C-DAC Chennai

8th floor, D Block, TIDEL Park, No. 4, SH 49A,

Tharamani, Chennai,Tamil Nadu - 600113

Project Report

on

Diet Recommendation System

Submitted to ACTS,C-DAC Chennai

by

Piyush Soni (220960825031)

Project Guide: Mrs. S.Irene

In the partial fulfillment of

Post Graduate Diploma in Big Data Analytics

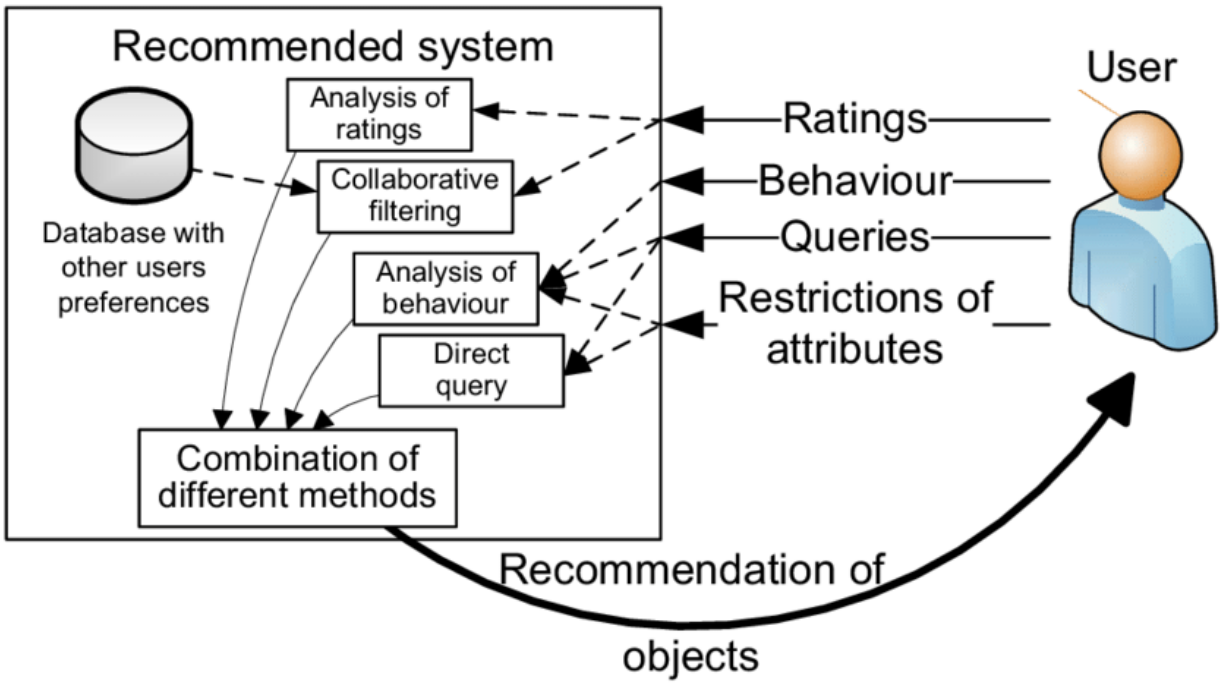
September 2022 – March 2023

ABSTRACT-

The food you eat every day affects your health - how you feel today, tomorrow, and in the future. Good nutrition is an important part of a healthy lifestyle. Combined with physical activity, diet can help you reach and maintain your goals. Healthy weight, reduced risk of chronic diseases (such as heart disease and cancer). and promote your overall health. A balanced diet is what gives your body Nutrients necessary for normal functioning. The number of calories in food is .It measures the amount of energy stored in that food. your body burns calories From food to walking, thinking, breathing and other vital functions. of The average person needs to consume about 2,000 calories a day to maintain energy levels. weight. However, a person's specific her daily caloric intake may vary depending on the condition. Age, gender and level of physical activity. Men generally need more calories Women and those who exercise need more calories than those who do not.

1. INTRODUCTION-

One of the most important factors for a healthy life is your daily diet and nutrition. People suffering from mild or severe illness eHealth Initiative and Research activities aim to provide novice end-users with a variety of ubiquitous applications improve their health.Various studies have shown that this is inappropriate and insufficient Food intake is a major cause of various health problems and diseases. the study According to estimates carried out by the World Health Organization (WHO), The total population of the world suffers from various diseases and 60% of deaths Associated with malnutrition in children every year. Another WHO study reported Inadequate and unbalanced food intake causes about 9% of heart attacks death,approximately 11% of ischemic heart disease deaths, 14% of gastrointestinal deaths Deaths from cancer worldwide. In addition, about 250 million children are suffering. Of those with vitamin A deficiency, 200 million suffer from iron deficiency (anemia) and he 700 million people suffer from iodine deficiency. the most important The focus of this work is nutritional support for a wide variety of people. You may have common illnesses or no illnesses at all. recommendation system Recommendation system ("system" may be replaced by synonyms such as: platform or engine) is a subclass of information filtering system intended to do this. Predict the "rating" or "liking" that users give to an item. they are Mainly used in commercial applications.



The recommendation process has basically three stages that are Information Collection Phase, Learning Phase and Recommendation Phase. The information is firstly collected about a particular problem and the various solutions related to that problem are categorized. After the collection of information Learning Phase comes

in which various conclusions are made out of that information which is gathered and in last phase i.e. Recommendation Phase an output is given in which various recommendations are made. In our system since it is a diet recommendation system so the recommendations will be about the diet plan like what all things you should eat, what is your BMI (Body Mass Index) which states whether you are healthy, overweight, or under-weight.

Techniques used for building a Recommendation System-

1.1 Content based Filtering Method- The content-based method is a domain-dependent algorithm which focuses on much more on the evaluation of the characteristics of things to produce predictions. When files like pages, publications as well as news are being suggested, the content-based filtering strategy is probably the most profitable. In a content-based filtering technique, the suggestion is made based upon the person profiles with features obtained from the information in the things the person has examined in previous times.

1.2 Collaborative based Filtering Method- Collaborative filtering is a domain-independent prediction technique for content that cannot easily and adequately be described by metadata such as movies and music. Collaborative filtering technique works by building a database (user-item matrix) of preferences for items by users. In the newer, narrower sense, collaborative filtering is a method of making automatic predictions (filtering) about the interests of a user by collecting preferences or taste information from many users (collaborating). The underlying assumption of the collaborative filtering approach is that if a person A has the same opinion as a person B on an issue, A is more likely to have B's opinion on a different issue than that of a randomly chosen person.

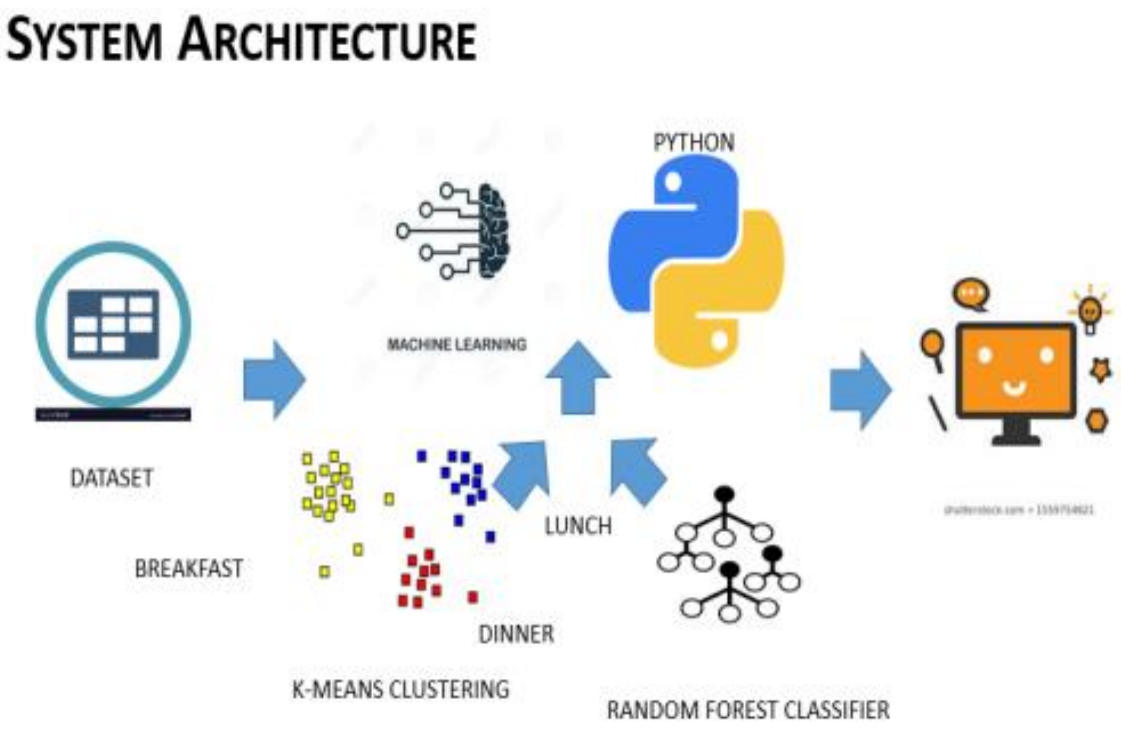
a. Memory based Filtering Method- The items that have been previously rated by the user before play a pertinent part in looking for a neighbor that shares appreciation with him. When a neighbor of a person is found, various algorithms could be utilized combining the tastes of friends to produce recommendations. Because of the usefulness of these strategies, they've accomplished extensive results in real-life applications.

b. Model based Filtering Method- In this approach, models are developed using different data mining, machine learning algorithms to predict users' rating of unrated items. There are many model-based CF algorithms. Bayesian networks, clustering models, latent semantic models such as singular value decomposition, probabilistic latent semantic analysis, multiple multiplicative factor, latent Dirichlet allocation and Markov decision process based models.

1.3Hybrid based Filtering Method- A number of applications combine the memory-based and the model-based CF algorithms. These overcome the

limitations of native CF approaches and improve prediction performance. Importantly, they overcome the CF problems such as sparsity and loss of information. However, they have increased complexity and are expensive to implement. Usually most commercial recommender systems are hybrid, for example, the Google news recommender system.

2. SYSTEM ARCHITECTURE:



1. User's will enter the necessary information like their age, gender, weight etc. on the website.

2. The information will then go through the ML model in following manner:

2.1 K-Means is used for clustering to cluster the food according to calories

2.2 Random Forest Classifier is used to classify the food items and predict the food items based on input

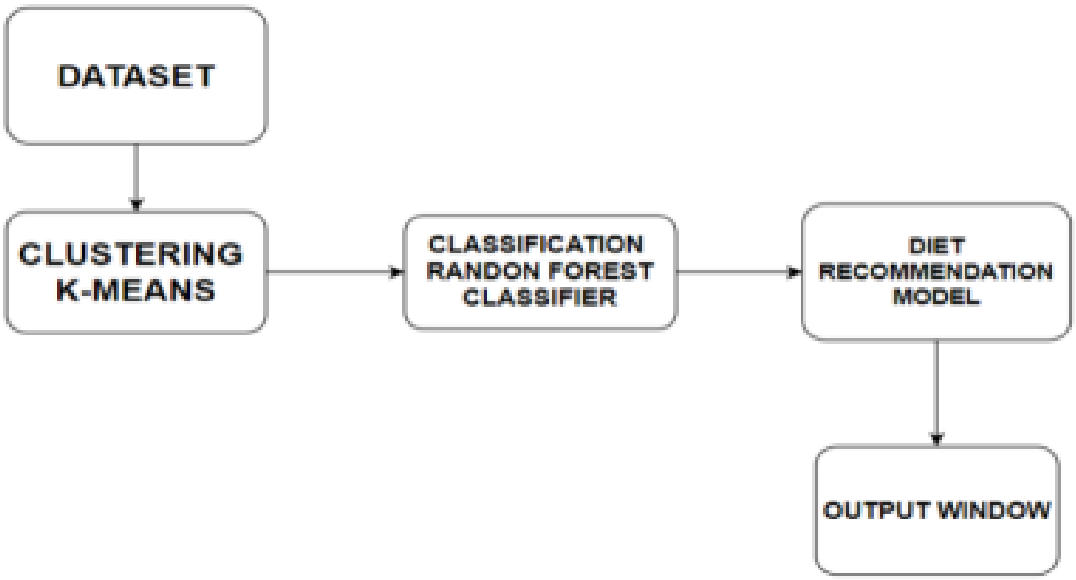
3. After analyzing all the data the system will respond by showing user's BMI and their current state (Overweight, Underweight, Healthy)

4. The System will then recommend diet to the users into three categories (breakfast, lunch, dinner) based on input

5. The Users can choose from multiple recommended items and make their diet plan.

6. After selecting food items the system will calculate selected food calories and show user's comparison between how much calories they chosen against how much they need to consume daily. 7. Accordingly then the User's will make its diet plan

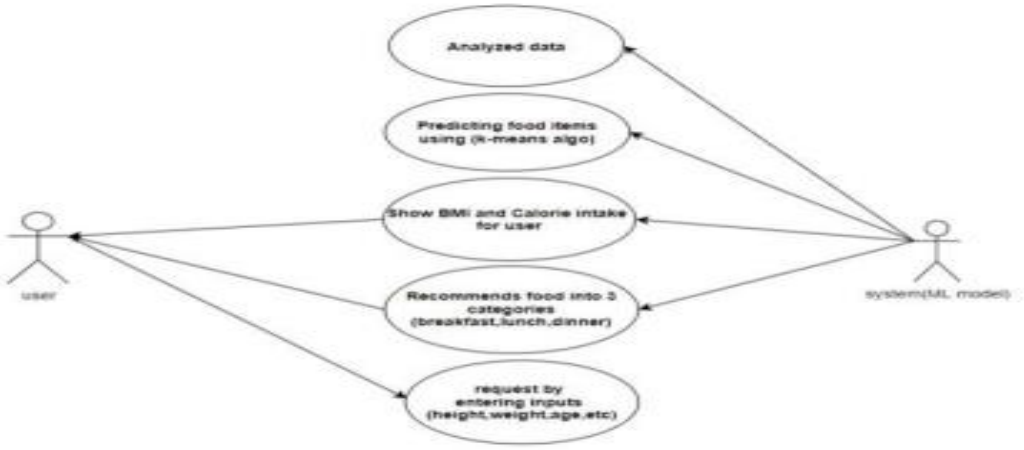
3. SYSTEM WORKFLOW:



3.1 . IMPLEMENATION AND DESIGN

USER FLOW-

User's will request to system by providing their physical information and after analyzing the data as a response the system (ML model) will recommend a diet which include (breakfast, lunch, dinner) based on the user information accordingly.



4. PROPOSED WORK:

The system operates in a machine learning environment where it analyses user data to determine the best diet to follow. We separated the dataset into three groups:

1. Lunch Data

2. Breakfast Data

3. Dinner Data

As a result, we train the ML model with various inputs to produce the intended user outputs. Here, we mainly employed 2 algorithms, which are:

1. K-Means

2. Random Forest

The model will produce a diet plan for the user based on the user's preference for a healthy diet, weight gain, or weight loss, as well as the data and category they have chosen

4.1.a K-MEANS ALGORITHIM-

Kmeans algorithm is an iterative algorithm that tries to partition the dataset into pre-defined distinct nonoverlapping subgroups (clusters) where each data point belongs to only one group. It tries to make the intra-cluster data points as similar as possible while also keeping the clusters as different (far) as possible. It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid (arithmetic mean of all the data points that belong

to that cluster) is at the minimum. The less variation we have within clusters, the more homogeneous (similar) the data points are within the same cluster.

The way k-means algorithm works is as follows:

1.Specify number of clusters K.

2.Initialize centroids by first shuffling the dataset and then randomly selecting K data points for the centroids without replacement.

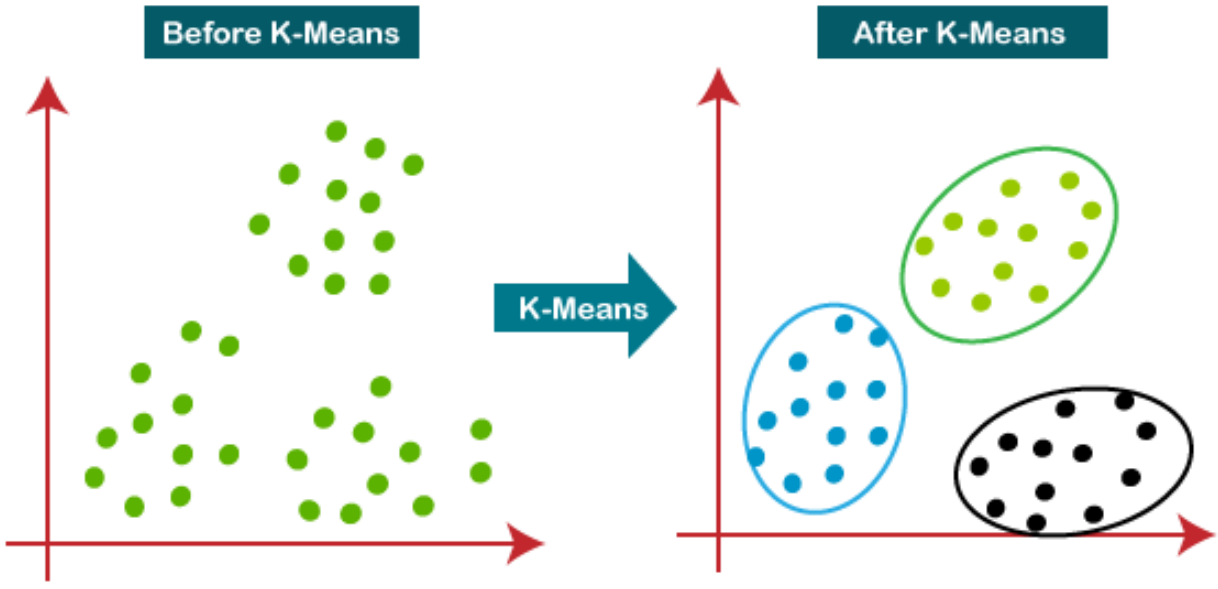
3.Keep iterating until there is no change to the centroids. i.e assignment of data points to clusters isn't changing.

4.Compute the sum of the squared distance between data points and all centroids.

5.Assign each data point to the closest cluster (centroid).

6.Compute the centroids for the clusters by taking the average of the all data points that belong to each cluster.

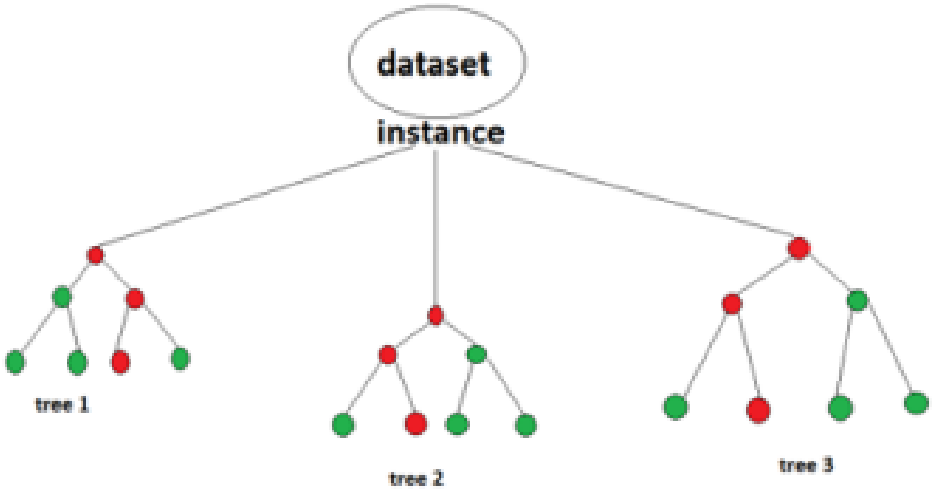
In our project the data set is divided into three categories lunch, breakfast, dinner with the help of k means clustering algorithm the below diagram shows how all three categories are separated from the cluster a dataset This helps us to finally divide the dataset into train and test dataset for all three categories and further the model is built in using the random forest algorithm.



4.1.b RANDOM FOREST ALGORITHIM-

Random Forest algorithm is a supervised classification algorithm. We can see it from its name, which is to create a forest by some way and make it random. There is a direct relationship between the number of trees in the forest and the results it can get: the larger the number of trees, the more accurate the result. But one thing to note is that creating the forest is not the same as constructing the decision with information gain or gain index approach. The decision tree is a decision support tool. It uses a tree-like graph to show the possible consequences. If you input a training dataset with targets and features into the decision tree, it will formulate some set of rules. These rules can be used to perform predictions.

When we have our dataset categorized into 3 category so now Random forest helps to make classes from the dataset. Random forest is clusters of decision trees all together, if you input a training dataset with features and labels into a decision tree, it will formulate some set of rules, which will be used to make the predictions.



4.1 Header Files Used:

```
import pandas as pd
import numpy as np
from tkinter import *
from sklearn.cluster import KMeans
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
```

Pandas was used to read datasets.

Numpy was used to convert features into numpy and then perform the further operations.

Tkinter was used to create interface.

KMeans was used to perform clustering.

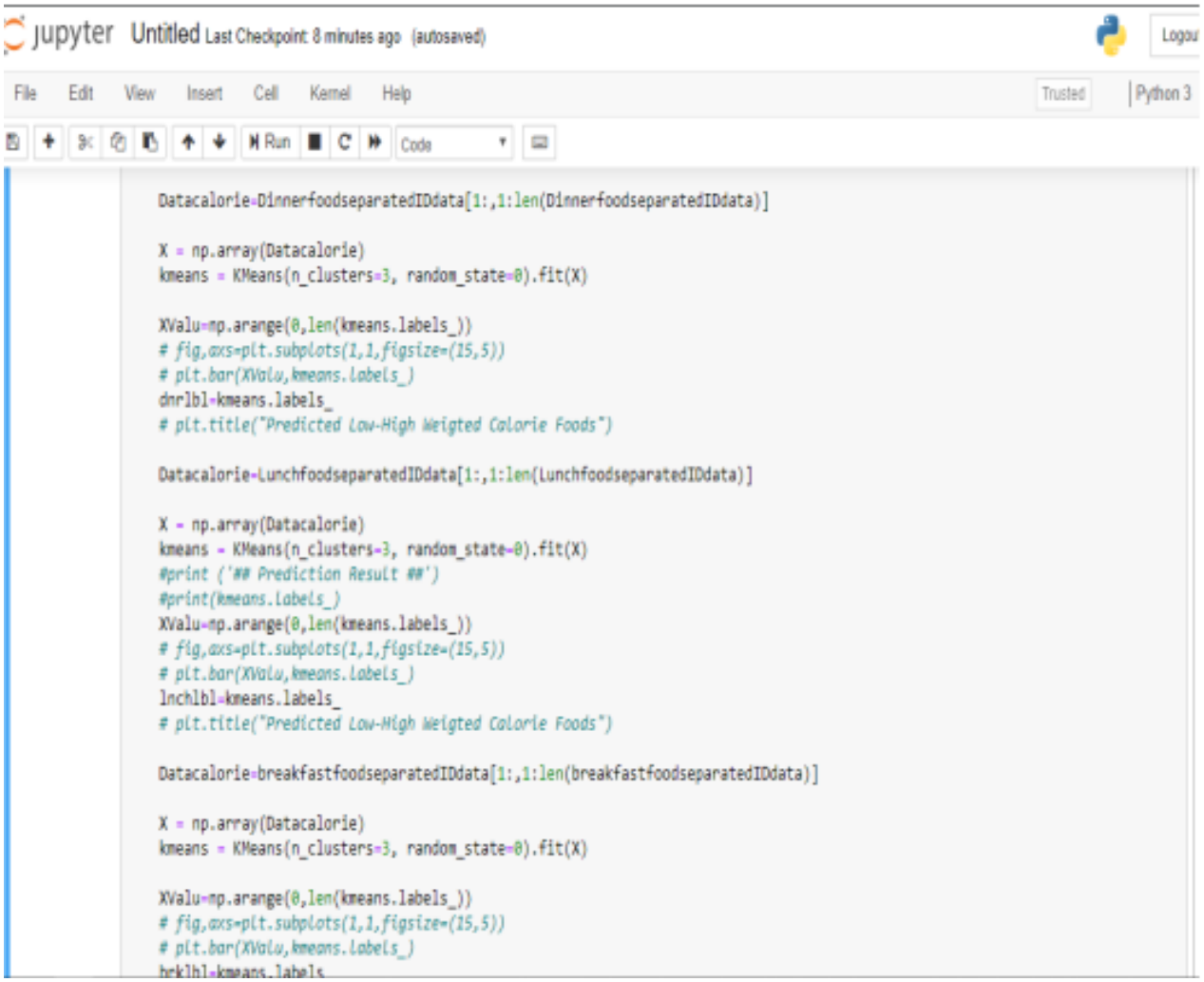
Train_test_split was used to divide the dataset into train and test portions to train and test the model.

RandomForestClassifier used to predict the food items based on clustered data.

4.2 Reading the Dataset:

```
data=pd.read_csv('food.csv')
Breakfastdata=data['Breakfast']
BreakfastdataNumpy=Breakfastdata.to_numpy()

Lunchdata=data['Lunch']
LunchdataNumpy=Lunchdata.to_numpy()

Dinnerdata=data['Dinner']
DinnerdataNumpy=Dinnerdata.to_numpy()
Food_itemsdata=data['Food_items']
```

4.3 Applying KMeans for clustering on Lunch_data, Breakfast_data, Dinner data:



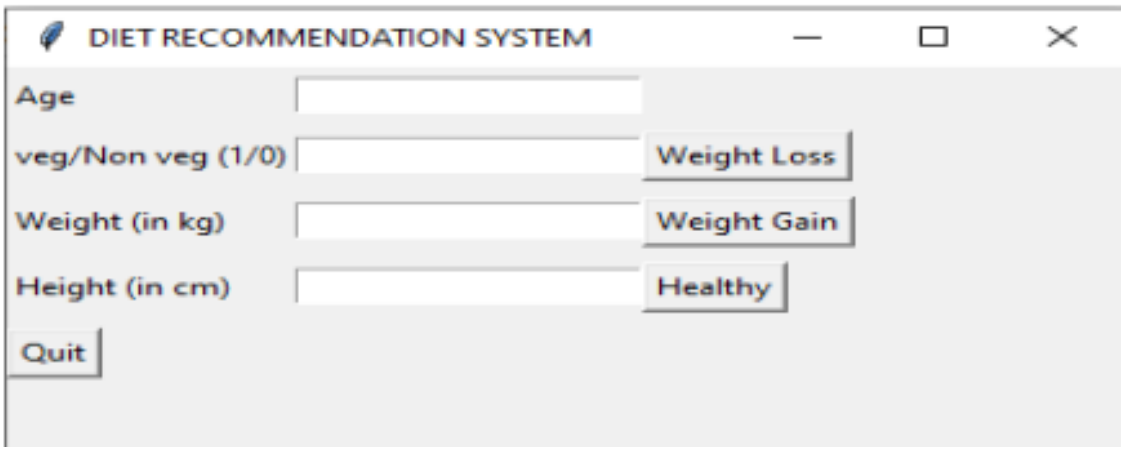
4.4 Applying Random Forest Classifier:

```
#Create a Gaussian Classifier
clf=RandomForestClassifier(n_estimators=100)

#Train the model using the training sets y_pred=clf.predict(X_test)
clf.fit(X_train,y_train)


X_test2=X_test
y_pred=clf.predict(X_test)
```

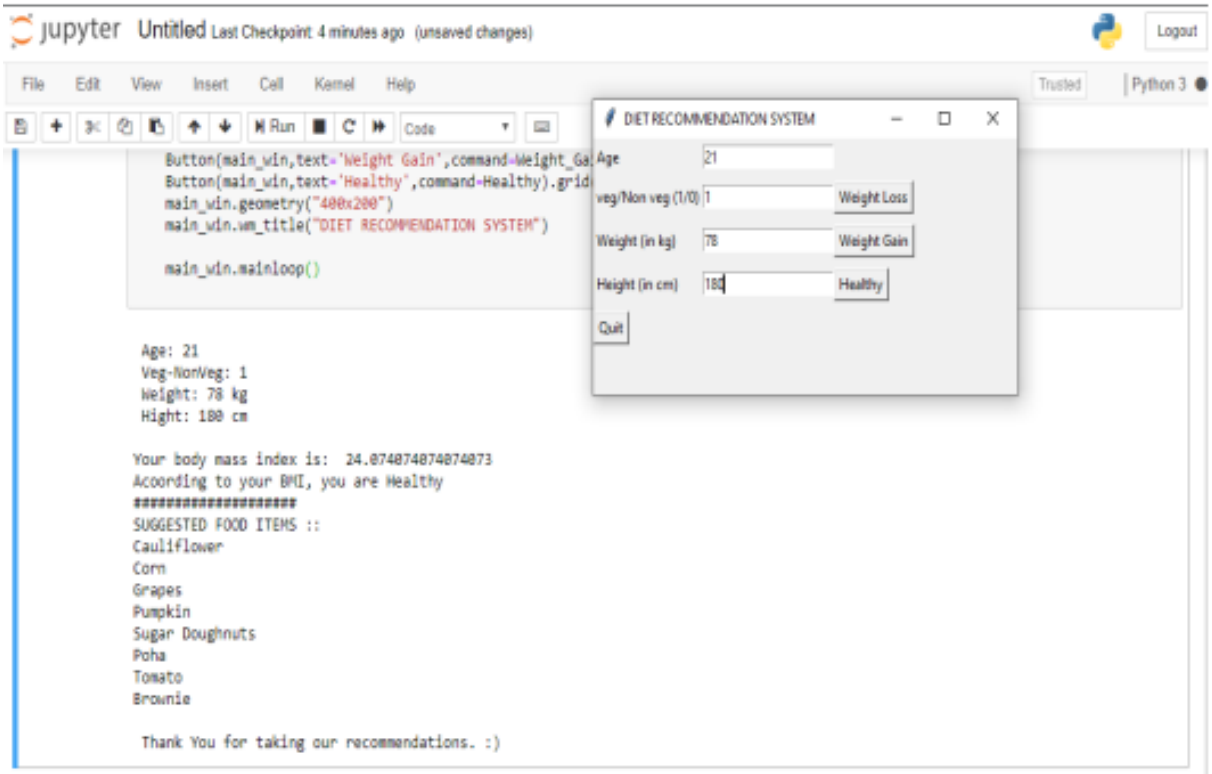
4.5 Creating Interface:

```
if __name__ == '__main__':
    main_win = Tk()

    Label(main_win,text="Age").grid(row=0,column=0,sticky=W,pady=4)
    Label(main_win,text="veg/Non veg (1/0)").grid(row=1,column=0,sticky=W,pady=4)
    Label(main_win,text="Weight (in kg)").grid(row=2,column=0,sticky=W,pady=4)
    Label(main_win,text="Height (in cm)").grid(row=3,column=0,sticky=W,pady=4)

    e1 = Entry(main_win)
    e2 = Entry(main_win)
    e3 = Entry(main_win)
    e4 = Entry(main_win)

    e1.grid(row=0, column=1)
    e2.grid(row=1, column=1)
    e3.grid(row=2, column=1)
    e4.grid(row=3, column=1)

    Button(main_win,text='Quit',command=main_win.quit).grid(row=5,column=0,sticky=W,pady=4)
    Button(main_win,text='Weight Loss',command=Weight_Loss).grid(row=1,column=4,sticky=W,pady=4)
    Button(main_win,text='Weight Gain',command=Weight_Gain).grid(row=2,column=4,sticky=W,pady=4)
    Button(main_win,text='Healthy',command=Healthy).grid(row=3,column=4,sticky=W,pady=4)
    main_win.geometry("400x200")
    main_win.wm_title("DIET RECOMMENDATION SYSTEM")

    main_win.mainloop()
```

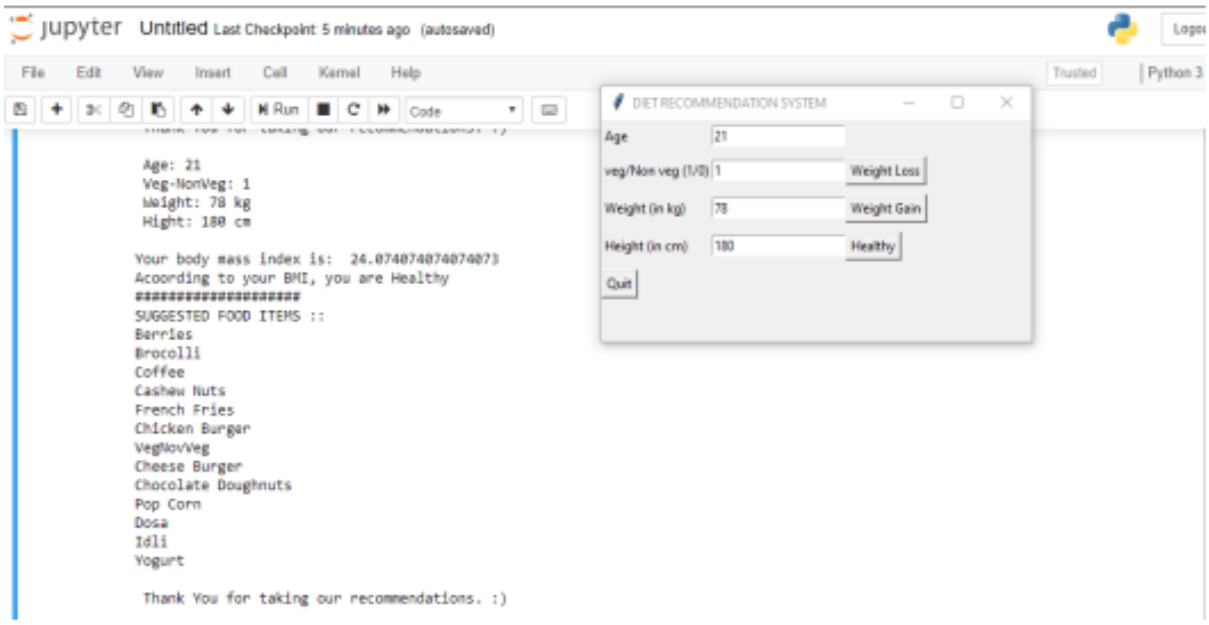
4.6 Taking Input:



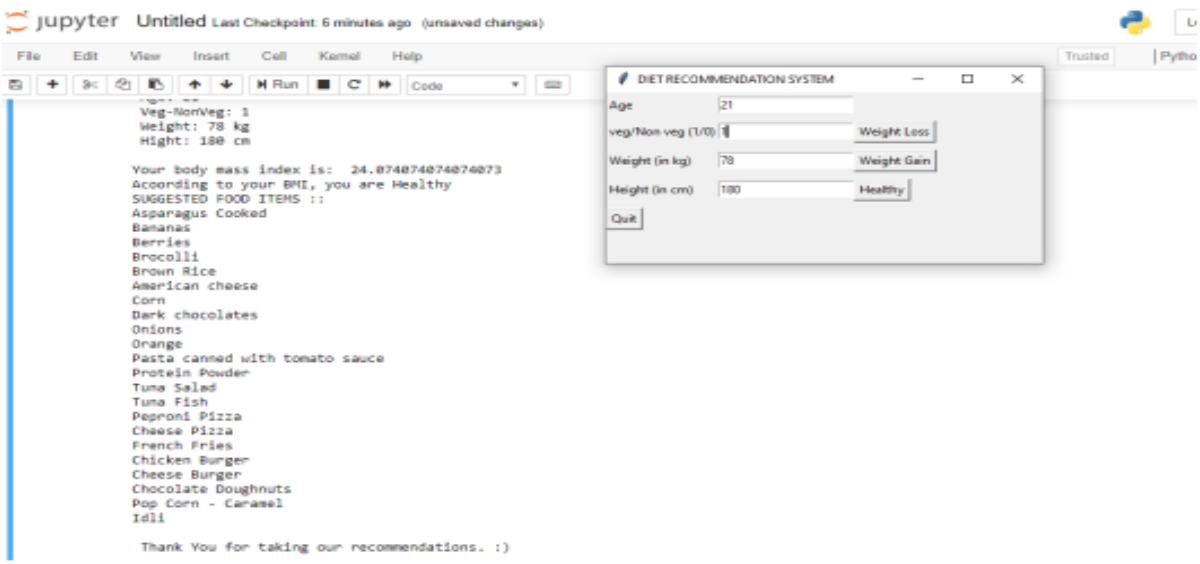
4.7 Predicting food items for Weight Loss Diet Plan:



4.8 Predicting Food Items for Weight Gain Diet Plan:



4.9 Predicting Food Items for Healthy food items:



5. IMPLEMENTATION PROCEDURE:

For training of the system, the initial process involves the segregation of food items depending upon the meal for which they are consumed i.e Breakfast, Lunch and Dinner. The clustering of various nutrients depending upon which are essential for the weight_loss, weight_gain and healthy is performed. After the clustering is performed, using Random Forest classifier, the nearest food items are predicted which best suited for the appropriate diet. As part of user interface, the inputs needed from the user are Age, Height, Weight and what the purpose for which the diet is required. Depending upon it, from the appropriate clustering, specific food items are classified and recommended to the user.

6.RESULT:

As soon as the project is launched, a dialogue box asks the user for information such as their age, preference of being vegetarian or not, weight in kilogrammes, height in centimetres, and aim, which can be either weight loss, weight gain, or simply being healthy.Our system analyses their BMI (Body Mass Index) and suggests a list of food items depending on the necessary amount of calories once the details are supplied and the option Weight Gain is selected.Alternatively, the option for weight loss is also available. It's the same with the Healthy option, which offers a list of foods to eat to keep healthy.

7.CONCLUSION:

Emerging technologies like artificial intelligence and machine learning are crucial to the growth of the IT (Information Technology) industries. We have utilised these technologies to develop a software for those seeking advice regarding their diets and to live a healthy life. To live a healthy and fit life, nutritional guidance is

becoming more and more important. A healthy diet plan is developed by the system by accepting the user’s preference and profile.