iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.43236

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Vitamin Deficiency and Food Recommendation System Using Machine Learning

Mrs. N. Naga Lakshmi[1], M. Jagadeesh Reddy[2], K. Hari Krishna[3], S. Sindhuja Reddy[4]
[1, 2]*Assistant Professor, Dept of IT, Anurag Group of Institutions, Telangana*
[1, 2, 3, 4]*Bachelor of Technology, Information Technology, Anurag Group of Institutions, Telangana, India*

Abstract: ***People are not paying attention to the quality of food they eat in our fast-paced and hectic world. They frequently ignore their eating routines and behaviours. Fast-food consumption is frighteningly increasing, which has resulted in the consumption of harmful foods. This causes a variety of health problems, including obesity, diabetes, and an increase in blood pressure, and so forth. As a result, it has become critical for people to have a well-balanced nutritionally sound diet. There are several applications that are thriving to assist folks in gaining control of their food and therefore can help individuals lose weight or maintain their fitness and health. The study article proposes healthy eating habits and patterns so that anybody may know the number of calories expended, macronutrient intake, and so on using data mining technologies. This technology is designed to uncover hidden patterns and client eating habits from various data sources. This approach will aid in tracking and improving an individual's health as well as the types of food that they should avoid in order to reduce their chance of disease. A balanced diet is one in which the intake of each basic nutrient meets its sufficient demand and real caloric intake equals calories burnt. Additionally, making a variety of dietary choices is vital for lowering the chance of acquiring chronic illnesses. This diet recommendation system tailors its recommendations to each individual depending on their eating patterns and body data. This study aids in the prediction of a healthy diet for any individual, as well as the construction of a diet plan based on the needs of the patient.***

Keywords : BMR, Healthy Diet, Recommender System, Harris Benedict equation, Nutrition, Calories, Data mining.

I. INTRODUCTION

A. Data Mining

Several Machine Learning (ML) applications have emerged, with data mining being the most effective. Data mining employs intelligent procedures, which play an important role in the process of information discovery in databases. At the junction of machine learning, statistics, and database systems, these intelligent technologies are used to extract hidden patterns. Data mining is the extraction of hidden knowledge from massive amounts of raw data. Data mining is the analytical stage in the "knowledge discovery in databases" process in KDD, which is an iterative process. Association rule learning, anomaly detection, classification, clustering, regression, and summarization are six major groups of activities that contribute to data mining. The availability of digital data nowadays primarily leads to the use of data mining to gather information and discover hidden patterns. Techniques from geographical data analysis, web technology, information retrieval, pattern recognition, signal processing, and computer graphics may be combined in data mining systems. Data mining systems are classified using the following criteria: database technology, statistics, machine learning, information science, visualization, and other fields.

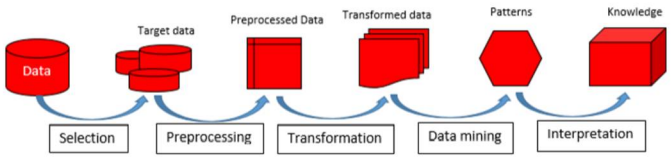


Figure 1:- The KDD Model

The goal of this study is to consider several significant components of the user's lifestyle and ensure that these parameters are integrated when the system develops a solution to create and suggest a healthy and nutritious diet for the user. A balanced, healthful diet and moderate physical activity can help you maintain a healthy weight. However, the advantages of healthy eating

©IJRASET: All Rights are Reserved

extend far beyond weight management. Good diet has a direct influence on our health and can help to minimize the risk of certain diseases such as heart disease, diabetes, stroke, certain cancers, and osteoporosis. When combined with physical exercise, food can help you achieve and maintain a healthy weight, lower your risk of chronic illnesses (such as heart disease and cancer), and boost general health. The 70/30 guideline applies to staying fit. To keep healthy, a person must focus 70% on his or her nutritional intake and 30% on physical activity/exercise.

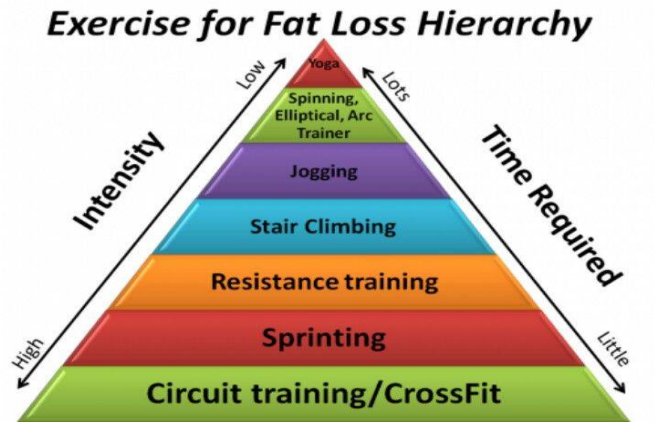


Figure 2: - Fat Loss Hierarchy

B. Data Mining Algorithms

The major purpose of the study paper is to collect data from patients (height, age, weight, physical activities, fitness goals, etc.) and evaluate the data in order to build a healthy diet plan based on individual goals. To develop and uncover any valuable and interesting suggested outputs, data sets are processed and data mining algorithms are used as information filtering tools. In the data mining process, we employ data mining methods such as classification, clustering, association rules, and the RandomTree decision tree algorithm, among others, to extract meaningful information about people's eating habits and to offer healthy diet plans. After comparing the J48, ID3, and RandomTree Decision tree algorithms, it was discovered that the RandomTree approach performed well for classification. First, the nutritional structure of each type of food is examined, and the amount of fat, energy, protein, and carbs needed in an individual's daily diet is estimated. The healthy diet report is prepared based on the quantity of calories and macronutrients required by each individual.

Classification is a data mining process that predicts categorical class labels. The ultimate aim of classification is to accurately anticipate the target class for each example in the data. The two phases of the Data Classification process are building the classifier or model and using the classifier for classification. Various algorithms are used in classification techniques. Classification is also known as Supervised Learning since a training dataset is utilized to build a classifier. \ The class assignments must be known ahead of time in order to begin the classification operation with a data set.

Clustering analysis groups together data elements that are related in some way. Beings in each group are more similar to entities in that group than to entities in other groupings. The purpose of clustering analysis is to locate high-quality clusters with low inter-cluster similarity and high intra-cluster similarity.

Clustering, like classification, is used to segment data. Clustering, on the other hand, is referred to as an unsupervised learning approach since references are chosen from unlabeled, unclassified, and uncategorized test datasets. Association rule learning is a rule-based machine learning approach for uncovering interesting relationships between variables in huge datasets. It consists of if/then phrases that aid in the analysis and prediction of client behavior. Association rule mining is used to identify strong rules uncovered in databases using various criteria of interestingness. As additional data is analyzed, this rule-based method creates new rules. It is essentially a pattern that asserts that when one event happens, another event happens with a particular probability. The ultimate purpose of Association rule mining is to assist a computer in mimicking the human brain's feature extraction and abstract association skills from new uncategorized data by assuming a huge dataset. A issue of association mining may be divided into two subproblems.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

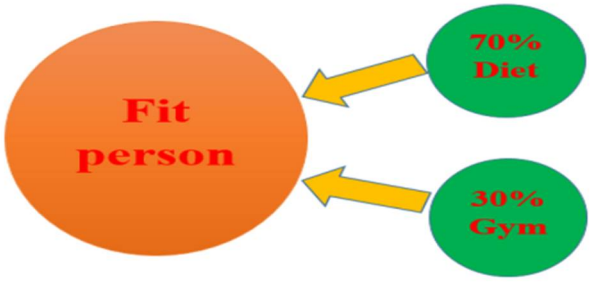


Figure 3: - The Fitness Mantra

II. LITERATURE REVIEW

The authors INGMAR WEBER and PALAKORN ACHANANUPARP [1] attempted to get insights from machine learning - diet success prediction, which would aid people attempting to stay fit and healthy by tracking their dietary consumption. The scientists studied the features of a failed diet using public food diaries from over 4,000 long-term active MyFitnessPal users. Authors specifically built a machine learning model to forecast going over or under self-set daily calorie targets frequently and then examined which factors contribute to the model's prediction, with research concentrating on "quantified self" data. The authors discovered that classification performance was adequate, and that the token-based model outperformed the category-based model, and that such data might be utilized for further in-depth data mining.

NANDISH SHAH and ISHANI SHAH [2] offered a suggestion for a healthy food habits and eating system based on online data mining, which would track eating habits and prescribe the forms of food that would promote health while avoiding the types of food that will increase the risk of disease. To extract meaningful information about people's eating habits, the authors employed data mining methods such as classification, clustering, and association rules. The nutritional structure of each type of food was examined, and the percentages of fat, energy, and vitamins in the recipe were computed. The classification mining method was then used to evaluate the composition data and determine if the diet was healthy or not. As a consequence, unique suggestions were made for each individual.

The authors AINE P. HEARTY AND MICHAEL J. GIBNEY [3] illustrated how a coding system at the meal level may be examined using data mining approaches in this work. They tested the usefulness of supervised data mining algorithms for predicting an element of nutritional quality based on dietary consumption using a food-based coding system and a novel meal-based coding system. Food consumption information from the North South Ireland Food Consumption Survey 1997–1999. A score for healthy eating (HEI) was created. Artificial neural networks (ANNs) and decision trees were used to predict HEI quintiles based on meal pairings. As a consequence, the ANN had somewhat greater accuracy than the decision tree in terms of predicting HEI.

However, on the basis of the meal coding scheme, the decision tree outperformed the ANN.

CHRISTY SAMUEL RAJU, SANCHIT V CHAVAN, KARAN PITHADIA, SHRADDHA SANKHE, and PROF. SACHIN GAVHANE [4] employed data mining to create a Fitness Advisor System. The authors' "Fitness Advisor" was a desktop program that recommended the user based on his/her problem with body weight through an efficient diagnostic and spreading correct information about the health concerns. The writers investigated several system parameters such as height, weight, body type, sex, smoking, drinking, health condition, physical activity, sleeping hours, and so on. The authors employed a combination of clustering, association, and classification algorithms to successfully give the best possible expert advice to the user's situation. The authors generated association rules using the Apriori method. The system's ultimate result was professional recommendations on nutrition and exercise. Lydia MANIKONDA, RAGHVENDRA MALLY, VIKRAM PUDIZ, and RAGHUNATHA RAO [5] employed mining questionnaires to measure the present knowledge of kid participants and how this knowledge increases after the training session. The work was broken into nine sections by the writers. Training was provided in the following areas: 1) Adolescent Phase 2) Breast Feeding 3) Food Types 4) Folates and Anaemia 5) Education in Family Life. The authors used data pre-processing to avoid inconsistencies. The primary algorithms employed for this goal were a few basic algorithms available in WEKA. On the dataset, three classification techniques were used: Naive Bayes, Bayes network, Decision Trees, and C4.5. The authors primarily employed two clustering techniques in their analysis: K-Means and DBSCAN. As a consequence, it was discovered that data mining was capable of inferring the outcomes far better than standard statistical analysis.

SHILPA DHARKAR and ANAND RAJAVAT [6] completed a performance study of a Healthy diet recommendation system utilizing online data mining, where the Design and implementation of a Healthy diet recommendation system was based on web

data mining. The authors discovered that the Healthy Diet Recommendation System included data gathering, data pre-processing, information filtering, data base design and implementation with a web-based user interface, and interaction between the user and the recommender system over HTTP. To categorize the healthy diet data set, the performance of a healthy diet recommendation system employed the ID3 and C4.5 decision tree classification algorithms. The authors chose the data set first, followed by the produced rule, during the implementation step. The authors proposed a novel framework based on data mining technologies for developing a Web-page recommender system, which would be utilized as the core framework for the healthy eating system. The authors evaluated two information filtering strategies for giving the desired information. The information filtering operations could be conducted before to the actual suggesting process by employing data mining techniques, which improved the system response time and made the framework scalable.

III. PROPOSED METHODOLOGY

A. System Methodology

The suggested meal recommendation system for a certain client is based on elements such as food preferences, food availability, medical information, illness information, caloric information for a food item, personal information, and each individual's activity level for a given food database. The most significant task in implementation is to propose a certain food item from the food database based on certain restrictions such as the likelihood of that food item, its availability, allergy to that food, and its nutritional components such as protein, carbs, and fats in that food. This tip aids in the selection of foods from the database so that nutritional deficiencies do not emerge in the near future and an appropriate diet plan is provided to each individual while meeting the daily calorie intake.

The main goal of the presented work is to build the decision tree until the appropriate classification is reached in order to select the proper food item based on food availability, user category (Fat, healthy, lean, etc.), likeness factor, user fitness goals, overall content of Nutrients in that food, decision rules and constraints on it are defined in order to design a healthy diet plan for each individual.

B. System Architecture

The overall process of developing a Healthy Diet Recommender System consists of seven steps: collecting user data, designing and implementing a database, acquiring data, pre-processing the data to filter and transform the food datasets based on user profiles, performing classification and clustering, applying the rules on the algorithm based on each individual profile, and recommending a healthy diet plan.

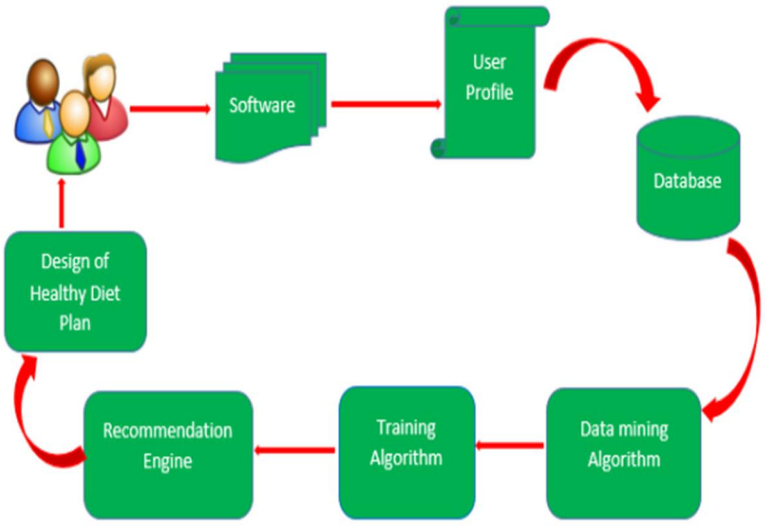


Figure 4: - Diet Recommendation system architecture

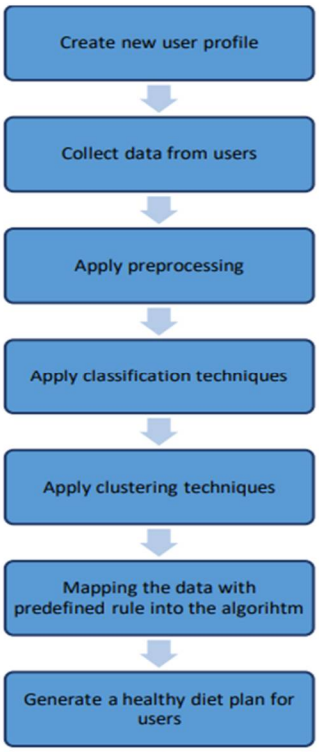


Figure 5: - Process flow of methodology

1) *Data Acquisition and Collection:* People log onto the system and immediately enter their health-related information and work lifestyle, which is then saved in the database. This data is collected on a regular basis in order to track people's health problems. This software-acquired health and lifestyle-related information is the first-hand material for this system, such as the users' height, weight, illnesses, allergic conditions, physical activity, habits, and so on. Data on food products and their nutritional worth are gathered from the USDA agricultural research service food databases. In general, the data collection module gets data from the outside online environment selectively in order to supply material and resources for the latter data mining.
2) *Data Pre-processing:* Data pre-processing primarily analyses and reconstructs the source data received during the data acquisition phase, as well as building a data warehouse of related topics to provide a foundation for the data mining process. It is a crucial data mining approach for converting raw data (inconsistent, incomplete, etc.) into a comprehensible format. Data pre-processing is the preparation for data mining, and it mostly tackles raw data difficulties through methods such as data cleansing, data integration, data conversion, data reduction, and so on.
3) *Information Filtering:* The information filtering step is the core process of the recommender system framework for removing unwanted or redundant data, in which data sets are analyzed and data mining algorithms are used as information filtering tools using (semi)automated or computerized methods to generate and discover any useful and interesting recommended outputs. In the data mining process, we utilize data mining methods such as classification techniques, clustering, and so on to extract meaningful information about people's eating patterns and anticipate the food items to be advised based on these behaviours and their health state. First, we examine the nutritional profile of each type of food and determine how much fat, energy, protein, and carbs that human needs in their daily diet. The preset criteria are then applied to the algorithms to process the composition data and provide the healthy diet reports.
4) *Recommending Healthy Diet:* We could acquire a lot of useful information after the information mining process in the last stage. Then we may recommend a healthy diet tailored to the consumers' personal circumstances. Our recommender system's advice would boost your organic process structure and raise your health standards. On the other side, we tend to concurrently track consumer preferences. This strategy may promote the connected diet try to meet the individual demands via the use of association rule mining. As a result, users will receive better service and expertise.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

C. Implementation

Decision tree learning is a data mining decision assistance tool that employs a tree-like model for decision making and computing the goal value with a separate function. It employs a branching strategy to present every conceivable choice outcome. The learnt function is represented by the decision tree. It is a straightforward format for categorizing examples. The decision tree is a popular non-parametric effective machine learning approach for creating tree-like classification or regression models.

The decision tree learning algorithms are commonly utilized for the following three reasons:

1) The decision tree may be widely utilized to derive conclusions from unseen examples of specific scenarios.

2) In these approaches, efficient computations that are proportionate to the observed occurrences are done.

3) Finally, when all computations have been completed, the created decision tree is easily comprehended by humans.

The suggested system uses the RandomTree algorithm for decision tree learning to make judgments such as which appropriate food item should be assigned while preparing the menu. The training dataset is supplied to categorize the decision tree in order to make this judgement. Entropy and information gain factors are calculated using training data. A condition might result in both negative and good results. The suggested system employs a decision tree to assess if a specific food item should be offered to an individual or not, considering parameters such as the Category of user fitness objectives, the Likeness Factor, and whether or not the individual is allergic to food products. Based on the favorable or bad consequence, one may efficiently recommend meals to be given to an individual. To make appropriate choices among accessible meals, the RandomTree algorithm is employed. When creating an appropriate healthy diet plan, user preferences might be considered. A larger training set produces more accurate results.

As the number of potential examples in characteristics grows, so does the training set. In this situation, the training data set is interpreted as all of the possible cases for each user, ensuring that accurate results are generated each time. To train RandomTree, a suitable collection of characteristics and an output choice must be given.

D. Proposed Mathematical Constraints and Algorithms

The Harris-Benedict equation, also known as the Harris-Benedict principle, is used to calculate a person's basal metabolic rate (BMR). The Harris Benedict Equation is used to calculate total daily energy expenditure (calories). Calories are consumed even when a person is inactive or not performing any work; the Basal Metabolic Rate (BMR) measures how many calories a person would burn if they slept all day. The BMR formula uses variables such as height, weight, age, and gender to compute the Basal Metabolic Rate (BMR). The Harris Benedict Equation is then used to calculate an individual's total daily energy expenditure (calories). To calculate the Harris Benedict Equation, an activity factor is used to BMR. We use basic mathematical calculations to calculate the quantity of calories to consume at each meal of the day.

Consider calorie intake for a certain individual as 'CAL' dependent on activity level and BMR.

These mathematical formulae assist in allocating the quantity of calories at various times of the day, such as morning breakfast, snacks, lunch, evening snacks, supper, and so on.

Equations are defined as,

Morning Breakfast: $a1 + a2 + a3 + ... + an1 = CAL/5$

Snacks: $b1 + b2 + b3 + ... + bn2 = CAL / 10$

Lunch: $c1 + c2 + c3 + ... + cn3 = 4*CAL / 10$

Evening snacks: $d1 + d2 + d3 + ... + dn4 = CAL/ 20$

Dinner: $e1 + e2 + e3 + ... + e5 = CAL / 4$

Where n1, n2, n3, n4, n5 represent the number of food items consumed at various times, such as morning drinks, breakfast, lunch, evening snacks, supper, and so on. ai, bi, ci, di, ei signifies the calories consumed from various food items "i" in the morning, snacks, lunch, evening snacks, dinner, and so on.

An algorithm is meant to decide which food products must be included in the diet plan in order to meet the nutritional requirements of the user. Each macronutrient that must be included in the diet plan has an upper and lower bound set by default. If adding paneer causes fat to surpass its upper bound, that food item must be eliminated and replaced by another food item that meets the requirements.

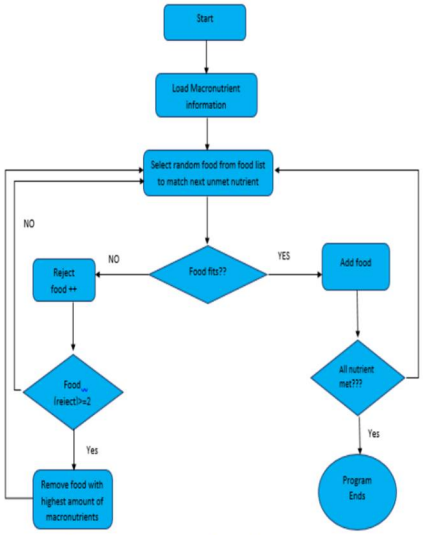


Figure 6: - Graphical representation of algorithm used to generate a food list

E. Diet Recommendation System Challenges

The three key challenges are as follows:

1) Problems encountered in relation to the user.

2) Algorithm-related difficulties encountered

3) Difficulties in describing the link between various types of dietary components.

IV. OUTCOMES AND FUTURE POTENTIAL

The RandomTree method is used to determine whether a specific food item should be supplied to an individual or not. To utilize RandomTree, you must first train it using the training dataset presented in Table 1. The decision tree is constructed after using the RandomTree method on the training data, as illustrated in Figure 8. The decision tree indicates whether or not a certain food item should be included in the client's diet plan. If a person is allergic to a food item listed in the system, the food is immediately deleted from his diet plan; otherwise, the system evaluates the liking factor of clients towards that food item. If a customer has a high liking factor for a certain food item, that meal will be included in his diet plan regardless of his fitness objective. If the likeness factor for a food item displayed is medium, the decision on whether that food item should be included in the diet plan or not will be made by taking the client's fitness goals into account; otherwise, if the likeness factor is low, the food item will not be added to the client's healthy diet plan. After using the classification approach to identify which food products are acceptable for an individual, the specified criteria are applied to the algorithm, and a healthy diet plan based on an individual's preferences is developed. In the future, an algorithm can be built to recommend a meal plan based on advanced nutrition levels such as salt content, phosphorous content, fibre content, manganese content, and so on. Along with the food products recommended for each meal, the system may be programmed to develop and deliver recipes that contain all of the food items recommended in the meal plan. Users can be given more freedom to add their own food items of their choosing, as well as the ability to incorporate cheat meals into the diet plan.

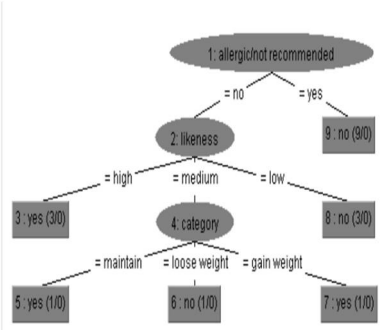


Figure 7: - Generated RandomTree decision tree

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

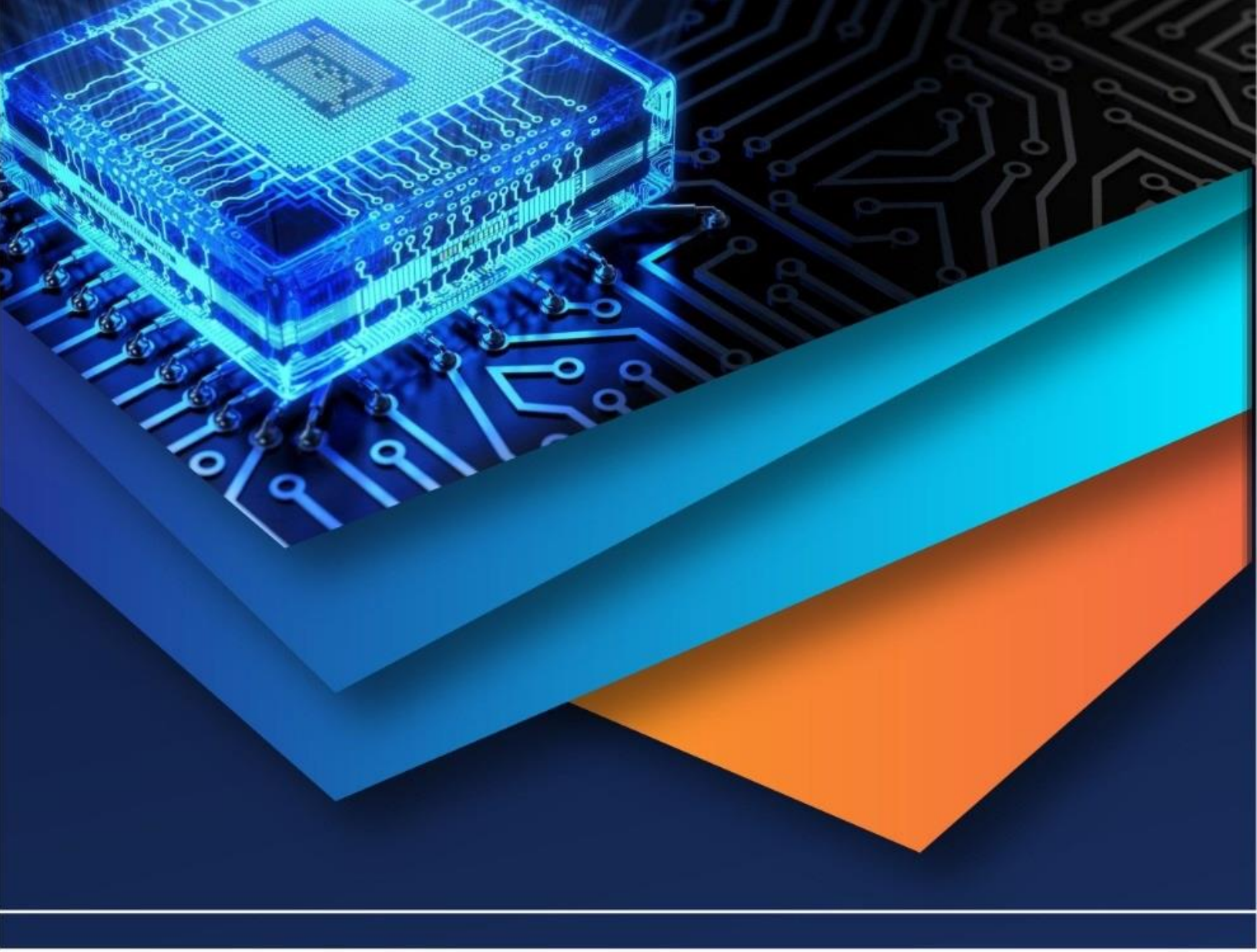
V. CONCLUSION

The relevance of dietary counselling in leading a healthy and fit life is growing by the day. A healthy diet plan may be developed by accepting the user's food choices and a user's profile in the system. Balancing the diet and developing a healthy diet plan by measuring calorie requirements based on an individual's preferences is often a time-consuming and labor-intensive procedure.
The research and their implementation shown that the decision tree learning method, RandomTree, performs effectively on any classification issue with non-repeated values in the dataset.

REFERENCES

[1] Akshay Mahajan, Ankita Dharmale, Ayushi Agarwal, Shriya Pawar, and Sneha Sunchu, Computer Engineering students at PCCOE in Pune, Maharashtra, India, "NUTRIEXPERT: A HEALTHY DIET RECOMMENDATION SYSTEM," IJARIIE-ISSN(O)-2395-4396, Vol-3 Issue-2 2017.
[2] Shilpa Dharkar and Anand Rajavat, "WEB DATA MINING FOR DESIGNING OF HEALTHY EATING SYSTEM", Department of Computer Science & Engineering, RGPV University Bhopal, India, International Journal of Internet Computing, Volume-I, Issue-1, 2011.
[3] Ingmar Weber and Palakorn Achananuparp, Qatar Computing Research Institute Doha, Qatar, Singapore Management University Singapore, "INSIGHTS FROM MACHINE-LEARNED DIET SUCCESS PREDICTION," Pacific Symposium on Biocomputing 2016, pp. 540-551.
[4] Nandish Shah and Ishani Shah, Department of IT Engineering, Institute of Technology, Nirma University, Ahmedabad (India), "DIET RECOMMENDATION SYSTEM USING WEB DATA MINING," International Journal of Advanced Technology in Engineering and Science, Volume No.02, Special Issue No. 01, September 2014, ISSN (online): 2348 – 7550, pages 334–341.
[5] American Society for Nutrition, "Analysis of meal patterns using supervised data mining techniques—artificial neural networks and decision trees," Am J Clin Nutr 2008;88:1632–42. Printed in the United States. American Society for Nutrition, 2008.
[6] Lydia Manikonda, Raghvendra Mally, Vikram Pudiz, and Raghunatha Rao Data Engineering Center Hyderabad, India "Mining Nutrition Survey Data," National Institute of Nutrition, Hyderabad, India.
[7] Christy Samuel Raju, Sanchit V Chavan, Karan Pithadia, Shraddha Sankhe, Prof. Sachin Gavhane, Information Technology Department, Atharva College of Engineering, Mumbai, "Fitness Advisor System Using Data Mining," International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 4, April 2016, DOI 10.17148/IJARCCE.2016.5451, pages 201–204.
[8] Shilpa Dharkar and Anand Rajavat International Journal of Scientific and Engineering Research, Volume 3, Issue 5, May 2012, ISSN 2229-5518, "Performance Analysis of Healthy Diet Recommendation System Using Web Data Mining."

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)