International Conference on Recent Trends in Data Science and its Applications
DOI: rp-9788770040723.097

Diet Recommendation System Using ML

Ananthajothi K
Computer Science and Engineering
Rajalakshmi Engineering College
Chennai, India
ananthajothi.k@rajalakshmi.edu.in

Suganthi M
Computer Science and Engineering
Rajalakshmi Engineering College
Chennai, India
suganthi.m.2019.cse@rajalakshmi. edu.in

Sujitha P
Computer Science and Engineering
Rajalakshmi Engineering College
Chennai, India
sujitha.p.2019.cse@rajalakshmi.edu.in

Visalatchi R
Computer Science and Engineering
Rajalakshmi Engineering College Chennai, India
visalatchi.r.2019.cse@rajalakshmi.edu.in

***Abstract*—Diabetes interacts with the capacity of our bodies to turn food into energy. It is a chronic health issue. Your pancreas is prompted to generate insulin when the sugar levels in your blood rise since an enormous portion of the food you consume gets transformed into sugar (glucose) and communicated into blood for circulation. Nowadays, deep learning is used in the healthcare industry to try and forecast diseases. Data is the most essential component of deep learning. p A deep learning model is constructed using the previous dataset, which is gathered. Univariate and bivariate analysis, among other essential pre- processing techniques, are used. A classification model is built using a machine learning algorithm only after data is shown to facilitate with feature interpretation. Algorithms are evaluated based on performance metrics like accuracy, F1-score, recall, etc.**

Keywords— BMI, Deep learning, Dietary plan, F1score, Diabetes

I. INTRODUCTION

Humans face a wide range of health problems, including mental and physical health problems. Numerous studies demonstrate that in adequate dietary intake and poor nutrition quality are the primary cause for numerous illnesses and health issues. Heart attacks, ischemic heart disease, and gastrointestinal cancer are the leading causes of death around the world, according to WHO research. Everyone should eat a well-balanced diet to overcome this. A healthy diet is essential for your organs and tissues to function properly. A healthy diet can help strengthen the immune system and prevent disease, according to medical research. A balanced diet includes a wide range of substances, notably water, vitamins, minerals, carbohydrate, proteins, and fats. Medical research has revealed that proper diet plan helps to build up the immune system and fight against diseases. Consumption of proper diet provides energy, vitamins, carbohydrates, proteins, fats, vitamins, minerals, and water multinomial analysis and random forest algorithm is to be integrated to provide healthy diet plan recommendation according to user characteristic.

This diet advice project intends to convey nutrition information in an adaptive manner to enhance the life of both healthy and sick people. The project offers food recommendations based on the attributes and assists users in keeping track of their calorie target based on their BMI. The system offers culinary recommendations to users using a ML algorithm. The system informs the user of the effects and reasons for ignoring the suggested diet.

The existing system just creates a week diet plan based on user attributes like height, weight, and BMI. This application technique develops a diet plan and health recommendations using the random forest algorithm. The system will generate a diet plan based on a healthy diet, weight growth or loss, past user assessments of food flavor and food type with suitable calorie levels, and information about weight increase or loss (carbohydrate, protein, vitamins, calcium, fats, fiber and iron). The strategy will help consumers become healthier and identify the kinds of foods to stay away from to lower their risk of illness.

II. LITERATURE SURVEY

The objective of the diet recommendation project is to improve the health and wellbeing of individuals, both those who are generally healthy and those with chronic illnesses. To achieve this, the project provides personalized nutrition guidance that takes into account the user's features and assists them in monitoring their calorie intake based on their BMI. The system offers nourishment suggestions that are tailored to the user's preferences and provides explanations for why it is important to follow the recommended diet. Users can also learn about the latest nutrition news and health tips in the system's section.

In [1], The system was built with a strong emphasis on developing a seven-day feeding plan based on an individual's personal needs and dietary preferences. It incorporates a content-based filtering approach that employs Euclidean Distance to offer alternatives for food allergies or dislikes. Due to the variety of preferences, likes, and dislikes among people, it also addresses the usage of the Pearson correlation coefficient for nutrient assessment and alternative meal selection.

The platform where consumers can locate their preferred food and its nutritional value. Anyone who cares about the health or wants to lose weight can benefit from this. This application can be used independently or as part of a more sophisticated application. Web scraping was used to collect the data set, which was then pre-processed based on attributes. The Content Based Filtering Algorithm was then used in the food recommendation system [2].

The approach was created to take into account the latest innovations in dietary evaluation setups, which are a more beneficial way to monitor daily food intake and manage dietary patterns.

A framework for estimating food, calories, and nutrition is incredibly helpful for individuals' groups to measure and manage their daily food consumption. AI will offer recommendations based on the benefits or prior experience that the customer has had. These computations are also used to make diet recommendations to the customer. [3].

DOI: rp-9788770040723.097

In [4], An Android application was developed to provide consumers a personalized diet. Similar to a genuine dietitian, it serves as a nutrition advisor. This system operates similarly to a nutritionist. A person must provide the nutritionist with certain details, such as their weight, height, gender, etc., in order for them to gain knowledge about their diet plan. In a similar manner, this system also offers a diet plan based on the data given by the user. To provide the user with the diet plan, the system collects all of the user's information and analyzes it.

The system was designed to operate in a machine learning environment. It evaluates user information comprising age, gender, height, weight, and body fat percentage, as well as choices for weight loss or weight gain, and then recommends dietary habits in three categories for lunch, breakfast, and dinner. Problems encountered in relation to the user. This system uses K-Means and then Random Forest algorithm approach. Problems encountered in relation to the user. The Nutrition dataset from Kaggle is used for recommendation [5].

The study suggests that individuals with diabetes should have snacks due to their specific circumstances and the risk of hypoglycemia between meals. To generate a list of significant items based on user characteristics, environment, and behavior, the approach employs various strategies such as collaborative sifting, knowledge-based, roulette wheel, and content-based sifting (CF) algorithms. In this paper, a recommendation system for various food ingredients is proposed, considering the small test measurement during the assessment phase and the absence of the primary meal. The proposed recommendation system suggests ingredients that are frequently consumed together. [6].

In [7], The recommendation system makes use of a web interface to take into account user inputs and make recommendations for food items based on the user's current requirements. Random Forest Classification, K-Means Clustering, LSTM and Python-based Tkinter GUI. The user interface and dataset in this system need to be improved. Regarding the recommendation of recipes, a system plan for putting recipe search into action takes into account the combination of numerous food ingredients, nutrition, and budget. This app is extremely useful for cooking a wide range of recipes with minimal internet research. Additionally, this app lets users create and save recipes for later reference. It will save people time and effort searching for recipes for both everyday meals and special occasions.

The suggested algorithm improves classification accuracy while producing effective classification. There are numerous approaches for the classification of medical data utilizing various metrics and techniques, and the subject of categorization in medical data has been well investigated. The approaches still struggle to improve classification accuracy performance. The classification algorithm was developed using MLIIM is employed in this study to increase classification effectiveness. By removing the incomplete information, the tool estimates the submitted collection of data to resolve the interference problems. During the subsequent stage, the algorithm periodically and at various levels examines the impact of the measure. Additionally, the method determines the class influence weight (CIW) for distinct classes. [8].

The approach offers personalized suggestions mostly to customers who are dealing with issues like diabetes or heart disease. Dynamic product recommendations enable for the provision of an appropriate product that adapts to consumer preferences over time. The algorithm for product categorization is in charge of data cleansing, tokenization, and frequency calculation. The association between the product and the user should be established via genetic algorithms. The food goods in this application must be acquired by the product buying module. The project doesn't include more product categories so that it can offer recommendations to different people who like items with distinctive nutritional benefits [9].

In [10], The system proposed the target by using their BMI as an input and by taking into account their allergies to let them track their daily calorie intake. Collaborative Filtering algorithm with fuzzy logic. As it is, collaborative filtering systems might need substantial financial support as well as a lot of computational power. Algorithms of machine learning are used in this research. Custom recommendations are a feature that has been added to this project specifically for individuals who have had coronary artery disease, diabetes, or hypertension. Here, nutrition is divided up into four categories: salts, sugar, energy, and proteins. Although the user's or buyer's genetic information is always constant, the product recommendations are dynamic in order to deliver an accurate product that consistently meets user preferences.

The Proposed is a system that aims to provide a comprehensive solution to the intelligent systems layer. The objective of the food recommendation approach is to establish personalized meal plans based on the user's nutritional needs and food preferences. Previous research has focused on developing computational tools for food intake recommendations, but few have integrated user preferences and dietary data. This approach combines a short-term intelligent system with an optimization scenario that considers both tangible and intangible nutritional and personal taste knowledge, as well as nutrient context systems that use an MCDA approach to filter out inappropriate food choices. [11].

In [12], The system uses machine learning techniques to examine users' data such as age, gender, height, weight, body fat percentage, and preferences, such as weight loss or weight gain, to suggest diet plans for three categories: breakfast, lunch, and dinner. The system employs the K-Means and Random Forest algorithms to address individual challenges. Additionally, the system uses a dietary dataset from Kaggle for recommendations.

The proposed system that considers both the major meals, diabetic patients should have appetizers because of a scheme that considers their special circumstances and the danger of hypoglycemia. This system offers suggestions for a list of items that could be useful based on the user's traits, situations, and behavior. In this study, a system for recommending snacks to type II diabetics is developed and evaluated. It makes use of a roulette wheel algorithm, a

DOI: rp-9788770040723.097

knowledge-based approach, and constraint-based reasoning. This system implements a content- based filtering system that uses Euclidean Distance to suggest substitutes for food intolerances or distaste and generates a seven-day food plan based on an individual's requirements and food tastes [13].

In [14], The proposed system recommends a list of food items based on the patient's medical conditions and nutritional value. The user selects a main culinary component, and the system provides a recommendation for the ideal food to eat from the database based on the patient's health. The system evaluates the user's health and suggests foods based on any chronic illnesses or ailments they may be dealing with. The user's medical history is obtained from the specific hospital as soon as they log in, and the application offers the appropriate meal based on their health issue or disease. The system operates precisely and fully maintains the user's health.

The proposed system aims to enhance people's diets by responding to dietary advice with personalized, nutrient-rich recommendations. To calculate these recommendations, the system will rely on the USDA nutrient dataset, which contains valuable information on the nutritional value of various foods. Additionally, users will need to provide their BMI values as an input, which will be used to tailor the recommendations to their specific needs. Another key input for the system is the user's daily food intake, which will be used to calculate the deficit nutrition and identify the missing nutrients that need to be included in the recommendations. To ensure the recommendations are as effective as possible, the input nutrients dataset will be sorted based on the user's BMI value. Finally, any deficit nutrients will be filled by selecting the appropriate foods from the sorted grocery dataset, resulting in a personalized, well-balanced diet recommendation for the user. [15]

III. METHODOLOGY

A. *Data Generation*

Data generation involves collecting datasets from various sources, which can be labelled or unlabeled. If the data is unlabeled, it needs to be manually labelled to represent the class to which the object belongs. This labelling helps the learning model to identify the particular class when it encounters data without a label. As there were no publicly available datasets that considered specific user conditions such as kidney stones and diabetes, the data for this project was collected using web-scraping. Since the proposed method involves multi-label classification, the raw unlabeled data was manually labelled into chronic, anemic, and diabetic data, which were further labelled as breakfast, lunch, or dinner.

	A	B	C	D	E	F
1	cholesterc	glucose	Nutritiona	Hemoglob	age	gender
2	193	77	49	14.9	19	female
3	146	79	41	15.9	19	female
4	217	75	54	9	20	female
5	226	97	70	14.9	20	female
6	164	91	67	14.7	20	female
7	170	69	64	11.6	20	female
8	149	77	49	12.7	20	female
9	164	71	63	12.7	20	male
10	230	112	64	14.1	20	male
11	179	105	60	14.9	20	female
12	174	105	117	13	20	male
13	193	106	63	16.7	20	female
14	132	99	34	13.4	21	female

Fig. 1. Data Generation

B. *Data Preprocessing*

Data pre-processing is performed on raw data. It has traditionally been an important preliminary step. It includes cleaning, instance selection, transformation, feature extraction, and selection. It eliminates the irrelevant noises in the given input. The data collected in this project is obtained through web-scraping. Hence pre-processing of them had to be done to adjust them to the format of machine learning methods. The collected raw data may be labelled or unlabeled. The raw unlabeled data are manually labelled into chronic and diabetic data.

C. *Multinomial Analysis*

Supervised artificial intelligence faces several challenges, and one of the most common ones is multiclass classification. Multiclass classification problems arise when there are at least two types of outputs that need to be classified. This involves organizing data into categories based on their similarities and differences. To achieve this, we need to identify and select relevant features or independent variables that are critical for grouping the data into different categories. Multiclass classification is the process of separating data into multiple categories based on our dependent variable of interest.

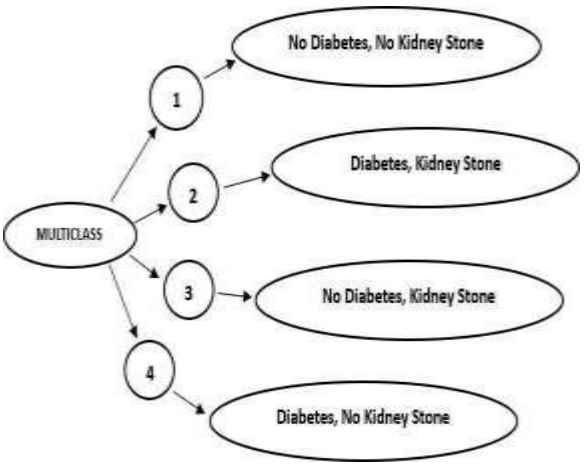

Fig. 2. Multiclass classification

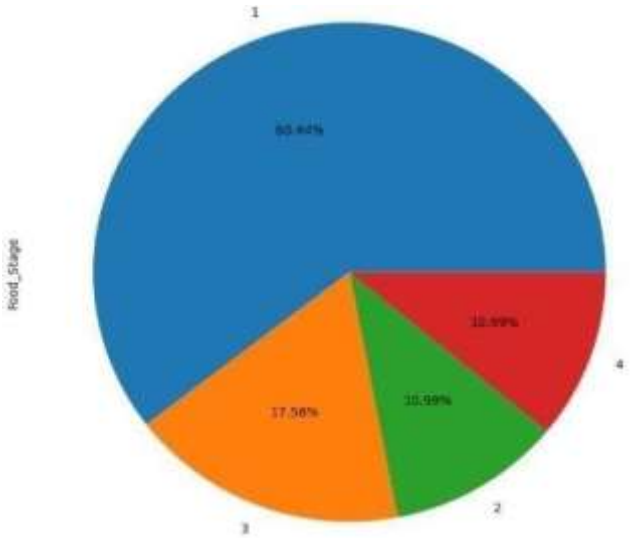

Fig. 3. Multinomial Analysis

DOI: rp-9788770040723.097

D. Experimental Environment Setup

The system architecture diagram provides a visual representation of the system's architecture, illustrating the connections between various system components, along with the purpose of each component. The general system representation depicts the primary system functions and the connections between different system parts. Pre-processing is done on the food dataset here, and the pre-processed data is then subjected to multinomial analysis and clustering to generate diet recommendations.

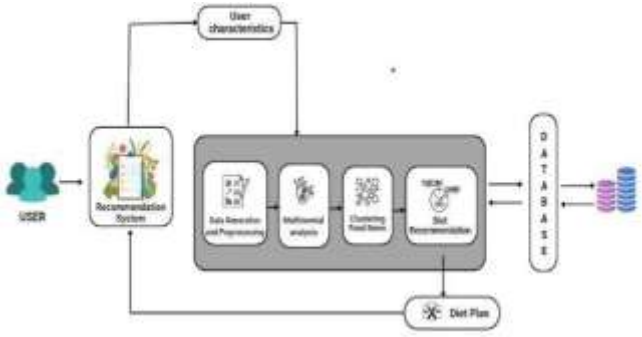

Fig. 4. System Architecture

E. Machine Learning Algorithms

1) Support Vector Machine

SVM among the most commonly employed methods for supervising learning, that is commonly used for classification and regression task. With the help of the SVM algorithm, researchers can classify new data points to make it easier in the future by creating a line of best, or decision boundary, that divides the n-dimensional space into groups. A hyperplane is this optimum decision boundary. SVM operates for producing the hyperplane by choosing the extreme points and vectors. The scientific name for the optimization technique as an SVM— one of the most popular techniques for Supervised Methods sustained vectors, and these extreme examples are defined as supported vectors.

2) Random Forest Classifier

Random Forest is popular ml algorithm that is commonly used as regression and classification task. The algorithm based on concept of ensemble learning, which involving combining many classifiers to enhance the accuracy of the particular model. It is a type of model that uses an ensemble of the decision tree, which are trained on using subsets of the provided datasets. The algorithm then takes mean of that predictions of all decision tree to enhance accuracy of the overall prediction. Unlike relying on just one decision tree, a Random Forest makes predictions based on the collective assessments of all decision trees in the ensemble, selecting the result based on the majority of predictions. By using multiple decision trees in the ensemble, Random Forest can increase accuracy and avoid the sampling.

3) Decision Tree Classifier

Decision trees classifiers can also be used for feature selection, where the most important features are identified based on their ability to split the data and classify instances correctly. This can help can also help to reduce the dimensionality of the input space. A decision tree is composed of a decision node and a child node. A leaf node represents the outcome of a decision without any subsequent branches, while a decision node represents the decision and has multiple branches. To make decisions, decision trees consider the features of the provided dataset. A decision tree is a visual representation of every possible solution for solving a problem or making a decision on given parameters. It is known a decision tree because it starts at root of the tree and grows in subsequent branches, resembling a tree. The Regression and Classification Tree (CART) technique is used to construct a decision tree.

4) MLP Classifier

Neural networks called multilayer perceptron produce result from predetermined inputs. A graph which is directed connect output layer and input layer of mlp consists of a several layers of modern output nodes. A multilayer perceptron is a powerful learning technique that creates an organization. Originally developed for image recognition, perceptron commonly referred to as computations. It carries out the human-like task of seeing, viewing, and it carries out human-like task of seeing, viewing, and identifying images. In essence, a multilayer perceptron is a cutting-edge neural structure with three different sorts of layers: input, output, and hidden. An input flag is delivered to an input layer for processing. The output layer carries out activities like prediction and categorization.

F. Clustering Food

The system could recommend a healthy diet plan according to the personal condition. Food is clustered into breakfast, lunch, dinner for recommendation. For the diet recommendation, we have devised our own algorithm which splits the foods into nutrients like fats, fiber, vitamins, etc.

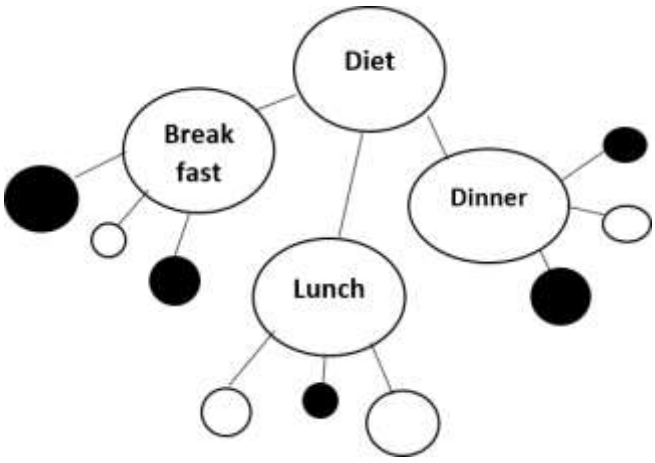

Fig. 5. Clustering food

And based on the nutritional requirements needed for a day a balanced diet is provided. The system also displays the BMI value under consideration and provides a diet plan to maintain them at normal. system could recommend the related diet plan to fulfil personalized needs by using multinomial analysis. Therefore, it would provide better service and experience for users.

DOI: rp-9788770040723.097

IV. Results and Analysis

Evaluating recall, F1 score and precision are commonly used to know the performance of classification models in accurately identifying instances belonging to a particular class. These metrics are crucial in determining the suitability of machine learning models for real-world applications and assessing their effectiveness. The Random Forest algorithm is a highly effective method, capable of handling complex nonlinear problems, managing large input datasets, and providing fast predictions. It has demonstrated an accuracy of 99% on various datasets.

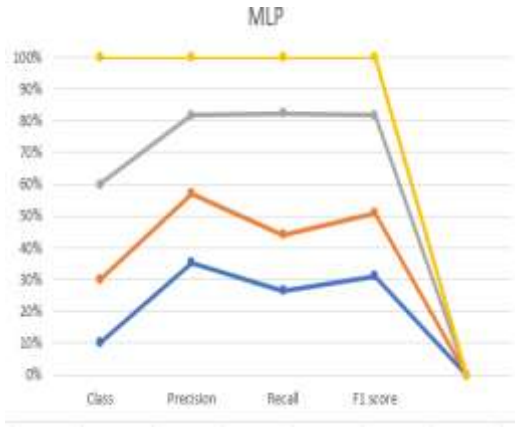

Fig. 6. MLP Analysis

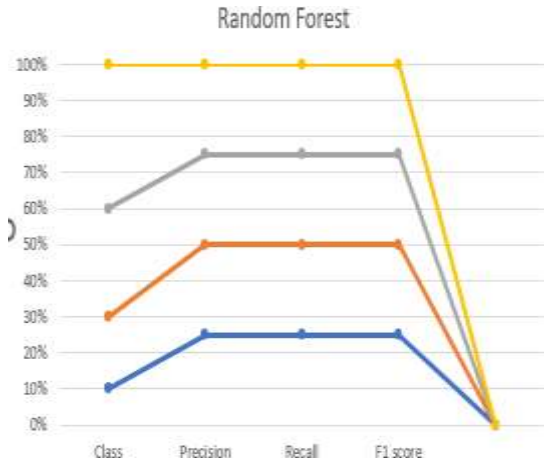

Fig. 7. Random Forest Analysis

Fig. 8. SVM Analysis

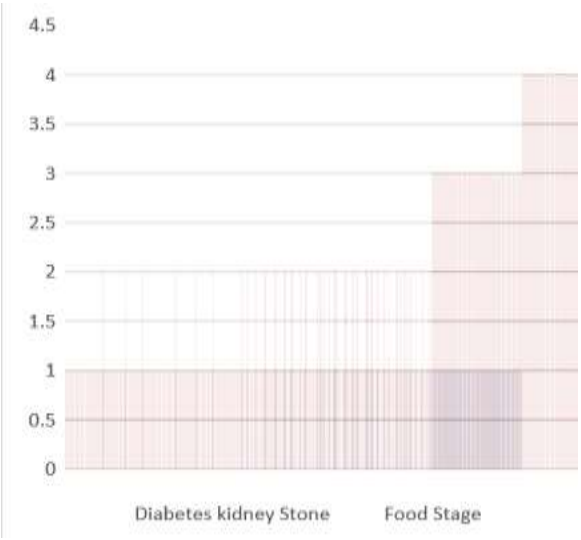

Fig. 9. Diabetes Kidney stone analysis

V. Conclusion

This study aimed to develop a web application that could provide personalized diet plans for patients based on their health status and other characteristics. The system used a machine learning algorithm to generate nutritional recommendations tailored to each patient's unique needs and medical conditions. Patient data was collected from various online sources, sorted, preprocessed, coded, and analyzed for similarities in order to train the model. Rapid prototyping was used to create and test the models using web scraping techniques. The results of the training and testing phases demonstrated that the proposed system was highly accurate and precise in generating personalized diet plans for patients.

References

[1] B.A. Ojokoh, "A personalized healthy diet recommender system", IEEE Conference IEEM-2014, pp. 798–802, 2016.

[2] Kundan Shumsher Rana, "Food recommendation based on content filtering algorithm", 2nd International Conference IT-DREPS, Amman, Jordan, pp. 1–5, 2018.

[3] Prajkta Bhosale, "Diet Recommendation Using Machine Learning", in Journal of Emerging Technologies And Innovative Research (JETIR), pp. 163–174, 2022.

[4] Garvita Gehlot, "Dietexpert – Android Application For Personal Diet Consultant", International Journal of Engineering Science and Computing, vol. 6, pp. 7113-16, 2020.

[5] V. Paraloganatha, "A Nutrient Based Diet Plan Recommendation System using Machine Learning", 5th International Research Conference of Uva Wellassa University, IRCUWU2021, Paper ID: IRCUWU2021-122, 2021.

[6] Zhengxian Li, "Recipe Recommendation Based on Ingredients using Machine Learning", International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), 2018.

[7] Jiazhao Shen, "Recipe Recommendation Based on Ingredients using Machine Learning", International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), 2018.

[8] V.Nallarasan "Improved Predictive Learning Approaches for Customized Diet Suggestion Framework in Medicinal Services",2020.

[9] 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.

[10] Manoj P, "Feedback Food Recommendation System Using Hybrid Deep Learning", International Journal of Innovative Science and Research Technology, ISSN No: -2456-2165, 2020.

[11] R. Prabakaran, "A Survey on Recommendation Systems using Collaborative Filtering Techniques", 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), 2023.

[12] 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.

[13] Zhu, M, "Content boosted matrix factorization for recommender systems: experiments with recipe recommendation," in Proc. of the fifth ACM conference on Recommender systems, pp. 261-264, 2011.

[14] Sajad Ahmadian, "Healthy Food Recommendation Using a Time-Aware Community Detection Approach and Reliability Measurement", International Journal of Computational Intelligence Systems, vol. 15, no. 1, 2022.

[15] Megh Shah, "Diet System Based On Different Machine Learners", 2nd International Journal on Artificial Intelligence and Smart Energy (ICAIS), pp. 2456-2165, 2022.