iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** IV **Month of publication:** April 2018

DOI: http://doi.org/10.22214/ijraset.2018.4394

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

Virtual Dietitian using Machine Learning

Ameya Parab[1], Akshaykumar Lilani[2], Jayesh Motwani[3], Charusheela Nehete[4]
[1, 2, 3, 4] *Department of Information Technology, Vivekanand Education Society's Institute of Technology, University of Mumbai*

Abstract: ***The Virtual Dietitian is a Web Application to suggest a diet plan to its users. It will act as a diet consultant like a real dietitian. A person, in order to know his/her diet plan needs to give some information to the dietitian such as its body type, weight, height, daily exercise, lifestyle and body goal details. In a similar way, this application will interact with the user as per the information entered by the user. The Application will query all this data to the user and process it to provide the appropriate diet plan to the user. Thus, the user can get the required diet plan by querying the application and the user does not need to visit any dietitian. The user will need to register and make an account wherein the user will be required to enter details, only then he/she will get a generated personalized diet plan, as there would be different diet plans for different needs of the user. In case the user is not satisfied the application will also give an alternative diet plan depending on the user feedback. Hence, the application will give immediate & personalized diet plans, and the application will have user autonomy.***
Keywords: ***Data Mining, Machine Learning, Random Forest, System design, Fitness***

I. INTRODUCTION

According to a special article that was written by Princeton University, the per capita calorie intake in India is declining, as is the intake of many other nutrients; indeed, fats are the only major nutrient group whose per capita consumption is unambiguously increasing. Today, more than three-quarters of the population live in households with per capita calorie consumption below 2,100 per day in urban areas and 2,400 per day in rural areas – numbers that are often cited as "minimum requirements" in India.

Nutrition is vital for your body and all of its systems to function properly, by having good nutrition it will help you maintain a healthy weight, reduce body fat, provide your body with energy, promote good sleep and generally make you feel better. By having good nutrition, it has been proven that you are less likely to develop many of the present-day diseases.

When we think about dieting, most of the people think about setting a weight-loss goal that will determine how long we watch what we're eating. However, people these days are focused on gaining and maintaining their body weight and staying fit. A better way to think about it might be to ask yourself the question:

How long do I want to try to avoid developing a chronic disease? Put in those terms, it's easy to see that getting control of BMI and eating foods that contribute to continuing good health (or that don't directly contribute to the development of dangerous disease conditions) is not a short-term goal.

For many of us, doing what it takes to get our BMI into a healthy range and to keep it there means learning to live our lives in a new way.

II. OBJECTIVES

A. Intake of Healthy foods

Improve nutrition and health outcomes of vulnerable segments (children and women) of the populations, through the availability of foods that would increase intake of vegetables and fruits, decrease caloric intake and increase micronutrient intake.

B. Personalised Diets

Clients receive personalized diets made exclusively for them to increase its effects.

C. Nutritional Awareness

To increase the health and Nutrition awareness among the population.

III. SYSTEM DESIGN

A. Architecture

The architecture of the Virtual Dietitian System consists of 5 layers, they are Presentation layer, Business Layer, Service Layer, Data Access Layer and the Data layer. Following are the various layers of the System Architecture:

1) *Presentation Layer:* It consists of application user interface which is used to take the necessary information from the user and display the diet plan as well as take the feedback from the user.
2) *Business Layer:* It acts like middleware for the application and consists of the Mifflin - St. Jeor Calorie Calculator and various constraints and rules used for building a diet plan for the user.
3) *Service Layer:* This layer provides the classification model for the diet plans depending on the category of users.
4) *Data Access Layer:* The layers above this layer can access the data stored in the database through this layer.
5) *Data Layer:* It contains the databases namely User database, Food and Nutrition database, and the Diet Plan database.

B. Flow Chart

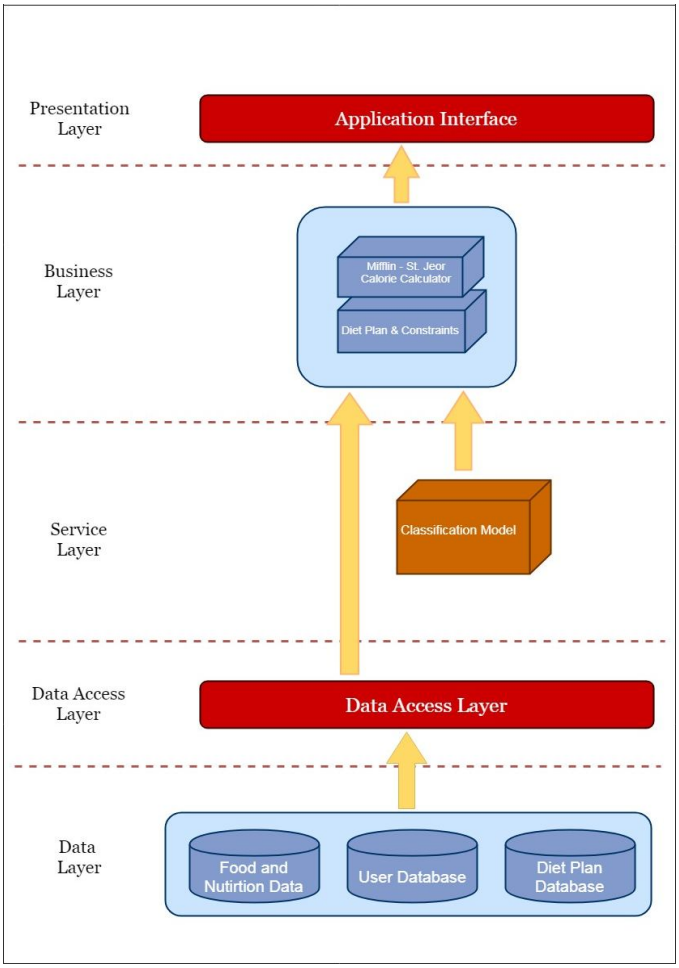


Fig. 1 System Architecture

IV. IMPLEMENTATION

This is a web application. Technologies required to develop are Python, Django (python web framework), Bootstrap, Python Machine learning libraries and sqlite3 database. The front-end part of the application will use Bootstrap and the Django views and templates. The backend will use sqlite3, Machine learning libraries of python, Django views, Django models, Django forms, etc.

The basic flow of the application is as, first the user is authenticated, the authenticated user is then prompted to fill a form which will be used to extract the diet plan. After the user has submitted the data prompted, the system then generates the diet plan which will be specific to the user. The generated diet plan is then showed to the user.

The system asks user to comeback after a month to the system in order to tell the system whether, the user is satisfied with the results. This feedback which is taken from the user, will then be used to improve the accuracy of the algorithm.

The goal (lose weight, gain weight or stay fit) of the user is asked in the form. In order to achieve the goal, one needs to increase or decrease the calorie intake depending upon what the goal is. The other nutrients such as, protein, carbohydrates, fats (saturated or unsaturated), etc also play a vital role in deciding what a person must eat in order to achieve the goal. This whole decision process

of generating the diet plan is done by the Application. User is also asked about his preferences in the form (which was prompted at the start), veg, non-veg or veg with eggs.

The starting point of this application is the calculation of the number of calories which the user must take in order to achieve the goal. To calculate the calorie intake, the same form asks user its height, weight, goal, preferences, lifestyle and profession. Not all the attributes mentioned formerly are used to calculate the calorie intake. Also, the calorie is calculated using Miffin – st. Joer Calorie Calculator.

Machine learning is implemented using the feedback loop. Loop because, every time user opts for a diet plan, the diet plan will be generated for a certain period of time. In this time period, user will use the diet plan. After this time period has expired, the user will be prompted by the system to input the feedback of the diet plan which was given to him/her last time i.e. how well the diet plan performed? This feedback will be used to update the training data set. Positive as well as negative feedback will be handled to improve the performance of the system.

A. Data Collection And Pre-Processing

A survey form was shared among 300 individuals who follow a diet plan. The data collected using this survey is used as training data. Attributes such as height which is required in cm are to be converted to cm if the value given isn't in cm. Only the required features such as height, weight, age, gender, lifestyle, and fitness goal are used for constructing classification model. Data taken from the user for generating the diet plan is the test data. The data which is obtained is categorical in nature. However, in python, unlike R, there is no option to represent categorical data as factors. Hence, the categorical data is converted to numerical data using pandas Series in python.

B. Calorie Calculator

Initially the data such as height, weight, age, gender, lifestyle, profession and fitness goal is taken from the user. For obtaining a diet plan required amount of calorie intake per day is calculated using calorie calculator.

The Calorie Calculator is based on the Mifflin-St Jeor Equation which calculates basal metabolic rate (BMR), and its results are based on an estimated average. The basal metabolic rate is the amount of energy expended per day at rest. There exist other equations including the Harris-Benedict Equation which was used until 1990 before the introduction of the Mifflin-St Jeor Equation and the Katch-McArdle Formula which calculates resting daily energy expenditure (RDEE) by taking lean body mass into account. The Mifflin-St Jeor Equation is considered the most accurate equation for calculating BMR, with the exception that the Katch-McArdle Formula can be more accurate for people who are leaner that know their body fat percentage. The Mifflin-St Jeor Equation is as follows:

1) Calculation of BMR

a) BMR FOR MEN = 10 X WEIGHT (KG) + 6.25 X HEIGHT (CM) – 5 X AGE (Y) + 5

b) BMR for Women = 10 x weight (kg) + 6.25 x height (cm) – 5 x age (y) – 161

2) Calculation of Activity factor

a) Sedentary = BMR X 1.2 (little or no exercise, desk job

b) Lightly active = BMR X 1.375 (light exercise or sports 1-3 days/week)

c) Moderately active = BMR X 1.55 (moderate exercise or sports 3-5 days/week)

d) Very active = BMR X 1.725 (hard exercise or sports 6-7 days/week)

e) Extremely Active = BMR X 1.9 (hard daily exercise or sports & physical labor job or 2 X day training, football camp, etc.)

3) Calculation of Calorie intake per day

a) *Staying fit:* For staying fit the calorie which is obtained after the calculation of activity factor is considered.

b) *Gaining weight:*

i. For gaining 1 kg per week, Calorie intake = Calorie obtained after calculation of activity factor + 1000

ii. For gaining 0.5 kg per week, Calorie intake = Calorie obtained after calculation of activity factor + 500

c) *Losing Weight:*

i. For losing 1 kg per week, Calorie intake = Calorie obtained after calculation of activity factor – 1000

ii. For losing 0.5 kg per week, Calorie intake = Calorie obtained after calculation of activity factor – 500

C. Methodology

After pre-processing a classification model is constructed using Random Forest classifier. The Random Forest is appropriate for high dimensional data modelling because it can handle missing values and can handle continuous, categorical and binary data which is suitable for this dataset. The bootstrapping and ensemble scheme makes Random Forest strong enough to overcome the problems of over fitting and hence there is no need to prune the trees. Besides high prediction accuracy, Random Forest is efficient, interpretable and non-parametric for various types of datasets the training data and then the results are obtained for test data.

The comparison of four classification models has been done i.e. Random Forest classifier, Gaussian I Bayes classifier, and Decision Tree classifier using Gini index and information gain. The size of the training data is up to 300 instances and the test data used for comparison of classifiers is 100 instances. The nodes of the tree are expanded till a maximum depth of 10 and accuracy of each model is calculated. The observed accuracy for the models under comparison are shown in Table 1.

TABLE I
ACCURACY OF CLASSIFICATION MODELS

Classifier	Accuracy
Random Forest Classifier	0.9154
Decision Tree Classifier (gini index)	0.6769
Decision Tree Classifier (information gain)	0.6346
Gaussian Naïve Bayes Classifier	0.5615

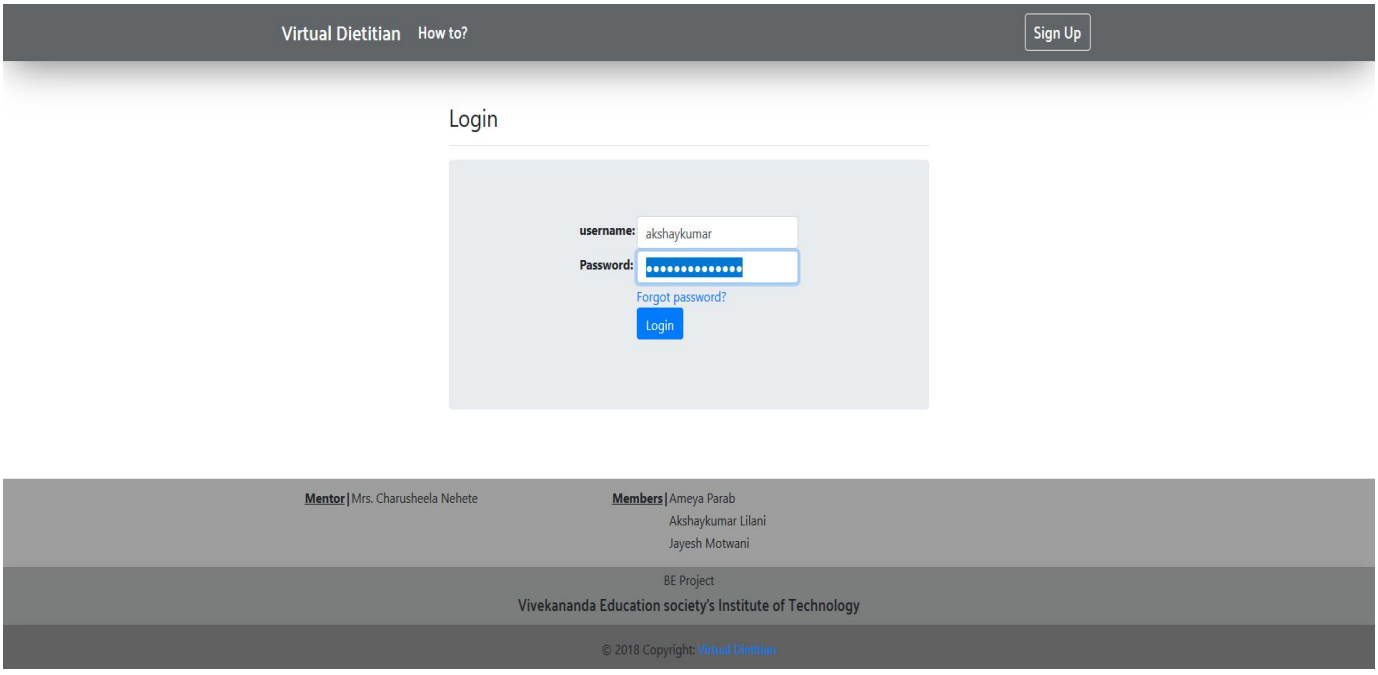
The diet plan is suggested to the user as per results of the classifier on the test data which is the data obtained from the user. Depending on the duration for which the diet plan has to be followed by the user, a feedback is taken from the user after the specified duration of time. The feedback taken is the data about weight of the user, which is compared to the previous weight of the user before suggesting the diet plan and result is obtained i.e. the diet plan which was suggested is suitable for that type of user (positive feedback) or it didn't work for that user (negative feedback).

In case of the positive feedback, the user data which was previously considered as test data for obtaining a diet plan is now added to the training data and the model is reconstructed every time the feedback based on the new training data.

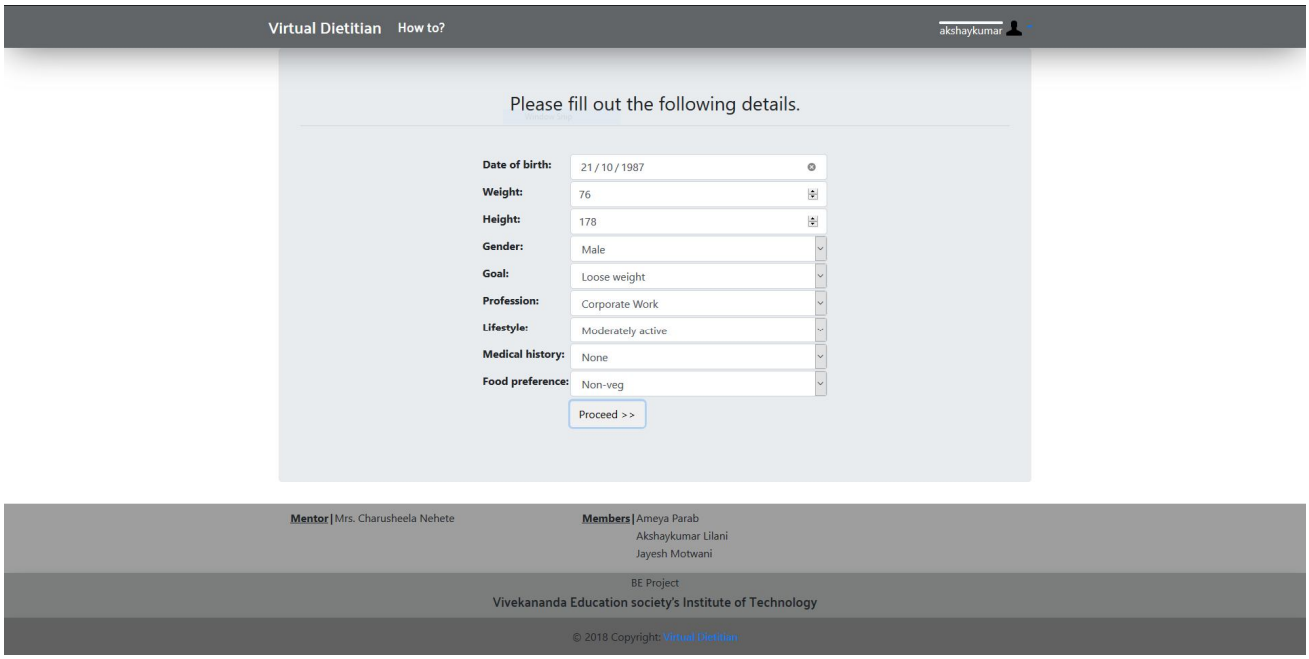
V. RESULTS

Following are the screenshots, of the process of generation of the Dietplan via the Virtual Dietitian

A. Login

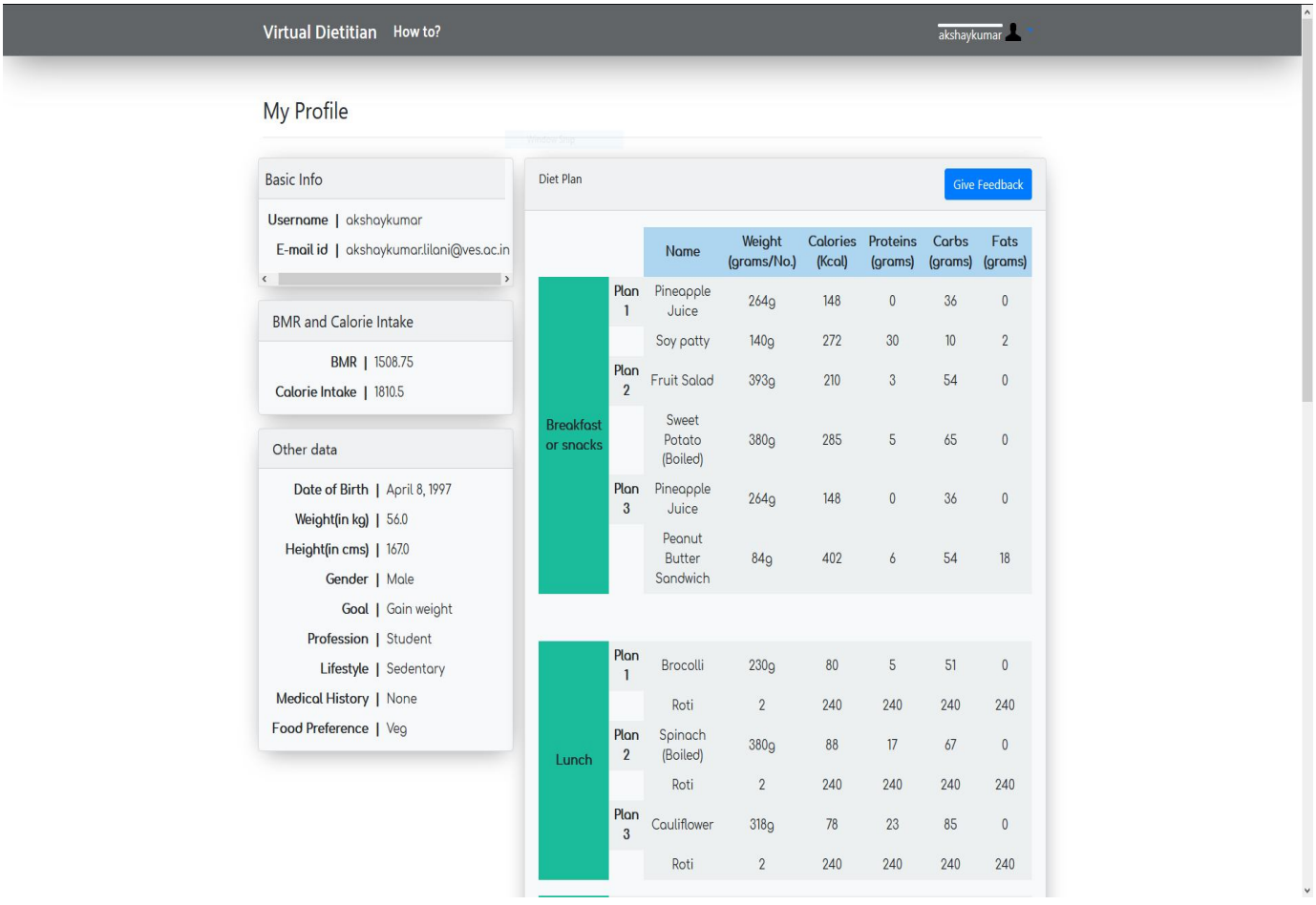


B. Filling the form which asks for the details required to generate Plan

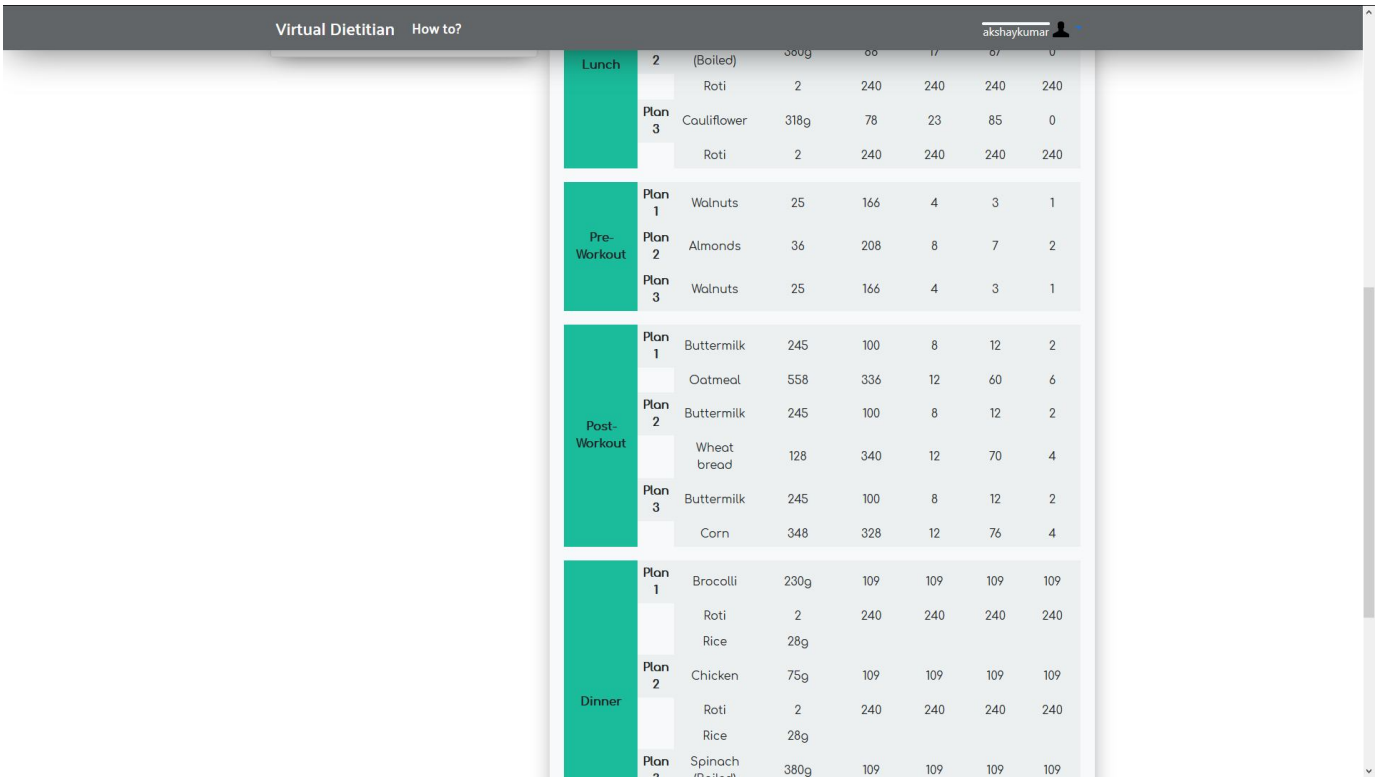


C. Going to My profile page to see the plan

1)

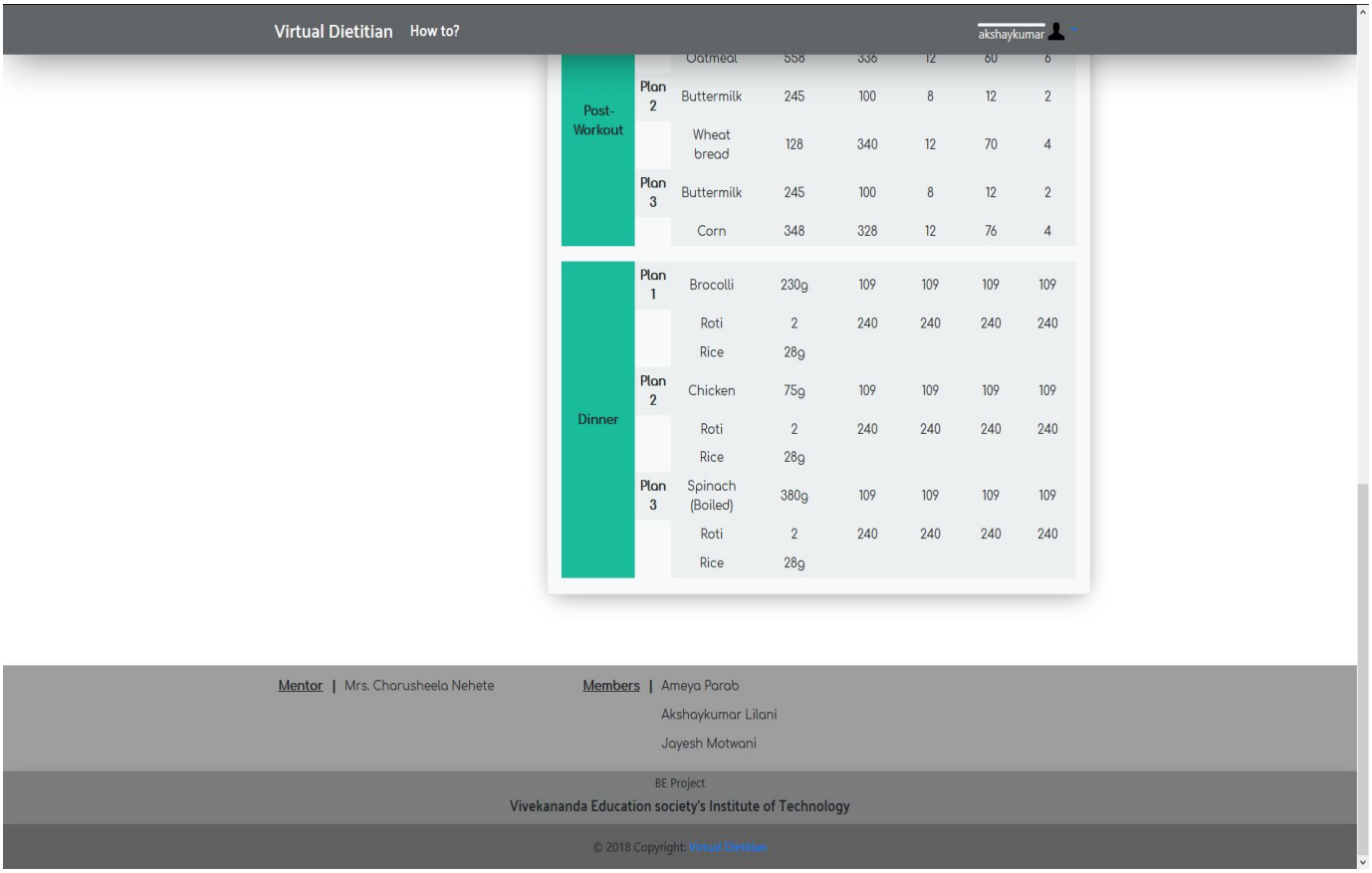


2)



Lunch	2	(Boiled)	380g	88	17	87	0
		Roti	2	240	240	240	240
	Plan 3	Cauliflower	318g	78	23	85	0
		Roti	2	240	240	240	240
Pre-Workout	Plan 1	Walnuts	25	166	4	3	1
	Plan 2	Almonds	36	208	8	7	2
	Plan 3	Walnuts	25	166	4	3	1
Post-Workout	Plan 1	Buttermilk	245	100	8	12	2
		Oatmeal	558	336	12	60	6
	Plan 2	Buttermilk	245	100	8	12	2
		Wheat bread	128	340	12	70	4
	Plan 3	Buttermilk	245	100	8	12	2
		Corn	348	328	12	76	4
Dinner	Plan 1	Brocolli	230g	109	109	109	109
		Roti	2	240	240	240	240
		Rice	28g				
	Plan 2	Chicken	75g	109	109	109	109
		Roti	2	240	240	240	240
		Rice	28g				
	Plan 3	Spinach (Boiled)	380g	109	109	109	109

3)

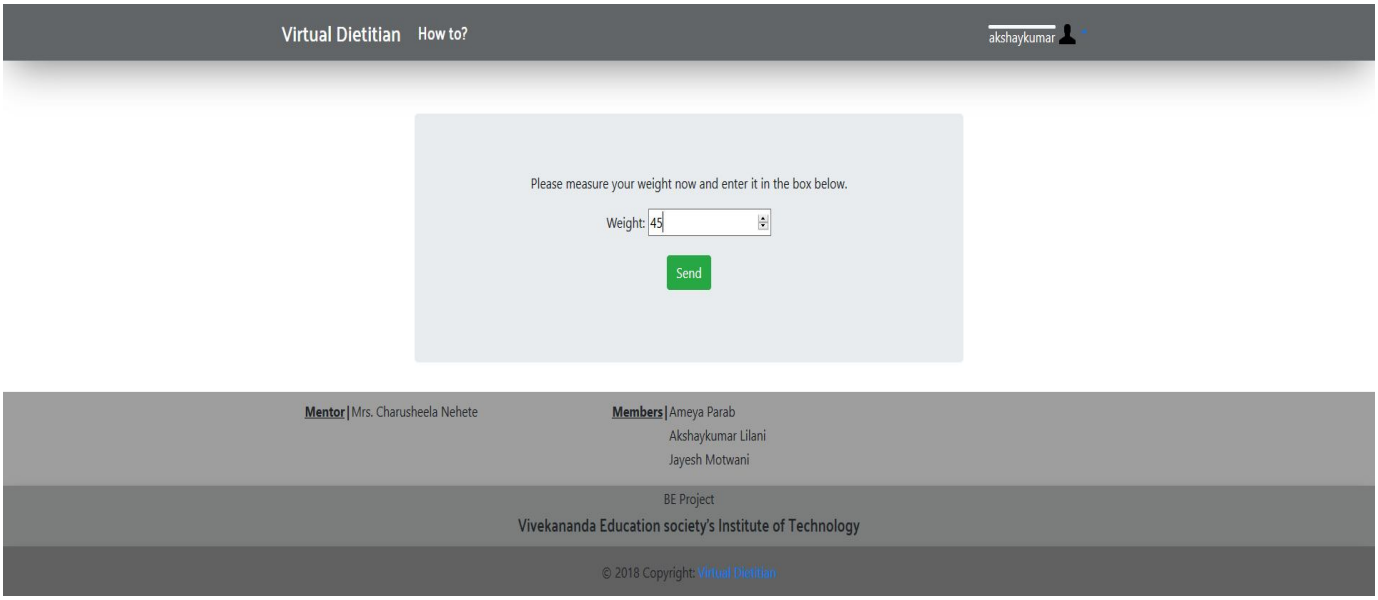


Post-Workout		Oatmeal	558	336	12	60	6
	Plan 2	Buttermilk	245	100	8	12	2
		Wheat bread	128	340	12	70	4
	Plan 3	Buttermilk	245	100	8	12	2
		Corn	348	328	12	76	4
Dinner	Plan 1	Brocolli	230g	109	109	109	109
		Roti	2	240	240	240	240
		Rice	28g				
	Plan 2	Chicken	75g	109	109	109	109
		Roti	2	240	240	240	240
		Rice	28g				
	Plan 3	Spinach (Boiled)	380g	109	109	109	109
		Roti	2	240	240	240	240
		Rice	28g				

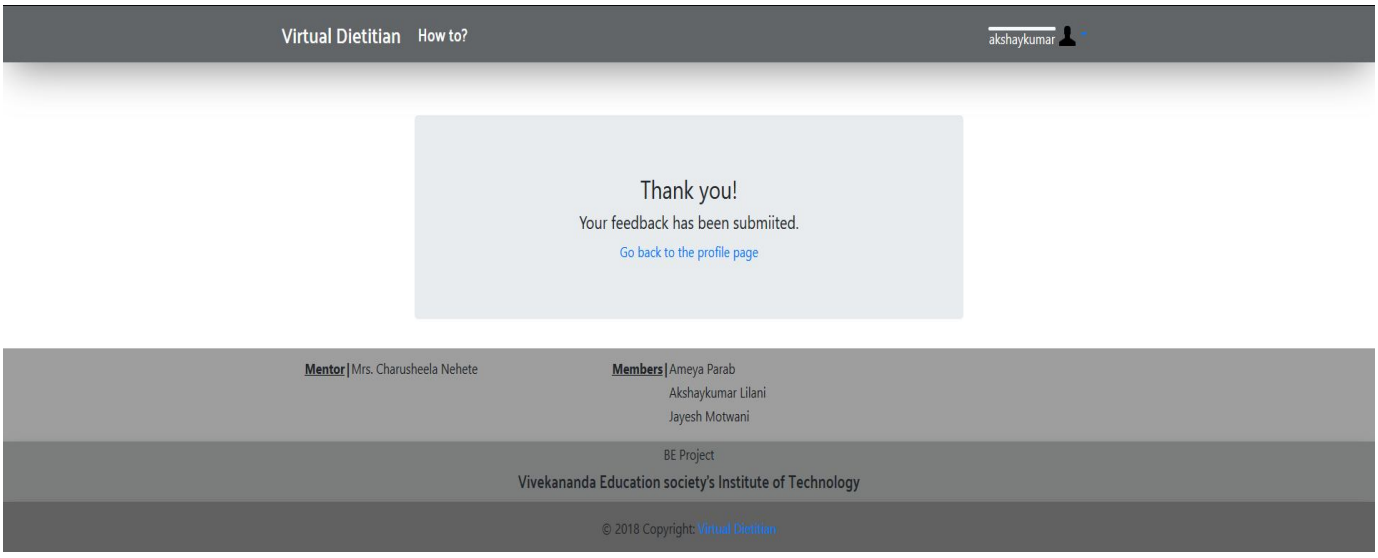
D. Giving feedback

After clicking on the 'Give feedback' button in my profile section, user lands on the following page and is asked to fill in his/her current weight. The algorithm then decides whether feedback is to be considered positive or Negative based on the goal chosen by the user at the time of generation of diet plan.

1)



2)



VI. CONCLUSIONS

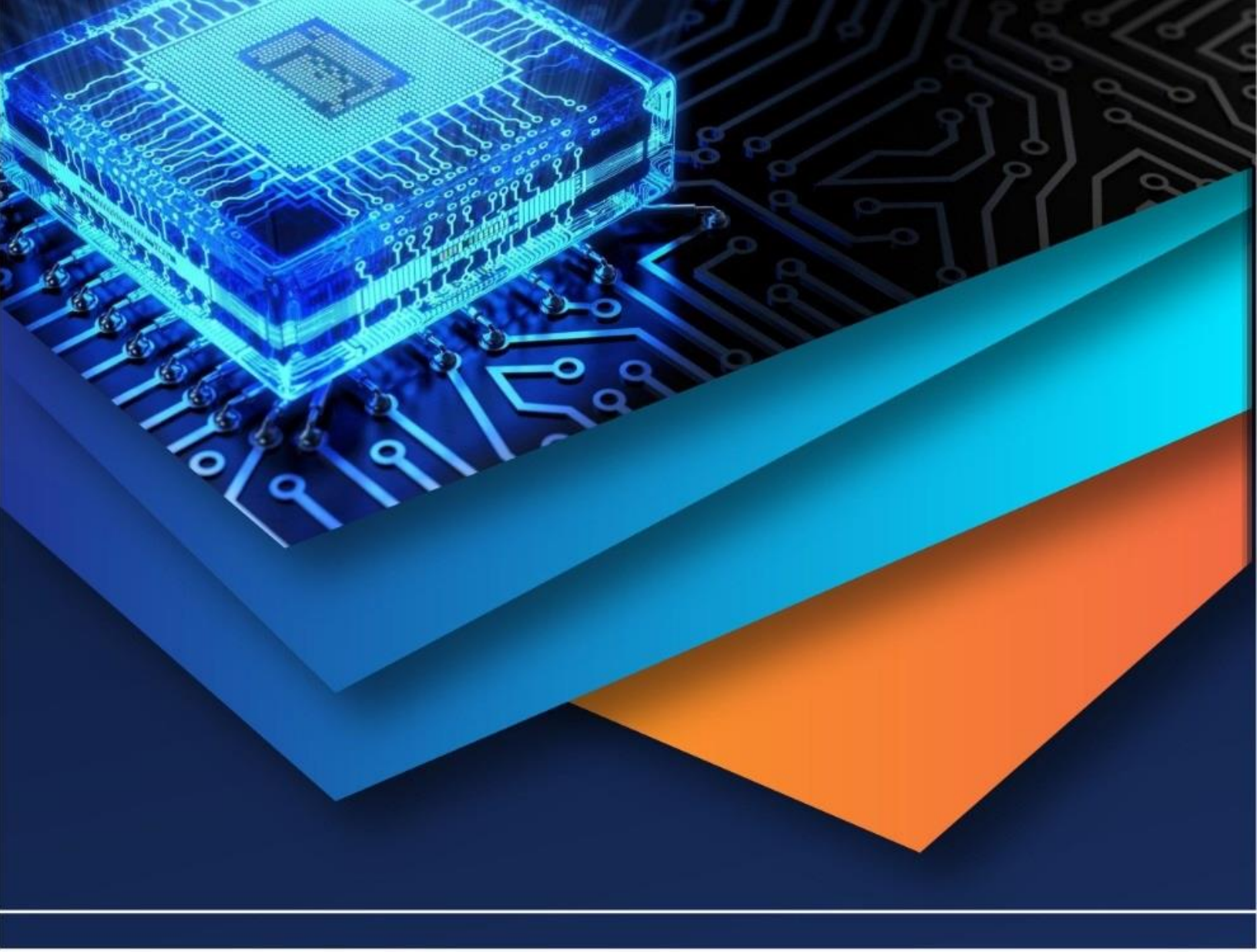
In this paper, the process of developing this application is studied. Ways in which the diet plan will be generated is explained. The algorithm which will be used to generate the classification is random forest. Why this algorithm is better than its competitors is described. Machine learning is implemented using the feedback loop. Based upon the feedback, positive or negative, the training dataset of the application will be updated so as to improve the performance of the system.

VII. FUTURE SCOPE

A. Accuracy of the predictions can be improved by optimizing feedback algorithm.
B. System can be enchanced so that it supports more diseases.
C. A feature to replace items from the diet plan can be provided.

REFERENCES

[1]. Jehad Ali, Rehanullah khan, Nasir Ahmad, Imran Maqsood, *"Random Forests and decision trees"* IJCSI International journal of Computer science issues. Vol 9, issue 5, No. 3, September 2012.
[2]. Andy Liaw, Matthew Wiener, "Classification and Regression by Random Forest" R News, ISSN 1609-3631. Vol. 2/3, December 2002.
[3]. Y. Claire Wang, MD, ScD, Amber Hsiao, MPH, C. Tracy Orleans, PhD, Steven L. Gortmaker, PhD," Average *Caloric Impact"* August 2013.
[4]. Mifflin, M. D., S. T. St Jeor, et al. (1990). *"A new predictive equation for resting energy expenditure in healthy individuals."*
[5]. Government of Canada. Nutrients Data. [Online]. Available link:https://www.canada.ca/en/health-canada/services/food-nutrition/healthy-eating/nutrient-data.html
[6]. Health Wire, Authority Nutrition. [Online]. Available link: **https://www.healthline.com/nutrition**
[7].myfitnesspal [Online]. Available link **https://www.myfitnesspal.com/**



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY