International Conference on Recent Trends in Data Science and its Applications
DOI: rp-9788770040723.216

Diet Recommendation System Using Machine Learning

Yuvraj Chibber
Data Science and Business Systems
SRM Institute of Science & Technology
Kattankulathur, Chennai, India
yc9714@srmist.edu.in

Dushyant Betala
Data Science and Business Systems
SRM Institute of Science & Technology
Kattankulathur, Chennai, India
db1624@srmist.edu.in

Srividhya.S*
Data Science and Business Systems
SRM Institute of Science & Technology
Kattankulathur, Chennai, India
srividhs1@srmist.edu.in

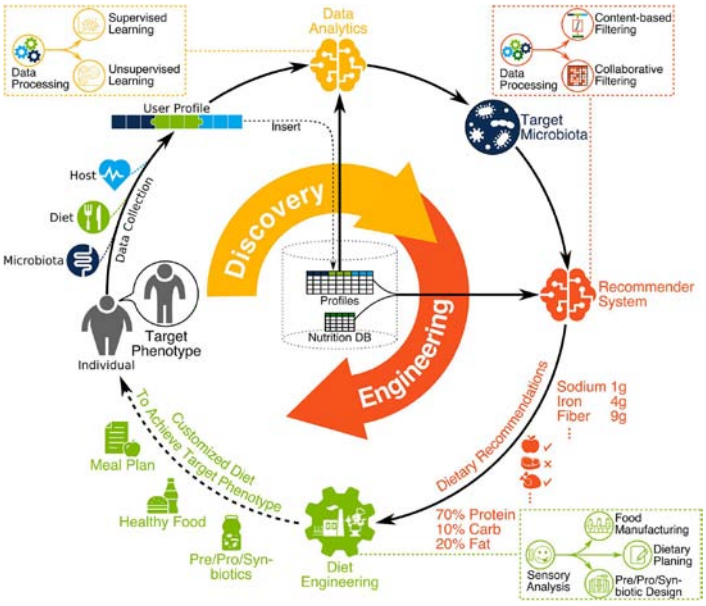
***Abstract* — A healthy diet is vital for maintaining overall health and well-being. It is essential to ensure that we consume a variety of nutrient-rich foods. The importance of a healthy diet is underestimated and we as humans often don't focus on the food we are consuming and if it is fulfilling the nutritional requirements of our body. Having a healthy diet can boost our immune system, It aids in maintaining a healthy weight, enhances mental and emotional well-being, and lowers the chance of developing chronic diseases. The quantity of energy provided by food is measured in terms of calories. Every day, we use calories to perform essential bodily functions such as breathing, moving around, walking, running, and more. An average person requires 2000 calories in their daily life but the precise amount depends on the physical attributes of a person like body weight, height, age, BMI etc. The food we consume everyday has a significant impact on our body. Consuming a balanced diet in conjunction with regular physical activity can promote overall health and well-being, leading to a better quality of life. Following a healthy diet is a difficult task firstly, unhealthy food options are often more readily available, cheaper, and more convenient than healthier alternatives. Additionally, many people have busy schedules that leave little time for meal planning and preparation, hence there is a need for a personalized diet recommendation which is based on a person's preferences. Diet recommendation system can save time and effort for individuals who may not have the knowledge or resources to create a healthy diet plan themselves. By providing personalized recommendations, the system can simplify the process of meal planning and make it easier to maintain a healthy diet in the long run. Machine Learning can be used to provide a solution to this problem, using Machine learning an alternative diet can be provided to the user which is familiar to their daily diet and fulfils the nutritional requirement of the body. The recommendation system provides a healthy diet for any person and devises a meal plan tailored to meet the requirements of the individual. It additionally suggests a diet program for people suffering from specific nutrition deficiencies such as Goiter, Osteoporosis, Anemia.**

Keywords— BMI, Machine Learning, Recommendation System, Calorie, Nutritional Deficiency.

I. INTRODUCTION

Health plays an important role in every human's life. What we end up consuming as humans impacts our health , hence it is essential to have a complete food diet which provides all the essential nutrients required by the body. Due to our hectic daily routine, it becomes difficult to follow a diet which provides all the nutrients required by our body. Not having a healthy diet can have a negative impact on our body, hence a solution to this problem is proposed by the diet recommendation system. The suggested system aims to assist the user in suggesting dietary plans for weight loss, weight gain or if they are suffering from a nutrition deficiency. In this paper we have modelled a diet recommendation system which provides a solution weight loss and weight gain in the form of a healthy diet plan. It also focuses on some specific diseases caused due to nutrition deficiency. The diseases covered by our system are Anemia, Osteoporosis, Goiter. Anemia is caused due to deficiency of iron, Osteoporosis due to deficiency of Calcium and Goiter due to deficiency of iodine. These illnesses are the most prevalent among people, necessitating regular health monitoring and treatment. The recommendation system is designed to offer information that aligns with the user's specific needs and limitations. Our system has been divided into two modules 1. Diet for Weight Loss / Weight Gain 2. Diet for specific nutrition deficiency.

The upcoming part of the paper is structured as follows: A literature review is present in Section 2. Section 3 presents the suggested methodology. Workflow of the system is described in Section 4. Section 5 provided the output and the result of the system. Section 6 provides with the conclusion of the project. Section 7 provides directions for future research.



DOI: rp-9788770040723.216

Fig.1. Working of a Diet Recommendation System

II. Literature Survey

In [1], a health care recommendation system was developed by utilizing an ontology framework for providing recommendations regarding exercise and food. The available dataset was used to extract user information using the decision tree technique. Maintaining the Integrity of the Specifications.

In [2], four different type of algorithms were used (namely k-nearest neighbors, Random forest, AdaBoost, Support vector machine SVM) to monitor the health of patients with LQTS genetic disorder. Random forest algorithm yielded better accuracy and score (60% to 70%) than the other three algorithms. The system identified patients who are at a high risk of cardiac complications by analyzing their ECGs.

[3] employed three different types of algorithms (random forest, logistic regression, gradient boosting) to monitor and verify the Fitbit credentials of a user. The logistic regression algorithm achieved higher accuracy and score (87%) compared to the other two algorithms.

In [4], the model uses the correlation between several resources to build the recommendation. Resources are referred to in content-based recommendation systems as a vector of attributes. Based on the features of the objects the user has assessed, the system then develops a profile of the user's interests. The algorithm looks at the connections between other products and the products that users have evaluated in order to forecast customer preferences. It accomplishes this by comparing the features or qualities of the various items. This helps the system determine which products are most likely to appeal to a given user based on their historical behavior and preferences.

The system created by Romeshwar Sookrah et al. [5] has a recommendation engine that gives hypertension patients tailored food regimens using content-based filtering and machine learning techniques. The recommendation engine takes into account various factors such as age, allergies, food preferences, alcohol consumption and smoking, blood pressure, and dietary intake to generate personalized diet plans for each user. This system is available as a smartphone application that is user-friendly and portable. Based on a survey, the application has been found to assist users in regulating and decreasing their blood pressure.

A recommendation system has been created for dietary food in [6], which takes into account patients' health conditions and other features, by utilizing gated recurrent network and K-clique algorithms. The dataset obtained from the internet has been processed, encoded, and organized based on similarities before being utilized to train the model. The gated recurrent network algorithm was utilized to enable faster training of the model. The developed design has undergone training, testing, and cross-validation. The outcomes of these processes demonstrate that the proposed system outperforms other machine learning and deep learning procedures that are similar to RNN, MLP, Navies' Bayes and Logistic Regression and in terms of precision and accuracy.

III. Proposed Methodology

The diet recommendation system utilizes algorithms and data analysis to provide personalized diet plans and recommendations to users. The systems take into account various factors such as gender, age, height, weight, body mass index. Personalized diet is provided with the use of K-means algorithm which categorizes the given food items on the basis of breakfast, lunch and Dinner then Random Forest algorithm is used to choose the food items based on its attributes such as calories, amount of Calcium, Sugar, iron, iodine, etc.

A. *About the Dataset*

The dataset consists of nutrition values of 90 food items. The columns are display the different nutrition values for example: Amount of Calories, fat, sugar, Calcium, Potassium present in a particular food item. The feature selection depends on the specific food nutrient on which the recommendation is focused.

	A	B	C	D	E	F	G	H	I
	Food_iter	Breakfast	Lunch	Dinner	VegNovV	Calories	Fats	Proteins	Iron
2	Asparagus	0	1	1		22	0.2	2.4	0.91
3	Avocados	1	0	0	0	160	15	2	0.55
4	Bananas	1	0	0	0	89	0.3	1.1	0.26
5	Bagels ma	0	1	1	0	250	1.5	10	2.76
6	Berries	1	0	0	0	349	0.4	14	6.8
7	Brocolli	0	1	1	0	25	0.5	3.8	1.27
8	Brown Ric	0	1	1	0	362	2.7	7.5	1.8
9	Cauliflow	0	1	1	0	32	0.3	3	0.72
10	American	1	0	0	0	331	24	20	0.84
11	Coffee	1	0	0	0	2	0	0.3	0.02
12	Corn	1	1	1	0	97	1.4	3.3	0.55
13	Dark choc	0	0	1	0	556	32	5.5	2.13
14	Grapes	1	0	0	0	93	2.1	5.6	2.63
15	Milk	1	0	1	0	97	6.9	3.8	0.12
16	Cashew N	1	0	0	0	553	44	18	6.68
17	Onions	0	1	1	0	40	0.1	1.1	0.21
18	Orange	1	0	0	0	97	0.2	1.5	0.8
19	Pasta canr	0	1	1	0	71	0.7	2.2	0.91

Fig. 2. Dataset used

B. *K – Means Algorithm*

K-Means algorithm is an un-supervised algorithm which is used to group unlabeled dataset into different clusters.

The k-means algorithm begins by selecting the desired number of clusters (k) and randomly assigning centroids to them. Each data point is then assigned to the nearest centroid, creating k clusters. The centroids of each cluster are adjusted by computing the mean of the data points within that cluster.

The process is repeated until the centroids remains stable or until a maximum number of iterations allowed is attained.

The procedure for the algorithm is as follows:

1) *Selecting the Number of Clusters to be made*
2) *Selecting the centroids of clusters randomly.*
3) *Assigning datapoints to each of the predefined clusters based on the distance from the centroid of each cluster.*
4) *Calculating the variance and recalculating the centroids of each cluster.*
5) *Reassigning each datapoint to the new closest centroid of each cluster.*

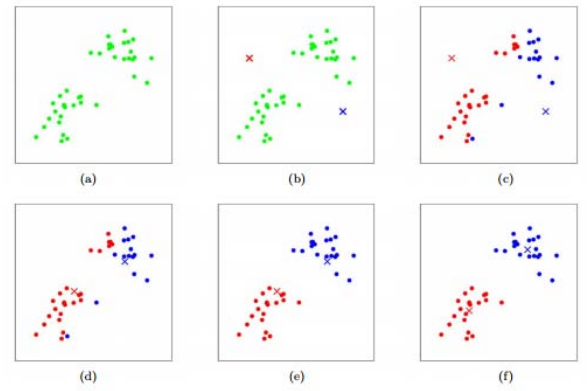


Fig. 3. Steps of K-Means Algorithm

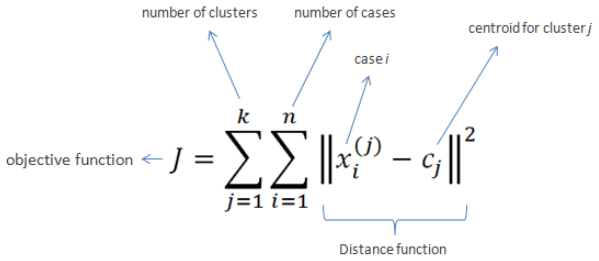


Fig. 4. Euclidean distance formula from centroid

C. *Random Forest Algorithm*

The popular machine learning algorithm Random Forest is a component of the supervised learning approach. It can be used to solve problems using ML Classification and Regression. The principle of ensemble learning, which is the act of combining different classifiers to solve a difficult problem and improve the performance of the model, serves as its theoretical underpinning.

In order to improve the predicted accuracy of the dataset, the Random Forest classifier, as its name suggests, "contains a number of decision trees on diverse subsets of the input dataset and takes the average." The random forest uses forecasts from all of the trees, rather than relying on just one, to predict the result based on which predictions earned the most votes.

Random Forest works based on the following assumptions, since there is a chance that while some decision trees will correctly predict the output, whereas some others decision tress might not. But together, these trees will correctly predict the right output. Therefore, these 2 assumptions are to be made for better forest classifier:

- There must be some actual values in the dataset for the feature variable to forecast actual outcomes as opposed to speculative outcomes.
- The prediction of each tree must have very low correlations between them.

First, N decision trees are combined to generate the random forest, and then predictions are made for each tree that was produced in the first phase.

The Working Process is as follows:

1. Select K random data points for the dataset.
2. Make decision trees connected to the selected data points. (Subsets)
3. Choose N for any decision trees you plan to build.
4. Repeat the 1st and 2nd steps.
5. Locate the forecasts from each decision tree for any new data points, then group them into the category with the most support.

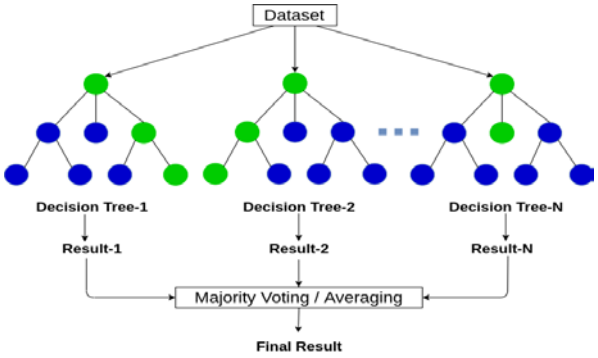


Figure 5. Working of the Random Forest Algorithm

IV. PROPOSED WORKFLOW

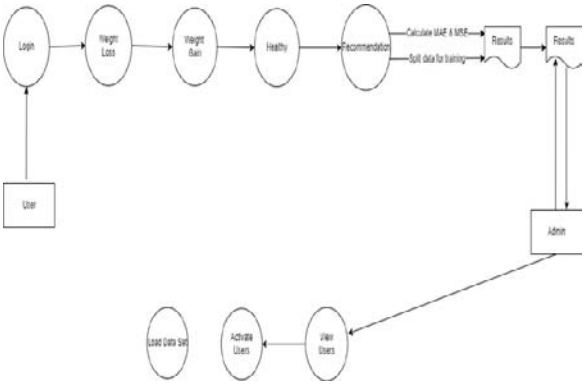


Fig. 6. Architecture Diagram for the Project

D. *Front End Workflow*

At the Front end, the user has to register to the application initially. After registering, the user enters their health information i.e. User weight, Height , Age and the nutrient deficiency they are suffering from. Based on the user information, the recommendation system in the back end of the project provides a suitable diet for the user which gets displayed on the user interface.

HTML, CSS and JavaScript languages are used to develop the application and Django is used to provide the back-end server-side framework for the recommendation system. The key reason of using Django is that it provides built in features which allow the development of the application to be quick and efficient. Django is also highly scalable and is capable of handling heavy workload.

E. *Data Preprocessing*

In order to separate the breakfast, lunch and dinner food items one hot encoding is performed on the data . Hence,

DOI: rp-9788770040723.216

food items are categorized based on the meal in which they are consumed.

F. Recommendation System

The clustering of various food nutrients based which are essential for weight loss, weight gain or dealing with a specific nutrition deficiency is performed. Clusters for the nutrients are made with the use of K–Means Algorithm. With the help of K–means algorithm nutrient specific clusters are formed. Using these clusters, food items are recommended based on the user input.

Random Forest Classifier takes the k -means based clusters as input and predicts the food item for the user based on the user's input. The nearest food items which are best suited for the user are hence predicted. Random Forest classifiers provides the output from various decision trees and it makes sure the best decision tree is used to predict an outcome.

This process is performed on breakfast, lunch and Dinner food items and a list of all the recommended items is displayed to the user as final output.

V. Result and Output

- Install Django, Pandas and other required packages to the device that will act as a server for the web application
- Run the "manage.py" script as sever using the command "python manage.py runserver" in the command prompt with the file location as target.
- Once the server starts running access it using any web browser
- The user has to register by selecting the "REGISTER" option in the top-right side of the home page and entering the required information of the user in the "User Registration Form".

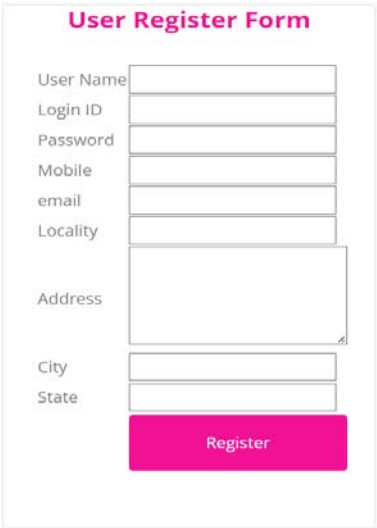


Fig. 7. User registration Form

- Once registered the user needs to be accepted by the admin in order to login and use the service.
- After register with username and creating a password to use this system. Login with registered username and password.

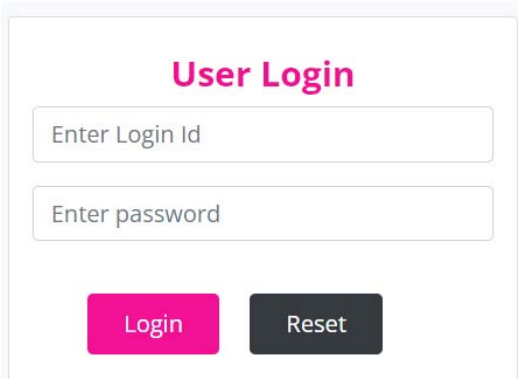


Fig. 8. Login page

- Once logged in, the user can get their diet based on their BMI or deficiency by selecting "BMI-DIET" or "DISEASE-DIET" on the top right corner.
- Once the user, enters their age, Veg/ Non_veg preference, weight and height in the "BMI-DIET" page and selects recommend option the system will give a recommended list of items for consumption based on the users BMI Index.
- User profile is created by height, weight and age, to estimate the BMI of user.

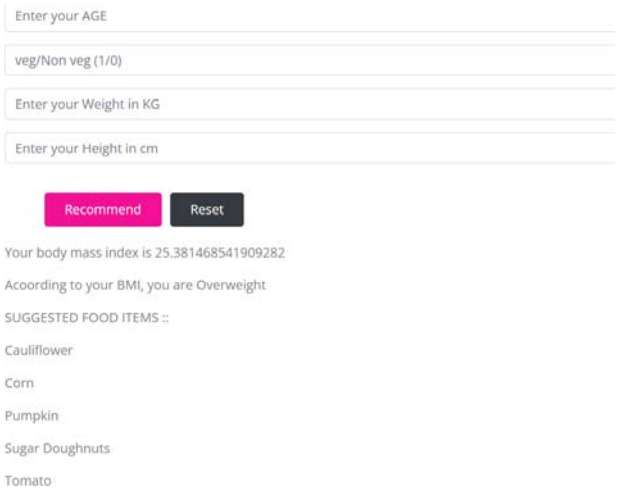


Fig. 9. BMI-DIET page with output

The diet is recommended in such a way as to return the users BMI back to the normal range by controlling the calories and making sure that no other nutrients levels are compromised.

This enables the user to follow a health and well-balanced diet based on his BMI in order to tackle obesity and underweight disorders.

- Once the user selects the nutritional disease on the "DISEASE-DIET" page the system recommends a list of

DOI: rp-9788770040723.216

food items recommended for consumption in order to prevent/fight the disease.

Figure 10. DISEASE_DIET page with output for diabetes

- The food items recommend in the figure 10, are suitable for users to fight/prevent against diabetes
- The nutritional content of the food items suggested are as given below in the figure 11

Food_items	Breakfast	Lunch	Dinner	Veg/Non/Veg	Calories	Fats	Proteins	Iron	Calcium	Sodium	Potassium	Carbohydrates	Fibre	VitaminD	Sugars
Asparagus Cooked	0	1	1		22	0.2	2.4	0.91	23	14	224	4.1	2	0	1.3
Brocolli	0	1	1	0	25	0.5	3.8	1.27	118	56	343	3.1	2.8	0	0.6
Cauliflower	0	1	1	0	32	0.3	3	0.72	32	259	278	6.3	3.3	0	0
Coffee	1	0	0	0	2	0	0.3	0.02	2	1	50	0.2	0	0	0
Onions	0	1	1	0	40	0.1	1.1	0.21	23	4	146	9.3	1.7	0	4.2
Pumpkin	0	1	1	0	18	0.1	0.7	0.57	15	237	230	4.3	1.1	0	2.1
Tomato	1	1	1	0	16	0.2	1.2	0.47	5	42	212	3.2	0.9	0	2.63
Mushrooms	1	1	1	0	22	0.3	3.1	0.5	3	5	318	3.3	1	7	2
Green Tea	1	0	0	0	1	0	0.2	0.02	0	1	0.2	0	0	0	0
Tea	1	0	0	0	1	0	0	0.08	2	1	9	0.2	0	0	0
Strawberries	1	0	0	0	32	0.3	0.7	0.41	16	1	153	7.7	2	0	4.9

Fig. 11. Nutritional info of the food recommended

The user that are suffering from diabetes need to take care of their sugar/glucose intake and minimize it in order to fight/prevent diabetes. These food items have very low amount of sugar and glucose content when compared to other food items.

These food recommendations not only cut down the users intake of sugar but also allows users to take a balanced amount of other nutrients and stay health and fit.

These food nutritional values are based on USDA organization. USDA is a food database that contains the food items and their nutritional values.

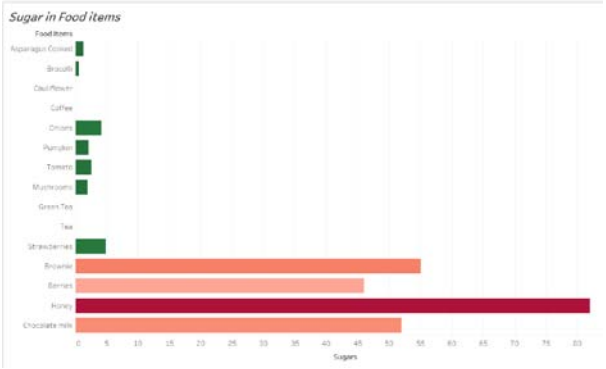


Fig. 12. Comparison of sugar content in recommended food items and other food items

Figure 12 shows the sugar/glucose content of the recommended food items in comparison with the other food items not being recommended by the system in order to prevent diabetes.

The food items that are not recommended by the system and have a high quantity of sugar content in them are highlighted in red, whereas the food items recommended are coloured green, these food items are suitable for consumption for the user as they provide with all the other necessary nutrients, making a well-balanced diet.

- Once the user is done with the information and the web page, he can logout of the webserver or close the webpage as a whole.

V. Conclusion

A diet recommendation system that caters to the user's nutritional demands was successfully built. Using the user's BMI and diet preferences, a personalized diet was generated from the list of food items available in the dataset used for this project which can help them improve their health and lead a disease-free life. The complex task of searching for a diet which is capable of providing all the nutrients required by the human body was made easy with the help of Machine Learning. By following the nutrition recommendations, users can maintain and enhance their overall health.

VI. Future Scope

As we all know, Exercise and a healthy diet are essential components of a healthy lifestyle. In addition to engaging in regular physical activity, it is crucial to maintain a balanced and diverse diet that includes fruits, vegetables, whole grains, lean proteins, and healthy fats in order to achieve optimal health. By combining these two lifestyle factors, the body can receive the necessary nutrients to function efficiently and maintain overall well-being. Keeping in mind the importance of physical exercise, an additional functionality to the present project which provides the user a regular exercise routine according to their health status can help improve the results in terms of a person's health. The user can follow the workout plan provided to them and be aware of the number of calories which are burnt while following the exercise routine. This functionality will improve the overall health of the user.

References

[1] Shreya B. Ahire, Harmeet Kaur Khanuja, 'A Personalized Framework for Health Care Recommendation', International Conference on Computing Communication Control and Automation, Pune, India, 26-27 Feb. 2015.

[2] Rajesh, M., &Sitharthan, R. (2022). Introduction to the special section on cyber-physical system for autonomous process control in industry 5.0.Computers and Electrical Engineering, 104, 108481.

[3] Poojitha Amin, Nikhitha R. Anikireddypally, Suraj Khurana, Sneha Vadakkemadathil, Wencen Wu,'Personalized Health Monitoring using Predictive Analytics,' IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService), Newark, CA.

[4] P.Kavya, M.Ramya Sri, Ch.Kavya Sri, K.Jagadish, S.Srikari, "Automated Diet Recommendation System", International Journal of Research in Advanced Computer Science Engineering, (IJRACSE), vol. 7, issue 1, pp. 1-9, June 2021.

[5] Sitharthan, R., Vimal, S., Verma, A., Karthikeyan, M., Dhanabalan, S. S., Prabaharan, N., ...&Eswaran, T. (2023). Smart microgrid with the

internet of things for adequate energy management and analysis.Computers and Electrical Engineering, 106, 108556.

[6] Samuel Manoharan, Sathesh Ammayappan "Patient Diet Recommendation System Using K Clique and Deep learning Classifiers" Journal of Artificial Intelligence and Capsule Networks (2020) vol.02, no. 02