DIET RECOMMENDATION SYSTEM USING MACHINE LEARNING

[1]19981a05e2@raghuenggcollege.in, [2]19981a05h0@raghuenggcollege.in ,
[3]19981a05f9@raghuenggcollege.in , [4]kavitha.chekuri@raghuenggcollege.in*

Reema Golagana, V. Sravani, T. Mohan Reddy, IV year B.Tech students, Dept. of Computer Science Engineering, Raghu Engineering College, Visakhapatnam, Andhra Pradesh, India

CH Kavitha Assistant professor, Dept. of Computer Science Engineering, Raghu Engineering College, Visakhapatnam, Andhra Pradesh, India

ABSTRACT

In today's culture, it is not easy to recommend a diet right away. Nowadays, people suffer from a range of diseases and chronic illnesses. And for most cases, an unhealthy or improper diet is the root cause of these ailments. On average, a person requires 2000 calories a day, however, the precise number of calories consumed depends on a person's physical characteristics, including BMI, age, gender, cholesterol, blood pressure, and other factors. In this paper, we propose a diet recommendation system based on the user's physics details and ailments. This research includes different machine learning and deep learning techniques. The system learns from its training data, which includes daily calorie intake and food consumption patterns, to predict the recommended meal plan for a user. We evaluated the system on a dataset, and the results show that the proposed model outperforms the existing diet recommendation systems. The proposed system predicts the meal plan based on the user's dietary nutritional requirements. It could achieve a personalized diet plan for individuals and help individuals lead healthier lives.

Keywords: diet, calories, BMI, diet recommendation, machine learning, deep learning

1. INTRODUCTION

People today struggle with a wide range of health issues, including mental health issues, poor diet, and fitness issues. Numerous studies have demonstrated that a poor diet is a major contributor to a wide range of health issues. A WHO study found that an inadequate and unbalanced diet is responsible for approximately 9% of deaths from heart attacks, 11% of deaths from heart disease (Ischemia), and 14% of deaths from gastrointestinal cancer worldwide, reportedly causing. In addition, approximately 250 million children suffer from vitamin A deficiency, 200 million from iron deficiency (anemia), and 700 million from iodine deficiency[7]. A diet recommendation system is a tool that assists people in making healthy food choices. Diet recommendation systems analyze diet preferences, nutritional needs, and other dietary factors to provide personalized recommendations for healthier eating. By the user's age, height, and weight, BMI calculates weight status categories which include underweight, healthy, overweight, and obese. This project will help users with daily diet recommendations along with BMI range, healthy food choices, eating behavior, and health problems. Seasonal foods, user-intreated foods, plant foods, and animal products are all included in the grocery data. This project will assist users with daily dietary guidelines, BMI ranges, appropriate food selection, eating behaviors, health issues, and behavior modification. [8,9]

 DOI:10.36893.DRSR.2023.V13I04.118-125

There are several parts to the paper. An introduction to the use of machine learning techniques for diet recommendations is given in the first part. A review of related literature is presented in the second part. The third part of the study's methodology includes a description of the dataset, data pre-processing, algorithms used, and various techniques like Random forest, K-means, and LSTM algorithms. The findings and their analysis are presented in the fourth section. The conclusion, which concludes the essay, also goes over its potential scope.

PROBLEM STATEMENT

The alarmingly high rate of consumption of fast food has resulted in the intake of unhealthy food. Obesity, diabetes, high blood pressure, and other health issues are all a result of this. Thus it has become extremely fundamental for individuals to have a decent adjusted to a healthy, nutritional, and balanced diet. However, not everyone in today‘s society has the time or resources to afford a personal dietitian or a nutritionist to take care of their health and provide a balanced diet plan based on their unique circumstances. In this report, we discussed a person's poor nutritional eating patterns and made an effort to provide them with a solution that would enable them to lead healthy lifestyles.

2. LITERATURE REVIEW

This literature review examines papers and studies that investigated the application of machine learning algorithms to food recommendation systems for diets.

Celestine Iwendi et al. (2020) [1] investigate the data collection potential of their system. Machine and deep learning algorithms such as Naive Bayes, Logistic Regression, Multilayer Perceptron (MLP), Gated Recurrent Units (GRU), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) are the goals of this research framework and how they apply to IoMT data. For consideration in the clinical dataset, 30 people's information with 13 highlights of different illnesses, as well as 1000 things, were accumulated through the internet and hospitals. There are eight features in the product area. Before applying deep learning and machine learning-based techniques, the characteristics of this IoMT data were examined and further encoded.

Thi Ngoc Trang Tran et al. (2021) [2] show that their method can be used to present a comprehensive review of healthcare recommender system research: Besides, our exploration recognizes from past important outlines concentrates that it gives knowledge for suggested circumstances and approaches. Dietary recommendations, drug ideas, health status forecasts, service recommendations, and recommendations from healthcare professionals are all included in this kind of proposal. They also give students examples from real-world situations to help them fully comprehend recommendation systems.

Gao et al. (2017) [3] introduced a computational framework for a personalized diet recommendation system. The system used Bayesian personalized ranking along with matrix factorization to learn user preferences from a wide range of data. The results showed that the approach led to better-personalized recommendation performance than conventional collaborative filtering techniques.

Butti Gouthami and Malige Gangappa (2020) [4] The USDA nutrition dataset will be used to determine the user's suggested diet. A set of grocery shop information that takes the user's preferred food intake into account. The USDA database contains nutritional data for every food item. A USDA ID is used as the baseline value for the input values for every 100 grams. Since they will ultimately be used to estimate the suggested diet, the data required to compute BMI (body mass index) must be provided.

DOI:10.36893.DRSR.2023.V13I04.118-125

The second input utilized to determine the user's suggested diet is their daily food intake. The diet recommendation is initially created using the food that was eaten that day, and the input nutrient dataset is sorted by a nutritional deficit.

Yera et al. [5] proposed a food recommendation system that takes user preferences and nutritional data into account. The user's advised meal plan considered their preferences. This tool maintains both user preferences and nutritional information

Iigo Orue Saiz et al. [6] made an effort to find the research that has been done and the recommendation systems that have been used in the major databases for this purpose over the past five years. The results lead to the following conclusion: Previous works focus more attention on the recommendation system—typically collaborative filtering—than they do on the data or sample description; It is unknown which indices are used to calculate calories or nutrients. In this way, to be viable, it is basic to work with open information or all-around depicted information, which permits the experience to be rehashed by different gatherings, or at any rate to be comparable.

3. METHODOLOGY

I. DATASETS

The recommended diet for the user will be determined using the USDA nutrition information. Every food item's nutrition information is maintained in the USDA database. [5,19]

In this diet recommendation system, to calculate the recommended diet, the nutrient dataset is first sorted based on the BMI value, and deficit nutrition is determined using the food consumed that day. The recommender provides dietary advice. The samples of datasets used are in below figures (1) and (2).

Food_iten	Breakfast	Lunch	Dinner	VegNovVe	Calories
Asparagus	0	1	1		22
Avocados	1	0	0	0	160
Bananas	1	0	0	0	89
Bagels ma	0	1	1	0	250
Berries	1	0	0	0	349
Brocolli	0	1	1	0	25
Brown Ric	0	1	1	0	362
Cauliflowe	0	1	1	0	32
American	1	0	0	0	331
Coffee	1	0	0	0	2
Corn	1	1	1	0	97
Dark choc	0	0	1	0	556

Figure 1: Sample food dataset

DOI:10.36893.DRSR.2023.V13I04.118-125

Calories	Fats (gm)	Proteins(g	Iron(mg)	Calcium(m	Sodium(m
160	15	2	0.55	12	7
89	0.3	1.1	0.26	5	1
349	0.4	14	6.8	190	298
331	24	20	0.84	497	966
2	0	0.3	0.02	2	1
97	1.4	3.3	0.55	2	253
93	2.1	5.6	2.63	2	9
97	6.9	3.8	0.12	2	52
553	44	18	6.68	2	12

Figure 2: Sample nutrient dataset

II. DATA PREPROCESSING

Data preprocessing is a step taken to prepare data for modeling with LSTM (Long Short-Term Memory) networks. Preprocessing helps in transforming raw data into a format that is suitable for analysis, reduces noise, and enhances the quality of the data. Cleaning the data and making it suitable for a machine-learning model, which also improves a model's accuracy and effectiveness is a required task. Feature Engineering is the process of enhancing or selecting significant features from the dataset. The process involves identifying the relevant variables, removing unwanted features, and creating derived features.

Train-Test Split is partitioning the available dataset into training and testing sets is crucial in evaluating the model's performance. LSTM can memorize the training data too well and result in overfitting. Therefore, the model needs to be validated using an independent test set.

III. RANDOM FOREST

Random Forest is a popular machine-learning algorithm used in recommendation systems. It is particularly effective for solving high-dimensional and non-linear problems. Random Forest is a type of supervised learning algorithm that belongs to the family of decision trees. Unlike traditional decision trees that work with a single tree, Random Forest combines multiple decision trees to create a more accurate and robust prediction model. In a recommendation system, Random Forest can be used to create a model that predicts the user's preferences based on their past behavior and the behavior of similar users. The algorithm can analyze large datasets and extract patterns and insights that can be used to make personalized recommendations.[7,18]

IV. K-MEANS

K-means clustering is a popular unsupervised machine learning algorithm in various applications, including diet recommendation systems. If we have a dataset of users where each user has provided information about their food choices, allergies, and dietary restrictions, we can apply k means clustering to group similar users together. This can help in generating personalized diet plans, recipes, and food recommendations for each cluster. K-means clustering works by assigning each user to the nearest cluster based on their similarity.[18]

V. LONG SHORT-TERM MEMORY(LSTM)

An improved LSTM (long short-term memory) that accurately and precisely meets patient needs and overcomes this problem in the existing model. LSTMs use a combination of an input gate, an output gate, and a forget gate to selectively retain or discard information at each time step. This allows the

 DOI:10.36893.DRSR.2023.V13I04.118-125

network to maintain long-term dependencies in the input sequence and avoid the problem of information loss over time.

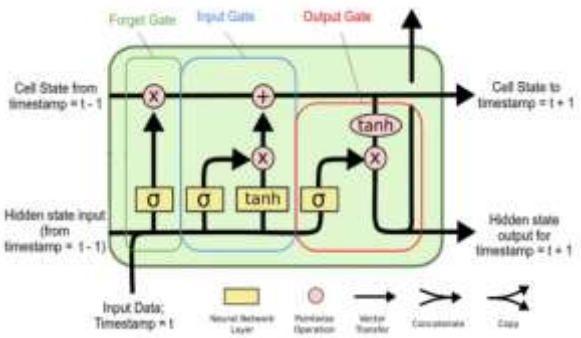


Figure 3: LSTM architecture[15]

- Forget Gate: The forget gate determines what information must be remembered and what can be forgotten. (C_t -1).
- Input Gate: The input gate is provided with new informational sets, which it learns and stores in memory (Ct).
- Output Gate: The output gate gives information to the next hidden state

LSTM model is trained on a dataset of food logs, where each log consists of a list of foods consumed in order. The LSTM could learn to predict what the user is likely to eat next based on their past eating habits and could generate recommendations for healthier alternatives. [13,15,16,17]

4. RESULT

This diet recommendation system can help individuals achieve their health goals by providing personalized dietary recommendations based on their age, weight, height, and preference, such as veg or non-veg. The recommender suggests 3 types of diets – weight loss, weight gain, and healthy. diet.

After entering the personal details in the input dialog box, the BMI is calculated and it will show based on the BMI if it's underweight, overweight, or healthy.

The Root Mean Squared Error (RMSE) is one of the two primary performance metrics. It determines the typical difference between actual values and a model's predictions. It provides an estimate of the model's accuracy (predictive power) for the desired value.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}$$

Figure 5: RMSE formula [14]

Where,

$\hat{Y}_1, \hat{Y}_2, \hat{Y}_3, \ldots, \hat{Y}_n$ - *predicted values*

$Y_1, Y_2, Y_3, \ldots, Y_n$ - *observed values*

 DOI:10.36893.DRSR.2023.V13I04.118-125

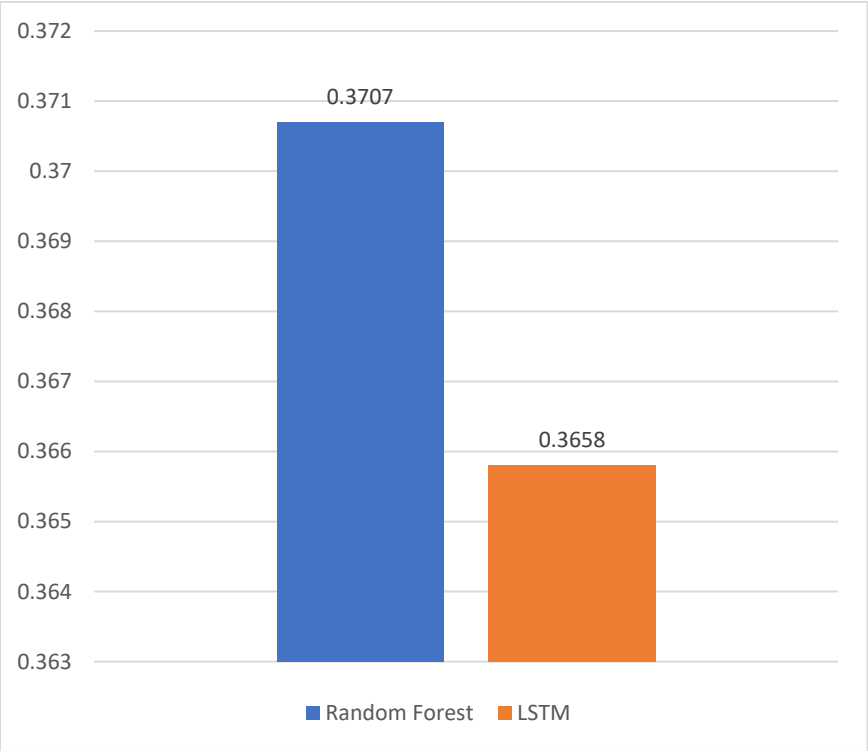
'n' *- number of observations*

These error rates based on RMSE (Root Mean Square Error) show the LSTM has a slightly higher accuracy than the Random Forest algorithm as shown in table (1) and graph (1) below.

Machine Learning Models	Error rate
Random Forest	**0.3707**
LSTM	**0.3658**

Table 1: Error rates

The graphical representation of error rates is in the below graph (1) :



Graph 1: Comparison Graph (Error Rates).

5. CONCLUSION AND FUTURE SCOPE

Machine learning algorithms have shown great promise in personalized diet recommendation systems. Motivating people to consume a healthy diet is the aim of nutrition education. Dietary interactions that are crucial for developing dietary guidelines are given special consideration. a health-based medical dataset that uses features like age, gender, weight, and height to automatically identify which foods

DOI:10.36893.DRSR.2023.V13I04.118-125

should be given to which patients based on their conditions. The machine learning and deep learning algorithms used in this study framework include Random Forest, K-Means, and Long Short-Term Memory (LSTM). The optimum eating strategy that yields positive effects are found in all of the models presented in this study. The development of personalized food advice systems could be furthered by investigating new important data sources and thinking about how to incorporate important variables like physical activity, lifestyle, and disease history.

REFERENCES

[1] Celestine Iwendi, Suleman Khan, Joseph Henry Anajemba, Ali Kashif Bashir, And Fazal Noor, "Realizing an Efficient IoMT-Assisted Patient Diet Recommendation System Through Machine Learning Model", IEEE Access, vol. 8,2020, pp. 28462-28474, doi: 10.1109/ACCESS.2020.296853.

[2]Thi Ngoc Trang Tran, Alexander Felfernig, Christoph Trattner and Andreas Holzinger, "Recommender systems in the healthcare domain: state-of-the-art and research issues", Journal of Intelligent Information Systems, vol. 57, 2020, pp. 171–201, doi: 10.1007/s10844-020- 00633-6.

[3] Jun Gao, Ninghao Liu, Mark Lawley, and Xia Hu, "An Interpretable Classification Framework for Information Extraction from Online Healthcare Forums", Journal of Healthcare Engineering, Cui Tao, vol. 2017, Article ID 2460174, https://doi.org/10.1155/2017/2460174.

[4] Butti Gouthami and Malige Gangappa, "NUTRITION DIET RECOMMENDATION SYSTEM USING USER'S INTEREST", International Journal of Advanced Research in Engineering and Technology (IJARET), vol. 11, 2020, pp. 2910-2919, doi: 10.34218/IJARET.11.12.2020.272

[5] Rachel Yera Toledo, Ahmad A Alzahrani And Luis Martinez, "A Food Recommender System Considering Nutritional Information And User Preference", IEEE Access, vol.7,2019, pp. 96695-96711, doi: 10.1109/ACCESS.2019.2929413.

[6]Inigo Orue-Saiz, Miguel Kazarez, And Amaia Mendez-Zorrilla, "Systematic Review Of Nutritional Recommendation Systems", Applied Science, vol. 11, no. 24: 12069, 2021, doi: https://doi.org/10.3390/app112412069**.**

[7] Shubham Singh Kardam, Pinky Yadav, Raj Thakkar, Prof Anand Ingle, "Website on Diet Recommendation Using Machine Learning", International Research Journal of Engineering and Technology (IRJET), vol.8, 2021, p-ISSN: 2395-0072.

[8] Salliah Shafi Bhat and Gufran Ahmad Ansari, "Predictions of Diabetes and Diet Recommendation System for Diabetic Patients using Machine Learning Techniques", International Conference for Emerging Technology (INCET), 2021, Belagavi, India, pp. 1-5, doi: 10.1109/INCET51464.2021.9456365.

[9] S. V. N. Sreenivasu, Sakshi Gupta, Ghanshyam Vatsa And Anurag Shrivastava, "Carbohydrate Recommendation for Type-1 Diabetics Patient Using Machine Learning", 5th International Conference on Contemporary Computing And Informatics, UP, 2022, pp. 600-604, doi: 10.1109/IC3I56241.2022.10072919.

[10] Aryan Jalali, Kumar P. Arjun Manoj, N.Amulya, Ayesha Siddiqui, Alok Kumar Singh, S.M. Hari Krishna, "Mobile Application for Personalized Food Recommendation", IEEE 7th International Conference for Convergence in Technology, Mumbai, 2022, pp. 1-6, doi: 10.1109/I2CT5491.2022.9824896.

[11]George A Bray, William E Heisel, Ashkan Afshin, Michael D Jensen, William H Dietz, Michael Long, Robert F Kushner, Stephen R Daniels, Thomas A Wadden, Adam G Tsai, Frank B Hu, John M

 DOI:10.36893.DRSR.2023.V13I04.118-125

Jackicic, Donna H Ryan, Bruce M Wolfe, Thomas H Inge, "The Science of Obesity Management: An Endocrine Society Scientific Statement", Endocrine Reviews, 2018, vol: 39, doi: https://doi.org/10.1210/er.2017-00253.

[12]https://gist.githubusercontent.com/sukanyabag/fec853cee5c77ec362820e2131777911/raw/8438f14bb7f48ba2f7dc8a7057e5bf23259a690a/LSTM.py

[13]https://www.nbshare.io/notebook/249468051/How-To-Code-RNN-and-LSTM-Neural-Networks-in-Python/

[14] https://towardsdatascience.com/what-does-rmse-really-mean-806b65f2e48e

[15]https://medium.com/hackernoon/understanding-architecture-of-lstm-cell-from-scratch-with-code-8da40f0b71f4**

[16]https://www.alpha-quantum.com/blog/long-short-term-memory-lstm-with-python/long-short-term-memory-lstm-with-python/

[17]https://bond-kirill-alexandrovich.medium.com/lstm-for-real-time-recommendation-systems-f5191d564be5

[18] https://towardsdatascience.com/building-a-random-forest-classifier-c73a4cae6781

[19]https://data.nal.usda.gov/dataset/usda-national-nutrient-database-standard-reference-legacy-release

 DOI:10.36893.DRSR.2023.V13I04.118-125