

## IMPROVED PREDICTIVE LEARNING APPROACHES FOR CUSTOMIZED DIET SUGGESTION FRAMEWORK IN MEDICINAL SERVICES

V.Nallarasan,

Department of Information Technology,  
SRM Institute of Science and Technology, Chengalpattu Dt., India

L.Anand

Department of Information Technology,  
SRM Institute of Science and Technology, Chengalpattu Dt., India

J.Prabakaran

Department of Information Technology,  
SRM Institute of Science and Technology, Chengalpattu Dt., India

**ABSTRACT-** In this modern world various people suffer from different types of diseases and illnesses. It is generally very difficult to suggest a diet as quickly as possible. Most people have a dire need to Lose Weight, Gain Weight or stay Healthy. Time has also become a possible constraint. The project makes use of a dataset which contains various nutrients in the correct amount. In the wake of the situation, we have tried to develop a program that recommends diet to the people. The items recommended are limited to three categories: Weight Loss, Weight Gain and Healthy category. Our project uses Machine Learning Algorithms named K-Means Clustering for clustering the data and Random Forest Classification to classify according to the categories listed. To predict the food items the Diet Recommendation System uses user inputs from a Graphical User Interface including age, height, weight, vegetarian or non-vegetarian food and selecting the above three categories. The working prototype of the Diet Recommendation System lists a set of food items as per the user inputs. The module uses the weight and height to calculate the body Mass Index (BMI) of the user and based on the preference of the kind of diet he wants the Recommendation System predicts the list of food items. The various kinds of nutrients and their calories are taken into consideration along with other relevant details like Fats, Proteins, Iron, Calcium, Sodium, Potassium, Carbohydrates, Fiber, Vitamin D and Sugar. A diverse range of food items are considered like major Vegetarian and Non- Vegetarian foods. Carbohydrates, Fats and Nutrients are the major contributors that are considered to make a food.

**Key words:** Machine Learning (ML), k-Means Clustering, Random Forest, Body Mass Index.

### I. INTRODUCTION

#### 1.1 OVERVIEW

What should I eat ? When should I eat ? How often should I eat ? How much should I eat ? These are some of the questions which arise in the minds of the people. A healthy and balanced diet is a key element in everyone's life. A balanced diet is one which contains sufficient amounts of all the nutrients namely carbohydrates , fats , proteins , vitamins , minerals , sugar etc which are necessary for keeping a person in perfect health.

'Dosha' , a Sanskrit word which means the root of a body helps us to infer that if the Dosha of our body i.e the root causes is disturbed then it may result in a number of diseases. Thus, maintaining a healthy diet is essential for everyone.

A diet recommendation system using machine learning algorithms has been proposed in this project which will suggest food items to the user from a predefined dataset depending upon user preferences for healthy food or food for weight loss or food for weight gain.

Fig 1 : Balanced Diet Nutrient Proportion



## 1.2 EVOLUTION OF HUMAN DIET

Through social development and changes in living space and biology, there have been various significant dietary evolution in human advancement, including meat eating, cooking, and those related with plant and animal hunting.

Right back from the Stone Age when it is said people were more prone to eating meat by hunting, with slow evolution to the Modern World where people have a vast variety of options for food. Diet evolves on several parameters like nutrition, availability, disease, technology. Examining the evolution of diet might help in finding the optimal solution for diet recommendation. The changing diet has also led to morphological changes in the human body. In this technological era, the diet is being suggested by machine learning models that are being developed after analyzing large sets of food data. With numerous options available for food, the diet of a person can be efficiently recommended based upon his requirements.

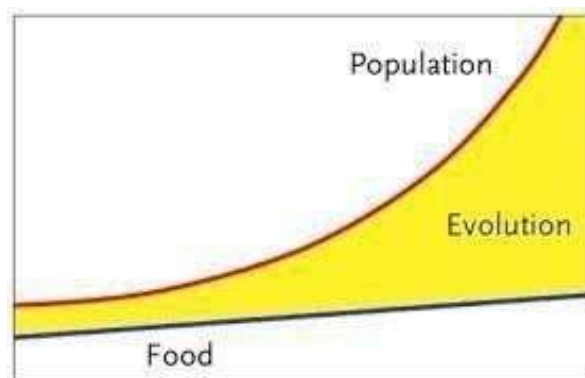


Fig 2: Evolution of food with population

## 1.3 BACKGROUND AND MOTIVATION

Considering the busy life of people in the present world, it results in situations when people tend to skip proper meals or have food at irregular intervals. In such a scenario, it becomes essential for someone to provide them with a proper diet plan which needs to be strictly adhered to. For the same, people have personal dieticians which provide them with a proper food chart regularly to keep them fit. However, it is not possible for humans to handle many cases and analyze all the elements which are needed to be considered for providing a healthy and balanced diet. Keeping this in mind, an idea struck among the researchers to develop certain systems which may implement machine learning algorithms to gather the user data such as age, weight, height, diseases etc. and analyze them to provide a menu of food items to the user for the proper meals at the right time. Studying such systems and considering the importance of it, we carried on with diet recommendation.

### 1.1 TYPES OF DIET PLAN

The need of diet has increased sharply such that it resulted in creation of various types of diet plans namely Keto Diet, Paleo Diet, Vegan Diet , Low Carbs Diet , Low Fat Diet etc. Each diet plans have certain benefits which can be used by people for weight gain or Weight loss or to remain fit.

### 1.1 PROBLEMSTATEMENT

Making decisions about what to eat is a major problem in our everyday lives due to a wide variety of ingredients, culinary styles, ethnicities, cultures, and personal tastes. Choosing the right dish at the right time seems to be a very difficult task.

Today, many diseases that were previously thought as hereditary are now shown to be seen connected to biological dis-function related to nutrition. Although being healthy and eating better is something the vast majority of the population want, doing so usually requires great effort and organization.

Thus, this project proposes an integration of machine learning algorithms to recommend the right food at the right time and with the right nutrition, calories, fat etc. here, we are using three attributes of food nutrition details, a person's physical details and time.

### 1.4 RESEARCHOBJECTIVES

Various machine learning algorithms which have been implemented in various ways to recommend diet to the user and which algorithms are more accurate and precise than the other. It briefed us about the idea of various elements such as age, gender, weight, height.

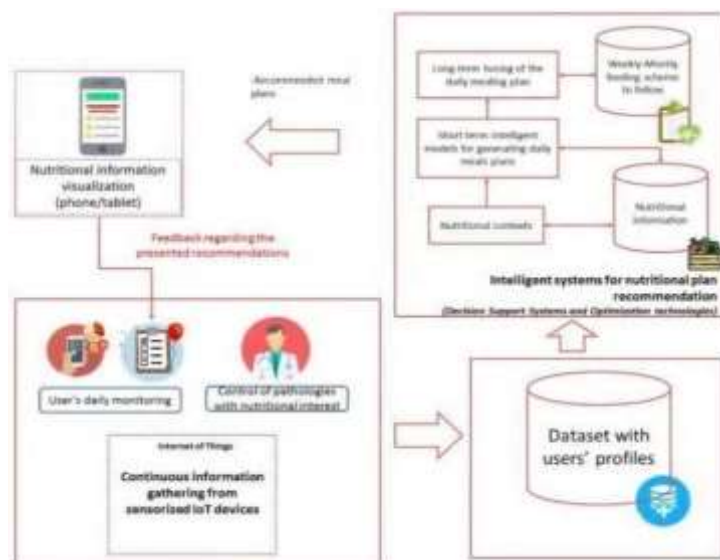


Fig 3 : A general Food Recommendation System

## II. LITERATURE STUDY

### 2.1 DIET RECOMMENDATION TO PATIENTS USING AHPMETHOD-2019:

Author(s): Sakshi Singh, Sanjay Kumar Dubey

The problem of malnutrition is of great concern in today's world. It can lead to various diseases which can be cured by a balanced diet in its initial stage. Marasmus - a disease which occurs due to malnutrition and which methods can be used to provide a diet plan has been discussed in this paper [1].

Marasmus is a disease which can occur due to lack of carbohydrates, fats, proteins, lipids, glycogen and other essential nutrients. When these nutrients are in deficient amount in our body, the stored nutrients are consumed to provide energy to the body and slowly the body starts consuming muscles and tissues energy as well. The end results of this include underweight, loss of fat mass and muscle

mass, dehydration and some other symptoms. It was found that by giving a decent eating routine containing all the essential supplements, for example, proteins, starches and so on, marasmus can be dealt with. It utilized AHP which helps in deciding the decent eating regimen which has all the essential supplements in the right extents and finally to approve the outcomes it utilized the Fuzzy Logic to make comparisons in the result.

The D. Rao et al. referenced in their article about lack of supplements. The lack of healthy sustenance results in predominated development and delicate insusceptible framework [2]. The article portrays an exhaustive audit of an examination quickly referencing the different strategies for the developing nations in evacuating the lack of healthy sustenance. A. Thapar and M. Goyal approached with a fuzzy method to analyze whether a newborn child is experiencing malnutrition [3].

The AHP method demonstrated made use of determining the main factors and the diet choices as prescribed by the dietician. After the analysis, a pairwise comparison matrix is developed to determine the Consistency Index and Ratio. Calculating these values, gave an idea about the ranking for various diets as which are more definite. The analysis result is confirmed by performing the fuzzy logic.

It made use of two techniques and gave an idea about how diet prediction can be provided but a more detailed and more efficient methodologies can be implemented to provide a more balanced diet.

## 2.2 DIET RECOMMENDATION USING FUZZY APPROACH-2019:

Author(s): Arushi Singh, Nandini Kashyap, Rakesh Garg

This paper describes how the human population lacks focus on diet and ends up having a dynamic eating routine with no appropriate food items and nutrients. To facilitate with a proper diet, a fuzzy methodology is provided [4].

In today's busy schedule, people tend to skip proper meals at proper times. It is important to ensure that the nutrients are perceived by the body in the right proportion and regularly, lack of which the person ends up with various illnesses starting with temperature, joint pains etc and ending up with serious diseases. Thus, it is necessary for a person to have a balanced diet. People tend to consult dieticians for providing a diet chart to them. However, it is considered that dieticians may tend to forget some parameters for consideration. Hence, a fuzzy logic method was implemented which considers a number of human parameters essential for diet plan depending upon the age group, health issues and other factors.

Fuzzy logic proposed makes use of input, output and a function providing values between 0 and 1. This method was used because the diet recommendation data is highly uncertain and vast. The inputs used were age group and nutrition which was fed into the fuzzy ontology and the output provided the percentage of the nutrition which should be received by the person per day.

As indicated by the test results, the suggested diet is more precise than the eating routine taken from the dietician. The drawbacks found for providing a diet included lack of food items list which can be also suggested to the user instead of just the nutrition proportion. Providing a list of food items for different meals can prove to be more efficient for a healthy diet.

## 2.3 PERSONALIZED EXPERT RECOMMENDATION SYSTEM FOR OPTIMIZED NUTRITION-2018:

Author(s) : Chih-Han Chen, Maria Karvela, Mohammadreza Sohbat, Thaksin Shinawatra, Christofer Toumazou

Studies have found strong correlation between genetic information of a person to the person's personalized diet. It is believed that all particles associated with human digestion are controlled straightforwardly or by implication of genes and in this manner the soundness of people can be enhanced through customized dietary guidance. To achieve what kind of system can be developed is discussed in this paper [5].

Nowadays, people conduct tests to gather detailed information about their genes so that they can plan their diet accordingly. Consequently, a system which would take into account/correlate the

grocery product information and genotypic information of a person was proposed. However, the fundamental obstruction of associating hereditary data to a diet is the multifaceted nature of information and the versatility of the applied frameworks.

The focus was to perform the operations of categorization and analysis of data and eventually with a recommendation of list of products. The system architecture was so developed that it would accept new product information including the nutritional values. The block was then trained with the new values and receives updates from the state machine block and stores the trained data. During training, food products are classified into different categories. The block sets up a threshold for nutrition using genetic algorithms. Ultimately, the decision system would recommend a suggestion of products based on phenotype of personal data, the categories of products and nutritional values.

## 2.4 IoMT ASSISTED PATIENT-DIET RECOMMENDATIONS SYSTEM-2020:

Author(s): Celestine Iwendi, Suleman Khan, Joseph Henry Anajemba, Ali Kashif Bashir

The need to increase the efficiency of recommender systems to predict the food items to patients resulted in implementation of machine and deep learning algorithms which is discussed in this paper [6].

Healthy diet prescribed by dieticians or artificial system automated diet on cloud servers can help to improve the life of a person; free from stress and diseases. However, the efficiency of such a system can be improved by training and implementing machine and deep learning algorithms namely Naive Bayes, Logistic Regression, Recurrent Neural Network, Multilayer Perceptron, Long Short-Term Memory.

The dataset was collected of some patients with details about their diseases, age, gender, nutrient levels, weight and height. On applying the machine and deep learning algorithms for it, it was found that LSTM provided more accuracy than other algorithms. From the dataset the essential features were analyzed using random forest classifier.

Specialists have found that a healthy diet is a good alternative to cure patients with diseases [7-8]. Researchers found that people are prone more to medicines rather than a healthy diet, the nutritional intake is not sufficient for them and to solve this issue a cloud based system was proposed. In this paper key points involve the usage of machine and deep learning algorithms. Provide an idea about how the model works under the items and patient ailment specifications. Study the various mechanisms for deeper understanding of patients' diets.

The entire process was carried out in five phases: Dataset Collection, Analysis of Optimal Features, Training and Testing, Evaluation. For the purpose, dataset was divided into training (70%) and testing (30%) sets. As an end result, it was concluded that LSTM has higher accuracy among all the other methods.

## 2.5 Food Recommendation System Considering Nutrition and User Preferences - 2019

Author(s): Racielyera Toledo, Ahmad A. Alzahrani, Luis Martinez

Health issues have always been a concern for the majority of the population. Researchers are implementing various methods for providing an efficient personalized diet suggestion. This paper deals with basic recommendations along with laying emphasis on nutrition and user preferences as well [9]. Research challenges involve inquiring about difficulties with the assortment of client data, the social occasion of healthful data from nourishments and plans, and the changing of eating practices.

For food recommendation, various factors such as personal attributes, lifestyle, needs and preferences, budget etc. are involved. Nutrition can vary from people to people, from healthy people to patients, vulnerable groups, allergic people and so on. Personal information can enhance the accuracy of menu generated frameworks. The focus was on two important points: building information model which can be obtained through questionnaires, flowcharts and processing nutritional information.

The methodology initiates with preparation of food profiles which have about 600 food items and about 20 nutrients. A template was also prepared for classifying the food depending upon breakfast, lunch and dinner. This acted as a decision table wherein the foods suggested are provided as an alternative depending upon the nutrients.

AHP Sort based pre filtering stage was used to classify important and unimportant food items. AHP Sort behaves as a multi-criteria decision analysis which helps in making decisions depending upon various alternatives. Furthermore, an optimization based approach for menu suggestion was implemented so that it focuses on user preferences also. It is dependent upon probabilistic research which recommends food on the basis as which food item is frequently consumed by people. A case study was carried on to discuss the efficiency of the system and improvements required.

This paper lacks using food history of consumers which can also be used for improving the recommendation, involving recipes of food items.

## 2.6 Personalized Nutrition Solution based on Nutrigenomics -2019:

Author(s): Jitao Yang

This paper brings in the concept of Nutri genomics and how it is related to the diet of a person[10].

Nutrigenomics refers to the impact of the diet consumption on the DNA of the person. Genome (a set of DNAs) consists of protein genes. These genes are responsible for maintaining the health of a person. Any infected gene can have an impact on the health of the person. By analyzing the variation in genes, various algorithms can be implemented to provide the nutrition details to a person. The DNA sequence will be studied by generic algorithms. Also considering lifestyle and other attributes, the application will provide a nutrition report.

A mobile application is developed which will take in the genetic information of the person and depending upon these it will suggest a menu to a person. The food items suggested can be analyzed with fine details of weight or nutrient content and other factors. Also recipes for various foods are provided which can be utilized for an efficient diet.

The paper just brought in the relation between a person's diet and his/her genes but lacked the concept of implementing any algorithms which can be used for the same. Also, the algorithm used in the application developed could have been explained for a better understanding. It however did not consider various other factors such as age, height, weight etc. of a person which are equally important for a diet.

## 2.7 Type-2 Fuzzy Logic for Diabetic-Patients Diet Recommendation -2018

Author(s): HebaAbdelgader Mohammed, Hani Hagra

This paper discusses the Fuzzy Logic method for recommending diet to diabetic patients [11].

The number of diabetic patients are increasing at a faster rate in the world. It is becoming difficult for the doctors to take care of each diabetic patient and prevent them from attaining complications. Eventually, a system was proposed so that an efficient diet can be recommended to such people so that it becomes easy for the doctors to handle them.

## 2.8 DIETOS: Diet Organiser System -2016

Author(s): Agapito G., Calabrese B., Guzzi P. H., Cannataro M., Simeoni M., Car'e I., Lamprinoudi T., Fuiano G., Pujia A.

This paper deals with a web-based application called DIETOS which provides the prototype and the flow in which diet can be recommended to people [12].

The application makes usage of health profiles of the user by collection the information through random questionnaires. Also the questionnaire is so prepared that it is adaptive depending upon the user's answer to different questions. The health profile of the user is prepared. The database involves Calbrain foods and the system can also provide information about which foods can be allergic or harmful to the user depending upon his health status.

It consists of the following modules: User Profile, Reminder, History, Security, Foods Filter. The User Profile module is concerned with the collection of the health information of the person along with other attributes. Reminder helps the person updated with the food menu and that he/she needs to regularly update the profile to increase efficiency. History keeps the records of all the updates made. Security is necessary because the details related to health is confidential to the person and hence the details collected are encrypted. Foods Filter basically filters out the food and then provides suggestions to the user depending upon the profile created by him.

The paper lacks the list of food items other than Calabrian foods. It aims to provide functions for adding food information by specialists themselves. Provide recipes of certain dishes. It overall aims to enhance the interface and functionalities so that more people use it and the work of doctors and dieticians become easy.

### III. PROPOSED METHODOLOGY

#### 3.1 FUNCTIONAL REQUIREMENTS

##### 3.1.1 GRAPHICAL USER INTERFACE

The main graphical user interface will consist of various input fields taking into consideration the user preferences for food like vegetarian/ non-vegetarian food items at Lunch/ Dinner/ Breakfast along with other details like an individual's height, weight and age. The Diet Recommendation System will recommend the list of desired foods based on the fact that whether the user is expecting a weight gain/ weight loss/ healthy diet.

##### 3.1.2 WEB APPLICATION

The Diet Recommendation System will be a web application. The application will have an interface as mentioned above. The interface will take the user inputs into account to make recommendation of the food items to the user based on the current requirements as stated by the user.

##### 3.1.3 TREND ANALYSIS

The Diet Recommendation System will analyze the food items being predicted under each category like Weight Gain/ Weight Loss/ Healthy and will maintain a record of the healthy food items under the trending category of healthy food items.

##### 3.1.4 USER ACCOUNT AND PREFERENCE

Diet Recommendation System will allow users to create an account and specify their interests. The user will have the option of choosing their interests once at the time of account creation and further change it anytime they desire. This makes the platform more versatile and allows it to cater to a wider range of audience.

##### 3.1.5 INTEREST PREDICTION

The final feature which Diet Recommendation System will implement will be tracking the user's interests and then classify them into groups (Weight Gain/ Weight Loss/ Healthy as per the Breakfast/ Lunch/ Dinner timings) and then display more food items related to the preferred group of the user.

### 3.2 NON-FUNCTIONAL REQUIREMENTS

Non-functional requirements include performance requirements, security requirements and portability requirements which is equivalently important for the proper functioning of the system.

#### 3.2.1 PERFORMANCE REQUIREMENTS

To improve the performance of the software it needs to be executed with a high internet speed so that no issues are observed in the existing working system. Various other measures need to be taken care of like the storage that of course requires a bigger workspace. As a result there is no performance issues observed and the module created is lightweight and can work on any platform.

#### 3.2.2 SAFETY REQUIREMENTS

In the time of an unexpected damage situation to the application a backup system needs to be established which can take care of software till the instances are revived.

#### 3.2.3 SECURITY REQUIREMENTS

Each and every user's data should be kept secret and will be only accessible through his own credentials in the system which any other user will not be able to access. The software needs to be hosted on cloud server to store all the user data. The rights of the user should be restricted so that whatever the commands that he must be running on the terminal or the interactive script that he is running should not cause any damage to the existing system or the working prototype of the module. This gives a high sense of rating to the system in terms of security measures.

### 3.3 OBJECTIVE

Wide variety of ingredients, cultures and personal tastes makes decision about what to eat a great problem. Many diseases that were previously thought as hereditary are now seen to be connected to biological dysfunction related to nutrition. Being healthy and eating better is something the vast majority of the population wants and doing so usually requires great effort. The working prototype accomplishes a Personalized Diet Recommendation System with integration of Machine Learning Algorithms to recommend the right food at right time and with the right nutrition, calories, fat etc.

#### DATASET

The dataset used to implement this project is a collection of various food items along with the time during which the food needs to be consumed i.e. Dinner/ Breakfast/ Lunch. The dataset also contains information whether a specific food item is veg or non – veg. For every food item the information about the calorie intake, the amount of Fats, Proteins, Iron, Calcium, Sodium, Potassium, Carbohydrates, Fiber, Vitamin D and Sugar. The dataset contains 90 unique items with all the 16 above mentioned attributes which will be used to train the Machine Learning model and make predictions whether the food item or a list of food items is suitable for a user or not to maintain his healthy food habits. The value 0 or 1 indicates if the particular food item successfully satisfies the column criteria or not. The data set is a collection of vital food items in an individual's life, each with all the following attributes like Food items, Breakfast, Lunch, Dinner, Veg, Non-Veg, Calories, Fats, Proteins, Iron, Calcium, Sodium, Potassium, Carbohydrates, Fiber, Vitamin D and Sugars. The dataset is self-prepared and the values that are taken for each attribute is verified from various official food calorie websites which maintain the correct amount of the measure for each nutrient for a particular unit of measure. In this Diet Recommendation System the values that are taken are considered in units per 100 g. A common standard is maintained for each and every nutrient so that the food value which is being taken into consideration is normalized and when the application makes use of the dataset it considers it in general amount and then makes recommendations for the user.



Fig 4: Dataset snippet -1

Food Items	Breakfast	Lunch	Dinner	Veg/Non-Veg	Calories	Fats (g)	Proteins(g)	Iron(mg)	Calcium(mg)	Sodium(mg)	Potassium(mg)	Carbohydrates (g)	Fibre (g)	Vitamins D (mcg)	Sugars (g)
1. Asparagus Cooked	0	1	1	0	22	0.2	2.4	0.91	23	14	224	4.1	2	0	5.3
2. Avocado	1	0	0	0	140	15	2	0.55	12	7	485	8.5	6.7	0	0.7
3. Banana	1	0	0	0	89	0.3	1.1	0.26	5	1	358	23	2.6	0	12
4. Bagels made in wheat	0	1	1	0	250	1.5	10	2.78	30	438	165	48	4.1	0	6.1
5. Berries	1	0	0	0	349	0.4	14	6.8	190	298	77	77	13	0	46
6. Broccoli	0	1	1	0	25	0.5	3.8	1.37	118	98	343	3.1	2.8	0	0.6
7. Brown Rice	0	1	1	0	362	2.7	7.5	1.8	33	4	266	70	3.4	0	0
8. Cauliflower	0	1	1	0	32	0.3	3	0.72	32	258	278	6.3	3.3	0	0
9. American cheese	1	0	0	0	131	24	20	0.84	497	988	363	6.3	0	0	0
10. Coffee	1	0	0	0	2	0	0.3	0.02	2	1	50	0.2	0	0	0
11. Corn	1	1	1	0	97	1.4	3.3	0.95	7	253	3.3	72	2.7	0	7.7
12. Dark chocolates	0	0	1	0	536	22	5.5	2.13	30	6	508	60	8.5	0	48
13. Grapes	1	0	0	0	93	2.1	5.6	2.83	283	8	272	17	11	0	6.3
14. Milk	1	0	1	0	97	4.9	3.8	0.12	989	52	178	5.2	0	0	0
15. Cantina Tasty	1	0	0	0	553	44	18	6.88	37	12	660	30	3.3	0	5.9
16. Onions	0	1	1	0	40	0.1	1.1	0.21	23	4	146	9.3	1.7	0	4.2
17. Orange	1	0	0	0	97	0.2	1.5	0.8	381	3	212	25	11	0	0
18. Pasta-corned with tomato	0	1	1	0	71	0.7	3.2	0.91	13	381	182	14	0.9	0	4
19. Peas	1	0	0	0	57	0.1	0.4	8.18	9	1	116	15	3.1	0	9.8
20. Pear	0	1	1	0	85	0.4	5.4	1.47	25	5	244	14	5.7	0	5.7
21. Protein Powder	1	0	0	0	411	17	46	8.57	500	329	1129	19	7.1	288	5.7
22. Pumpkin	0	1	1	0	18	0.5	0.7	0.97	15	237	230	4.3	1.1	0	2.1
23. Tuna Salad	0	1	1	1	187	9.3	18	1	17	402	178	9.4	0	0	0
24. Tuna Fish	0	0	1	1	184	6.3	30	1.31	10	50	323	0	0	0	0
25. Pepperoni Pizza	0	0	1	0	298	14	12	2.14	145	892	188	30	1.8	0	3.2
26. Cheese Pizza	0	0	1	0	276	11	11	2.47	183	580	170	23	2.1	0	2.5
27. French Fries	0	1	1	0	289	14	3.5	0.91	17	357	545	37	3.9	0	0.3
28. Chicken Burger	0	1	1	1	292	15	18	0.83	13	859	315	20	1.3	0	0

Fig 5: Dataset snippet-2

29. Chicken Sandwich	0	1	1	1	257	3.2	15	1.32	32	605	256	23	1.2	0	5
30. Sugar Doughnuts	0	1	1	0	424	29	5.2	1.06	60	402	102	51	1.5	0	32
31. Chocolate Doughnuts	0	1	1	0	452	25	4.9	4	24	326	201	51	1.9	0	27
32. Pop Corn - Caramel	1	0	1	0	385	1.4	2	0.8	18	288	110	90	2.5	0	65
33. Pop Corn	1	0	1	0	429	9.5	13	2.28	11	400	241	73	14	0	0.5
34. Donut	1	0	1	0	168	3.7	4.5	8	0.7	94	78	29	0.9	0	0.1
35. M&	1	1	1	0	166	1.7	5	17.2	4	207	63	30.2	2.1	0	0.74
36. Poha	1	0	0	0	130	1.5	2.8	3.18	1	391	117	28.9	1.1	0	0.5
37. Chopped	0	1	1	0	291	7.5	11	3.01	85	486	266	48	4.9	0	2.7
38. Tomato	1	1	1	0	16	0.2	1.2	0.47	5	42	215	3.2	0.9	0	2.63
39. Yogurt	1	1	1	0	90	4	3.1	0.08	180	70	234	7	0	1	7
40. Broccoli	1	0	0	0	407	8.2	4.4	3.81	17	457	51	94	2.9	0	55
41. Noodles	0	1	1	0	108	0.2	1.8	0.14	4	18	4	24	1	0	0
42. Uthappi	1	1	1	0	188	7.2	4.4	24	6.4	522	91	26.4	2.2	0	0.24
43. Bhaj Puri	1	0	0	0	101	2.4	9	37.4	3.4	438	180	29.3	1	0	1.35
44. Dal Makhani	0	1	1	0	109	8.5	2.1	35.2	8.3	243	366	6.3	1.5	0	3.29
45. Almonds	1	0	0	0	579	56	21	3.71	269	1	733	22	13	0	4.4
46. Muesli	1	1	1	0	22	0.3	3.1	0.5	3	5	318	3.3	1	7	2
47. Egg Yolk cooked	1	0	0	1	596	15	14	1.89	42	287	152	8.8	0	88	0.4
48. Sweet Potatoes cooked	1	1	0	0	76	0.1	1.4	0.72	27	230	19	2.5	0	5.7	0
49. Baked Potatoes	1	1	1	0	92	0.1	1.9	0.31	5	248	379	25	2	0	0.9
50. White Rice	0	1	1	0	360	0.6	0.9	4.36	9	1	86	79	1.4	0	0
51. Orange juice	1	0	0	0	45	0.2	0.7	0.2	11	1	200	10	0	0	8.4
52. Greek yogurt plain	1	1	1	0	73	1.9	10	0.04	115	34	141	3.9	0	0	3.6
53. Oat Bran Cooked	1	0	0	0	40	0.9	3.2	0.88	10	1	92	11	2.6	0	0
54. Green Tea	1	0	0	0	1	0	0.2	0.02	0	1	0.2	0	0	0	0
55. Chia seeds	1	0	0	0	488	31	17	7.72	631	16	407	42	34	0	0
56. Cottage cheese with vegi	0	1	1	0	95	4.2	11	0.1	36	483	85	3	0.1	0	0.4
57. Salmon	0	1	1	1	127	4.4	21	0.38	7	75	360	0	0	435	0
58. Corns Corn Flakes	1	0	0	0	384	0.9	5.9	19.4	2	571	107	68	2.7	286	7.8
59. Beans	0	1	1	0	31	0.2	1.8	1.05	37	6	211	7	2.7	0	3.3
60. Lentils	0	1	1	0	101	0.5	8.8	3.1	14	246	384	21	0	0	0
61. Pasta with corn tomatoes	0	1	1	0	128	0.7	2.8	0.25	1	0	31	28	4.8	0	0
62. Taw	1	0	0	0	1	0	0	0.08	2	1	4	0.3	0	0	0

### 3.4 ARCHITECTURAL DIAGRAM

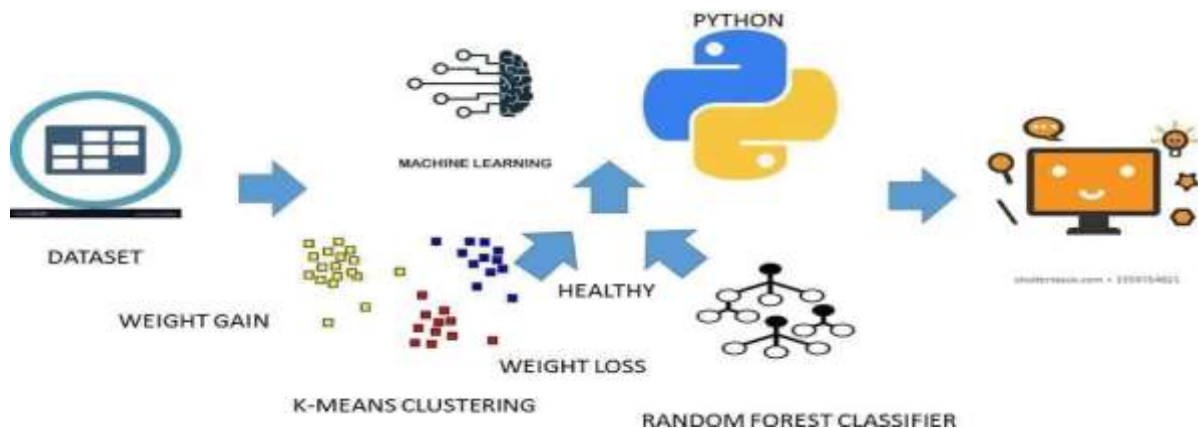


Fig 7: Architectural Diagram

The above figure shows the architectural diagram for the working prototype of the Diet Recommendation System. The code is written in Python which is used to implement various Clustering and Classification algorithms in Machine Learning for predicting a proper diet to the user. 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 suits for the appropriate diet. As part of user interface, the inputs needed from the user are Age, Height, Weight and for what the purpose the diet is required. Depending upon it, from the appropriate clustering, specific food items are classified and recommended to the user. The K – Means Clustering Algorithm takes into consideration various attributes and forms a cluster. The new attribute values being taken into account fall into one of the clusters and then the classification algorithms are used to predict the food item list based on the type of inputs given by the user. The input is taken from the Tkinter GUI which is designed using Python. The main graphical user interface will consist of various input fields taking into consideration the user preferences for food like vegetarian/ non- vegetarian food items at Lunch/ Dinner/ Breakfast along with other details like an individual's height, weight and age. The Diet Recommendation System will recommend the list of desired foods based on the fact that whether the user is expecting a weight gain/ weight loss/ healthy diet.

The GUI for the Diet Recommendation System is depicted below with all the input fields as stated above:



Fig 8.GUI Interface

## IV. IMPLEMENTATION

### 4.1 MODULES

The implementation of the Diet Recommendation System is accomplished using various modules namely,

- *Dataset*
- *Machine Learning Algorithms*

- *K – MeansClustering*

K-means algorithm is an iterative algorithm that tries to partition the dataset into  $K$  pre-defined distinct non-overlapping 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.

- *Random Forest Classification*

Random forest is a supervised learning algorithm which is used for both classification as well as regression. But however, it is mainly used for

classification problems. As we know that a forest is made up of trees and more trees means more robust forest. Similarly, random forest algorithm creates decision trees on data samples and then gets the prediction from each of them and finally selects the best solution by means of voting. It is an ensemble method which is better than a single decision tree because it reduces the over-fitting by averaging the result.

- *GUI (using Python's TkinterLibrary)*

#### 4.2 DATA FLOWDIAGRAM

The data flow in the below working prototype is as such, firstly, the dataset is taken into consideration which is utilized by the K – Means Clustering Algorithm to classify the entire data into clusters of Weight Loss, Weight Gain and Healthy and then training a Random Forest Classifier to make food items list suggestion to the user. After reading the dataset, the entire dataset is classified into Breakfast, Dinner and Lunch. The `to_numpy` function is used to convert the values to the numpy array which can be considered for further manipulation. All the details stored in each of the rows are retrieved in the specified category using the `iloc` method. A Python's list is created which contains the food items and all the nutrients and then the Transpose of the value stored is stored in another variable. As stated above the process is repeated for the Lunch, Dinner and Breakfast itself. The Diet Recommendation System calculates the Body Mass Index taking into consideration the Height and Weight of the user. An age range is also displayed to state the comparison of the BMI for a specific range group. Based on the BMI calculated the category is specified that the human is underweight, healthy, overweight or severely overweight. Now the application of K – Means Clustering Algorithm comes into play. The number of clusters decided for the module is three namely Weight Gain, Weight Loss and Healthy. All the food items at the time of Breakfast, Lunch and Dinner falls into any of these categories. We make use of the K- Means Clustering Algorithm after importing the functionalities present in this class which can be achieved using the statement from `sklearn`. `Cluster import K-Means`. The data is fitted to the model and then make predictions for the food items. Now we take the data of a specific cluster and divide it into the test and the train dataset which is further used to train the Random Forest Classification model to make recommendations of the food items to the user.

#### 4.3 TOOLS



The tools being utilized are Python 3 for making use of the Machine Learning Algorithms and the execution of the Diet Recommendation takes place using a Spyder IDE which makes use of the Tkinter library.

#### 4.4 PROCEDURE

##### 4.4.1 GRAPHICAL USERINTERFACE

The implementation of the Diet Recommendation System begins with the user input as accepted from the Graphical User Interface (GUI). The GUI is build using Python's library Tkinter. An object of this library class is created and used to build a GUI. To create a Tkinter based GUI we need to

import all its function using the statement from tkinter import \*. As soon as we get all the functions of the tkinter library, we can use it to create and design our GUI. An object main\_win is created of the Tkinter class. The GUI contains four labels and text entry boxes each for Age, Veg/Non-Veg, Weight and Height and align it to a specific position on the screen. Four Buttons stating the Weight Loss, Weight Gain, Healthy and Quit options are made available on the GUI so that the required functions can be called and then the food can be recommended as per the user's inputs.

Fig 10:GUI



#### 4.4.2 DATASET

After reading the dataset, the entire dataset is classified into Breakfast, Dinner and Lunch. The to\_numpy function is used to convert the values to the numpy array which can be considered for further manipulation. All the details stored in each of the rows are retrieved in the specified category using the i\_loc method. A Python's list is created which contains the food items and all the nutrients and then the Transpose of the value stored is stored in another variable. As stated above the process is repeated for the Lunch, Dinner and Breakfast itself. The Diet Recommendation System calculates the Body Mass Index taking into consideration the Height and Weight of the user. An age range is also displayed to state the comparison of the BMI for a specific range group. Based on the BMI calculated the category is specified that the human is underweight, healthy, overweight or severely over weight.

Fig 11: Categorisation of food items in various meals

```
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' ]
breakfastfoodseparated=[]
Lunchfoodseparated=[]
Dinnerfoodseparated=[]

breakfastfoodseparatedID=[]
LunchfoodseparatedID=[]
DinnerfoodseparatedID=[]

for i in range(len(Breakfastdata)):
    if BreakfastdataNumpy[i]==1:
        breakfastfoodseparated.append(Food_itemsdata[i])
        breakfastfoodseparatedID.append(i)
    if LunchdataNumpy[i]==1:
        Lunchfoodseparated.append(Food_itemsdata[i])
        LunchfoodseparatedID.append(i)
    if DinnerdataNumpy[i]==1:
        Dinnerfoodseparated.append(Food_itemsdata[i])
        DinnerfoodseparatedID.append(i)
```

```

#conditions
print("Your body mass index is: ", bmi)
if ( bmi < 16):
    print("severely underweight")
    clbmi=4
elif ( bmi >= 16 and bmi < 18.5):
    print("underweight")
    clbmi=3
elif ( bmi >= 18.5 and bmi < 25):
    print("Healthy")
    clbmi=2
elif ( bmi >= 25 and bmi < 30):
    print("overweight")
    clbmi=1
elif ( bmi >=30):
    print("severely overweight")
    clbmi=0

```

Fig 12 : BMI index values

#### 4.4.2 MACHINELEARNING

In the Diet Recommendation System there are three main functions Weight Loss, Weight Gain and Healthy that will use the Machine Learning functionalities like K- Means Clustering and Random Forest Classification Algorithm to recommend the list of food to the user. The user defined functions makes use of various in build Python's function used in Machine Learning to provide the actual functionality. We are using various libraries like Pandas, Numpy and Sklearn in general. The Pandas library is used to import the data set so that the machine learning model can make use of it and the list of food items can be predicted to the user. The Numpy is used for values conversion to array so that it can provide the values which can be used to train the Machine Learning model. The Sklearn package has many of the Machine Learning Algorithms. The one which we are using are K- Means Clustering and Random Forest Classification Algorithms.

##### 4.4.2.1 K-MEANS CLUSTERINGALGORITHM

###### A. INTRODUCTION

K-means clustering is a type of unsupervised algorithm, which is employed when you have unlabelled data (i.e., information without defined types or clusters). The objective of this algorithm is to find groups in the data, with the number of groups represented by the variable K. The algorithm works iteratively to allot apiece data point to one of K groups constructed on the features that are delivered. Data points are bundled based on feature resemblance. The outcomes of the K-means clustering algorithm are:

1. The centroids of the K clusters, which can be used to label noveldata
2. Labels for the preparation data (each data point is assigned to a single cluster)

Rather than describing groups before observing at the data, clustering permits you to find and analyse the groups that have made naturally.

Each centroid of a cluster is a assortment of feature values which define the resulting groups. Investigating the centroid feature weights can be used to qualitatively construe what kind of group each cluster signifies.

###### A. BUSINESS USES OF K-MEANS

The K-means clustering algorithm is used to search groups of data which have not been obviously



labelled in the data. This can be used to authorize business expectations about what types of groups occur or to recognize unidentified groups in complex data sets. Once the algorithm has been run and the groups are definite, any new data can be easily assigned to the correct group.

- Behavioral segmentation:
- Section by purchase history
- Segment by activities on application, website, or display place
- Define facades based on interests
- Create profiles based on activity monitoring
- Inventory categorization
- Group images
- Detect activity types in motion sensors

#### **B. STEPS FOR K-MEANS CLUSTERING**

1. Step one: Initialize cluster centers
2. Assign observations to the closest cluster center
3. Revise cluster centers as mean of assigned observations
4. Repeating above 2 and 3 steps until getting accurate clusters

#### **C. IMPLEMENTATION**

Now the application of K – Means Clustering Algorithm comes into play. The number of clusters decided for the module is three namely Weight Gain, Weight Loss and Healthy. All the food items at the time of Breakfast, Lunch and Dinner falls into any of these categories.

We make use of the K- Means Clustering Algorithm after importing the functionalities present in this class which can be achieved using the statement from sklearn.Cluster import K-Means. The data is fitted to the model and then make predictions for the food items. Now we take the data of a specific cluster and divide it into the test and the train dataset which is further used to train the Random Forest Classification model to make recommendations of the food items to the user. The same process is repeated for Weight Loss, Weight Gain and Healthy Categories as per the user requirements.

K- Means Clustering Algorithm allows to cluster the data and is a very convenient tool for discovering the groups in your dataset that might not have been predicted earlier. To identify the clusters just by observing the scatter plots as such is not an easy task. To help facilitate it we can use the K- Means Algorithm in Machine Learning. K stands for the number of clusters. In our case the number of clusters is 3. The way this algorithm works is by just selecting the centroids of each clusters, it can be any points in the dataset or any random points in the scatterplot. This step is followed by the assignment of several points to the nearest centroid. To make a cluster one of the approaches can be to divide the entire scatter plot into divisions such that if a line is drawn perpendicular to the line joining the two centroids (in case of 2 clusters) then the entire part is divided into two halves and each point on the line is equidistant from the two centroids of the clusters. If a point is above the line, it is near to the first centroid and thus is considered as a part of that cluster. On the other hand if it is in the lower half then it is near to the second centroid and thus forms the part of the second cluster. The Algorithm works on the simple principle of mathematics that states the concept of geometry. This centroid is the temporary centroid and is only for the initial stage of classification of the scattered points. Now comes into role the new and actual centroid that has to be placed in the formation of the cluster. This is calculated based on the centre of mass of the points that are falling into the specific cluster. This average value helps in finding the actual centroid. Now after the actual centroid has been found then all the scattered data points needs to be reassigned based on the new changes. If any reassignment takes place then the above step of finding the centroid again

has to be repeated and if this is not the case i.e. all the data points remain at the same point exactly then the centroid of the cluster is already decided and the process of finding the clusters is finalized. This can be referred as the intuition behind the K –Means Clustering. Another effective phenomenon is the Random Initialization Trap to form the clusters in the K– Means Algorithm. This approach considers a different approach to form the clusters. Select at random K points, the centroids (not necessarily from the existing dataset). Then the concept of equidistant lines comes into role and we divide the scatter plots into three parts based on the rule of simple geometry. Resulting into the formation of the final n clusters. Further the centroid is reallocated and the cluster is given a new form. Again the points are reassigned after the formation of clusters with new centroid. If the scatter points are at the same place with respect to the new centroid then no need to reform the entire cluster structure.

The K - Means++ Algorithm is a solution to the common problem when a situation arises about what is a True Cluster and what is False Cluster. The K-Means++ Algorithm solves the conditions above and results in the formation of true clusters.

In K – Means Clustering choosing the right number of clusters in an important decision. In a particular scenario whether 2 or n number of clusters is important can be decided as per the performance of the situation. The formula for deciding the number of clusters is given by:

$$WCSS = \text{Sum of all the points in Cluster 1 distance } (P_i, C_1)^2 + \text{Sum of all the points in Cluster 2 distance } (P_i, C_2)^2 + \text{Sum of all the points in Cluster 3 distance } (P_i, C_3)^2.$$

If we consider the initial situation of one cluster then we have only one centroid in a single cluster. On the other hand if we consider two clusters each having its own centroid then the formula for calculating the WCSS is more optimal.

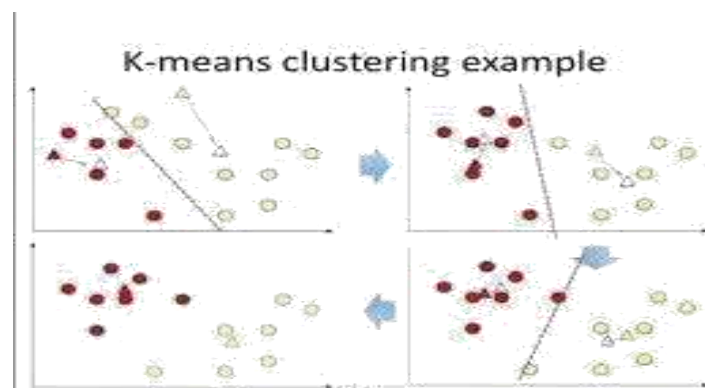


Fig 13: K means Clustering

#### 4.4.2.1 RANDOM FOREST CLASSIFICATION ALGORITHM

##### A. INTRODUCTION

Random Forest is a supervised machine learning algorithm which is widely used for classification problems. The working of this algorithm is based on construction of multiple decision trees at the time of training time and the output produced is the class with highest frequency or mean of the individual trees. Each decision tree constitutes a root node and each internal node of the tree represents an attribute and each leaf node defines class label.

##### B. WORKING OF RANDOMFOREST

The working consists of two steps:

##### 1. Creation of RF pseudocode:

- I. Randomly selecting n features from total m features where  $n \ll m$
- II. Calculation of node “i” using the best split from n features
- III. Splitting the node into child nodes
- IV. Repeating the above steps until k number of nodes has been reached
- V. Building the forest by repeating step 1 to step 4 for n trees

- .RF prediction pseudocode:
- Using the test features and rules of every created decision tree for predicting the result and storing the outcome or target
- Calculation of votes for every predicted target
- The final prediction is the most voted predicted target



## V.CONCLUSION

A Diet Recommendation System is implemented with the working functionalities like:

- 5.6.1 Desired food list prediction.
- 5.6.2 Diet list based on weight category.
- 5.6.3 Body Mass Index (BMI) Calculation.

Health is vital for an individual and can be achieved with this working module. Thus making life healthy. Wide variety of ingredients, cultures and personal tastes makes decisions about what to eat a great problem. Many diseases that were previously thought as hereditary are now seen to be connected to biological dysfunction related to nutrition.

Being healthy and eating better is something the vast majority of the population wants and doing so usually requires great effort. The working prototype accomplishes a Personalized Diet Recommendation System with integration of Machine Learning Algorithms to recommend the right food at the right time and with the right nutrition

## VI.FUTURE ENHANCEMENT

- Enhance the UI Layer – The UI has to be given various better functionalities and must be made a device compatible version which when released in production can provide a better functionalities to its users. The various technologies that can be used are Bootstrap and further enhanced tech stack for Web Design like HTML 5, CSS 3 and Flask. Validation into the user's login and Sign-up for data privacy can be the next majorstep.
- Diversifying the data being used in the working prototype – A lot more enhancement in the dataset can be a major step in future so that a lot more details about the various food items can be taken into consideration to provide a better functionality to the user. More Nutrients can be included in the data being utilized along with other inputs in addition to what is used now from the GUI concerning the user details. A Data Layer needs to be used in the working module to keep a backup of all the information and user details with the food database.



- Make it support on Cloud-Based Technology – A cloud based technology version of the current supported module can help in better use of the packed module and dependencies being used. Cloud version of the module being released can provide better security, reliability, mobility and unlimited storage capacity.

## REFERENCES

- [1] SakshiSingh , Sanjay Kumar Dubey, "Recommendation of Diet to a Patient using AHP and Fuzzy Approach", International Conference on Cloud Computing, Data Science and Engineering ,IEEE-2019.
- [2] D.Rao,C.Higgins,H.Margot,T.Lyle,S.Falls,E.Obeysekare,K.Mehta,"Micro nutrient Deficiencies in the Developing World:An evaluation of Delivery Methods",Global Humanitarian Technology Conference, 2016.
- [3] A. Thapar and M. Goyal, "A fuzzy expert system for diagnosis of malnutrition in children," 2016 IEEE Region10 Humanitarian Technology Conference (R10-HTC), Agra, India, 2016, pp. 1-6. doi: 10.1109/R10HTC.2016.7906819.
- [4] ArushiSingh ,NandiniKashyap , RakeshGarg, "Fuzzy based approach for diet prediction" , International Conference on Cloud Computing, Data Science and Engineering , IEEE-2019.
- [5] Chih-Han Chen, Maria Karvela, MohammadrezaSohbati, ThaksinShinawatra, ChristoferToumazou, "PERSON—Personalized Expert Recommendation System for Optimized Nutrition" ,IEEE-2018.
- [6] Celestine Iwendi, Suleman Khan, Joseph Henry Anajemba, Ali Kashif Bashir, "Realizing an Efficient IoMT-Assisted Patient Diet Recommendation System Through Machine Learning Model" , IEEE-2020.
- [7] S. Saini and S. K. Dubey, "Recommendation of diet to jaundice patient on the basis of nutrients using AHP and fuzzy AHP technique," Int. J. Intell. Eng. Syst., vol. 10, no. 4, pp. 91–99, Jul. 2017.
- [8] L. Yang, C.-K. Hsieh, H. Yang, J. P. Pollak, N. Dell, S. Belongie, C. Cole, and D. Estrin, "Yum-Me: A personalized nutrient-based meal recommender system," ACM Trans. Inf. Syst., vol. 36, no. 1, pp. 1–31, Jul.2017.
- [9] Racielyera Toledo, Ahmad A. Alzahrani, Luis Martinez, "Food Recommender System Considering Nutritional Information and User Preferences" ,IEEE-2019.
- [10] JitaoYang , "Personalized Nutrition Solution based on Nutrigenomics" , 2019 19th International Conference on Computational Science and Its Applications , IEEE-2019.
- [11] V.Nallarasan,Sampada Talwar "Turbo Wheels – A Vehicle Based 3D Game"International Journal of Pure and Applied Mathematics ', 2018 .
- [12] HebaAbdelgader Mohammed, Hani Hagra, "Towards Developing Type 2 Fuzzy Logic Diet Recommendation System for Diabetes" ,IEEE-2018.
- [13] Agapito G., Calabrese B., Guzzi P. H., Cannataro M., Simeoni M., Car' e I., Lamprinoudi T., Fuiano G., Pujia A. , "DIETOS : a recommender system for adaptive diet monitoring and personalized food suggestion." ,IEEE-2016.
- [14] RomeshwarSookrah, JaysreeDeveeDhowtal, Soulakshmee Devi Nagowah, "A DASH Diet Recommendation System for Hypertensive Patients using machine learning" , 2019 7th International Conference on Information and Communication Technology ,IEEE-2019.
- [15] T. Karthick, M. Manikandan,"Fog assisted IoT based medical cyber system for cardiovascular diseases affected patients", Concurrency Computation Practice Experience, (2018).
- [16] Anandan. M, Manikandan. M, Karthick. T, "Advanced Indoor and Outdoor Navigation System for Blind People Using Raspberry-Pi "in Journal of Internet Technology, 2019.
- [17] T. Karthick ,V.Nallarasan,G.parimala,, "A Review on IoT based Healthcare Systems and Technologies," Journal of Xi'an University of Architecture & Technology Volume XII, Issue V, 2020 ISSN No : 1006-7930.