Assagao, Goa, India

\*\*\*

**Abstract** - Health is a very important concern in today's generation and the first step is following the right diet. Many people suffer from many diseases due to improper consumption of food. The root cause of this can be the lack of knowledge regarding the nutrients available in the food product. Some of the common diseases that the majority of the people suffer from are diabetes, cholesterol and jaundice. These diseases are caused when certain nutrients in a food product are consumed in excess. We have taken these three in consideration and created an application that will help the people to know the contents present in the food item and guide them if they should be consumed by them or not with respect to their health condition. An application that will allow the user to provide the disease he is suffering, allowing him to scan the food products barcode and guiding him if that food product is suitable for consumption according to his health condition.

**Key Words:** Data Mining, Decision Tree, Entropy

### 1. INTRODUCTION

The process of finding anomalies, trends and correlations in massive data sets using techniques of machine learning, statistics and database systems to anticipate outcomes is known as data mining.

A root node, branches, and leaf nodes make up the structure of a decision tree. It is a supervised learning algorithm that can work on both discrete and continuous variables. It follows a top-down approach as the top region presents all the observations at a single place which splits into two or more branches that further split. A test on an attribute is shown by each internal node, the result of a test is indicated by each branch, and a class label is contained in each leaf node. The highest impure node within the tree is the root node. It splits the dataset into subsets on the basis of the most significant attribute in the dataset. How the decision tree identifies this attribute and how this splitting is done is decided by the algorithms.

Entropy is defined as the randomness or measuring the disorder of the information being processed in Machine Learning. In other words, it measures the impurity in that

particular system. When information is handled within the system, each piece of information has a specific value and can be used to reach conclusions. Therefore, if it is easier to draw a valid conclusion from information, then the entropy will be lower in Machine Learning, or if the entropy is higher, then it will be hard to draw a conclusion from that information.

### 2. LITERATURE SURVEY

#### 2.1 Decision Tree Algorithms

Process of extracting useful knowledge from a huge set of incomplete, noisy, fuzzy and random data is called data mining. Decision tree classification technique is one of the most popular data mining techniques. In decision tree divide and conquer technique is used as a basic learning strategy. A decision tree is a structure that includes a root node, branches, and leaf nodes. Each internal node denotes a test on an attribute, each branch denotes the outcome of a test, and each leaf node holds a class label. The topmost node in the tree is the root node. This paper focuses on the various algorithms of Decision trees (ID3, C4.5, CART), their characteristics, challenges, advantages and disadvantages.[1]

#### 2.2 Analysis of Decision Tree Algorithms

In Data mining there are various techniques available to mine large amounts of data. The data mining techniques are Classification, Association rule mining, cluster analysis, outlier analysis. In this data mining techniques classification technique is the most powerful technique to classify our large amount of data. In this classification, the Decision trees are the most prominent classification technique in data mining. Decision tree induction is the learning of decision trees from class labeled training tuples. A decision tree is a flow chart like tree structure. Decision trees have been used to solve a wide variety of implications in the data mining process. This is used in statistics, machine learning. Data mining is also a predictive model. Decision trees are classified by using the two phases, first phase is tree building phase and the second phase is tree pruning phase. From this paper we

learnt more on data mining decision tree algorithms like ID3, C4.5, J48, CART and the performance analysis of these algorithms.[2]

### 2.3 Data Normalization and Standardization

This paper aims to clarify how and why data are normalized or standardized, these two processes are used in the data preprocessing stage in which the data is prepared to be processed later by one of the data mining and machine learning techniques like support vector machine, neural network, etc. The two methods try to scale the data set. These two processes are helpful in some cases and necessary in some other cases, most of the data mining and machine learning tools include these two preprocessing techniques like in Weka or in MATLAB. This paper will simply define and present the use of these two data preprocessing techniques.[3]

### 2.3 Web Server

The HTTP Web Server is a method which runs on a machine and listens for incoming HTTP Requests by a specific IP and Port number and then sends back a response for that particular request. Python library has an in-built webserver that can be invoked for a simple web client server communication. This website gave an insight on how HTTPRequestHandler is used to create a custom handler.[4]

## 3. EXISTING SYSTEM

The traditional way in which people shop for food products is without checking the nutritional information of that particular product. This may also include some nutrients in that product that may be harmful to the user's current health condition. Consuming that product in excess can make his current health condition worse.

## 4. PROPOSED SYSTEM

The application will be installed on the user's phone which will allow him to scan the barcode of the food products while shopping. After scanning the products barcode, it will show if the product is suitable for consumption according to the health condition which he has given and accordingly the user can purchase that particular food product.

### 4.1 Overview

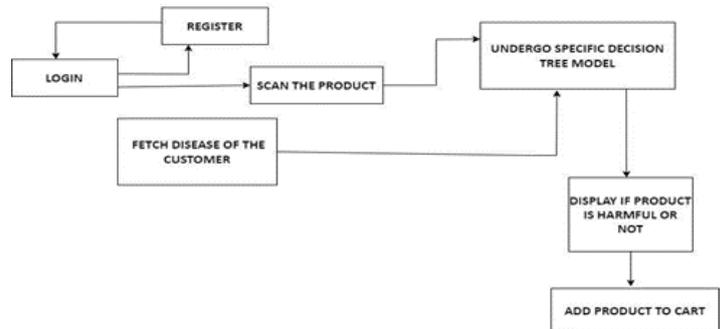


Fig. 1. Entire Process

The application will function as follows:

- The user will provide the required information through the registration process.
- If user is already registered, then he will just login and he will be directed to the scan page
- On the scan page the user will scan the products barcode. After scanning the products barcode, the products name and its price will be displayed. Also, it will display if the product is consumable or not with respect to health details provided during registration.

### 4.2 Overall look

The proposed project consists of a database designed in Firebase, an application, a web server and a website that have various functions related to the user using it.

The website is mostly handled by an authenticated admin and the application will be used by the customers. Firebase is used since it provides a real-time database that helps to store and synchronize data, provides authentication service and allows the content to be stored with ease. If the user forgets his password, then firebase sends an email containing a link. The user can click on the link and reset his password. After resetting the password, the user needs to sign in again.

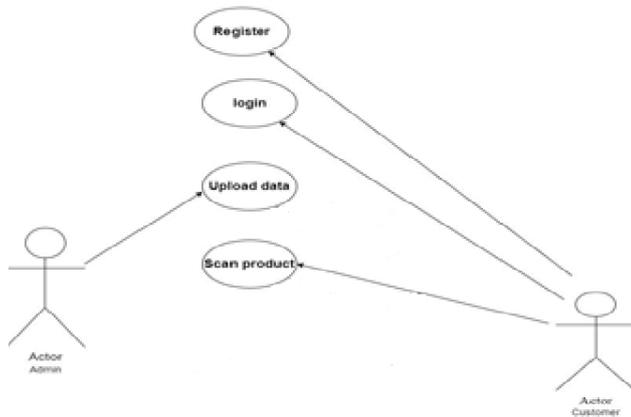


Fig. 2. Use Case Diagram

The above use case diagram shows the interaction.

The user through Signup will provide information like username, name, email-id, age, gender and the disease which includes diabetes, cholesterol and jaundice.

The authenticated admin needs to login in order to perform various operations. The basic login/logout functionalities are also given to the admin. After the admin login, he can perform the following functionalities:

- Add: Adding a new product to the database.
- Display: Displaying all the information of the product.
- Update: Updating information of a particular product.
- Delete: Deleting a product from the database.

The model consists of the following steps:

- Importing the necessary libraries.
- Fetching the data
- Separating data in target and independent variables.
- Normalizing the input data.
- Converting NumPy array to data frame.
- Splitting the data into test and train.
- Fitting the decision tree model.
- Getting credential access by using a json key.
- Initializing the application
- Getting reference to the specified barcode path.
- Fetching the values from firebase.
- Converting dictionary to data frame.
- Normalizing result on the raw input.

- Predicting the new value.

## 5. IMPLEMENTATION AND RESULT

The images below are the screenshots from the output we receive:

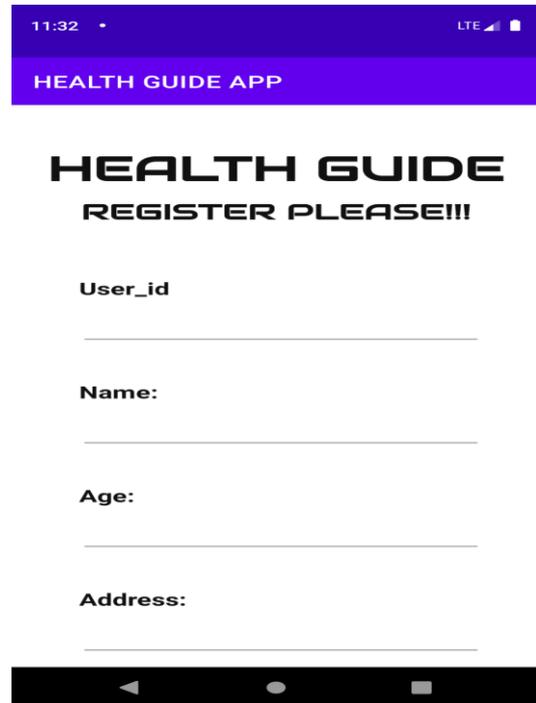


Fig. 3. User Registration

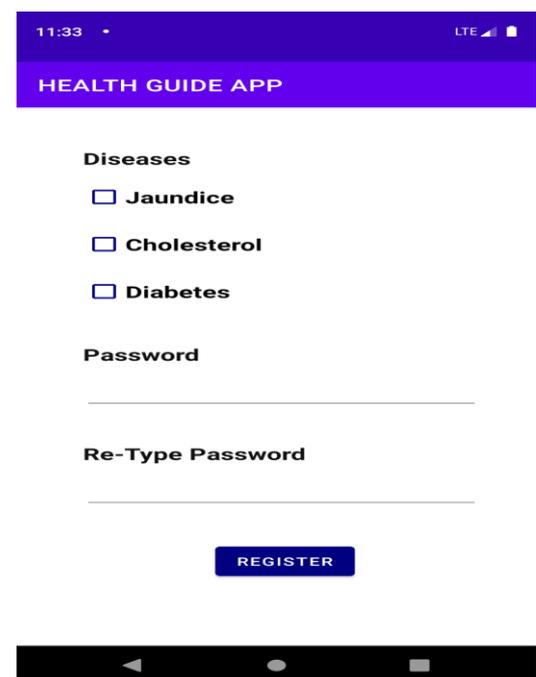


Fig. 3.1. User Registration



Fig. 4. Scanning Barcode

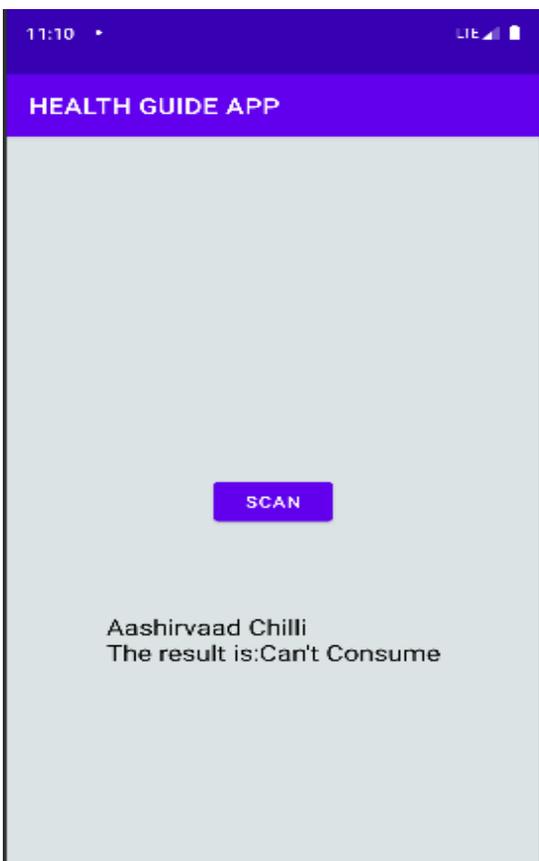


Fig. 5. Result

## 6. CONCLUSION

The project proposed an application that will help the customer in selecting a product that is healthy to him with respect to the disease he is suffering from. Through this he can be more alert whenever he is shopping for a food product.

This project can be further improved by adding payment gateway, adding security features which improves the authenticity and also showing the contents and percentage up to which the content can be consumed. Also, this app can be modified with respect to certain activities like gym which can help to track the calories and also in the medical field that will allow users to check for the contents of the medicine.

## REFERENCES

- [1] Sharma, H., & Kumar, S. (2016). A survey on decision tree algorithms of classification in data mining. *International Journal of Science and Research (IJSR)*, 5(4), 2094-2097.
- [2] Priyam, A., Abhijeeta, G. R., Rathee, A., & Srivastava, S. (2013). Comparative analysis of decision tree classification algorithms. *International Journal of current engineering and technology*, 3(2), 334-337.
- [3] Ali, P. J. M., Faraj, R. H., Koya, E., Ali, P. J. M., & Faraj, R. H. (2014). Data normalization and standardization: a technical report. *Mach Learn Tech Rep*, 1(1), 1-6
- [4] tutorialspoint. Python - HTTP Server. Tutorials Point. Retrieved June 29, 2022, from [https://www.tutorialspoint.com/python\\_network\\_programming/python\\_http\\_server.htm](https://www.tutorialspoint.com/python_network_programming/python_http_server.htm)