International journal of application or innovation in engineering and management

Ijaiem.com/June 2025/ Volume 14/Issue 1/Article No-1/893-901

ISSN: 2319-4847

An Online Health Tracking System that Uses Machine Learning to Make Nutrition and Exercise Suggestions

[1] Dr.B. Nageshwarrao, [2] P. Radhika,

[1]Assistant Professor, Megha Institute of Engineering & Technology for Women, Ghatkesar.
[2] MCA Student, Megha Institute of Engineering & Technology for Women, Ghatkesar.

Article Info

Received: 30-04-2025 ***Revised: 06 -06-2025*** ***Accepted: 17-06-2025*** ***Published:28/06/2025***

Abstract

A growing number of people are being diagnosed with diseases that are progressing to chronic stages because they are either too busy to adhere to healthy eating habits, exercise consistently enough, or pay enough attention to their illnesses while they are sick. We thus suggest a system that analyzes and monitors health metrics and the values from patients' most recent reports relating to the condition in order to suggest better eating and exercise regimens to patients suffering from different ailments. Patients with diabetes, hypertension, or thyroid issues were taken into consideration. Based on patients' most recent lab results and other health information, our system may help physicians make dietary and activity recommendations. We have divided our system into two main sections for this purpose: First, keeping tabs on your health; second, suggesting healthy eating and exercise habits. Until the reports return to normal, the system will recommend follow-up sessions in the Health Monitoring module. Decision trees are used as categorization algorithms in the Diet and Exercise Recommendation module. In particular, dietary and exercise suggestions are made using C4.5. With the use of our specialized datasets, a C4.5 decision tree will aid in making recommendations and deciding whether or not to provide a certain food item and exercise to a certain person.

Keywords

Machine learning, C4.5, health monitoring, dietary and exercise recommendation

I. INTRODUCTION

A person's health is also crucial to their survival. Health and fitness are taking a back seat as a result of people's hectic schedules and heavy workloads. The biggest issue facing today's youth is their lack of physical exercise. The key to staying fit is sticking to a regular exercise and eating regimen. Therefore, maintaining and improving one's health requires a certain level of nourishment. Users' dietary and exercise habits fluctuate across demographics such as age, gender, height, weight, and degree of physical activity. Physical activity and healthy eating go hand in hand. Keeping calorie consumption in check is essential for sugar maintenance. Therefore, the suggested method would enable physicians to prescribe medicine and dietary supplements to patients with diabetes, hypertension, or thyroid issues with the simple click of a mouse at each follow-up appointment. Our concept for a health monitoring

International journal of application or innovation in engineering and management

Ijaiem.com/June 2025/ Volume 14/Issue 1/Article No-1/893-901

ISSN: 2319-4847

system that includes dietary and exercise recommendations is detailed in this article. Diabetes, hypertension, and thyroid disorders are the only ones we focus on in this approach. Proper health monitoring and treatment are necessary due to the widespread prevalence of these illnesses among individuals. Based on the user's needs and limitations, the recommendation system will provide them information. Two separate components make up our system. Medical Tracking Device No. 1 Second, suggestions for food and physical activity. The C4.5 classifier is used by the Diet and Exercise Recommendation module. When compared to a standard decision tree classifier, it performs better because to its enhanced capabilities, such as pre pruning, managing continuous attributes and missing data, and rule induction. The recommendation system's optimum algorithm is found via several comparisons. ID3 and C4.5 both work well in terms of algorithmic qualities, and C4.5 meets all the optimal criteria in terms of the following characteristics. In TABLE I, we can see that ID3 and C4.5 vary in terms of the kind of data handled, speed, splitting criteria, correctness, and whether or not pruning occurs.

TABLE I DIFFERENCE BETWEEN ID3 AND C4.5

	ID3	C4.5
Data type handled	Categorical	Categorical, Continuous, Numerical, Missing
Speed	low	faster than ID3
Pruning	Not Supported	Pre-pruning
Splitting criteria used	Information Gain	Information Gain Ratio
Accuracy	low	higher than ID3

II. LITERATURE REVIEW

A. In regards to the Fitness Tracking System Using an ontology framework for dietary and physical activity recommendation, a health care recommendation system was constructed in [1]. The data-set was queried for user information using a decision tree method. With a score and accuracy of 60 to 70% for health monitoring that examined the electrocardiograms (ECGs) of patients with the LQTS genetic disorder and identified patients at high risk of cardiac events, Random forest outperformed the other three algorithms (k-nearest neighbors, Support vector machine, and AdaBoost) in the study [2]. The logistic algorithm outperformed the other two algorithms—random forest and gradient boosting—in terms of accuracy and score (87 percent) when it came to monitoring and authenticating a user's Fitbit credential ([3]). Various ML models, including Random Forest, Support Vector Machine, and Deep Learning, were used to analyze data for the purpose of remote health monitoring for the elderly in [6]. In [7], the system automatically classifies the various phases of CKD according to their severity using ML after extracting variables from the UCI Chronic kidney data set that are responsible for chronic kidney disease. B. In regards to the System for Recommendations on Diet and Exercise By using the USDA Food Composition Database and the WEKA categorization methodology, the authors of [8] suggest a Mauritian diet for hypertension individuals.A web-based approach that uses the Health Calabria Food Database to provide dietary recommendations has the potential to enhance the health of those impacted by chronic diseases, according to [11].

According to [12], a recommendation system was developed for both amateur and professional runners. The system uses social semantic web inputs to present users with food and training advice that are tailored to their needs. Section C: Algorithms A refined and easier variant of the ID3 method was presented in [4]. They find that the accuracy rises in tandem with the quantity of data recordings. Compared to the standard ID3 Algorithm, which

Ijaiem.com/June 2025/ Volume 14/Issue 1/Article No-1/893-901

ISSN: 2319-4847

achieved an accuracy of 88.9% with 1232 records, their enhanced algorithm achieved an accuracy of 92.6%. [5]Three distinct methods were employed: the naive Bayes algorithm, k-means, and the ID3 decision tree algorithm. When compared to the other two algorithms, ID3's data categorization accuracy was found to be 6 to 7% higher. After obtaining class labels using the Kmeans clustering approach, the J48 Decision Tree approach was used on the ARFF Dataset to construct a classifier model for the prediction of learner groups on the Test Data in [9]. There were 11,200 celebrity deaths recorded worldwide between 2006 and 2016, and the authors of [10] compiled this information from public and open access sources. In order to analyze the provided dataset, decision tree classifier models and lazy ones were used. With the generated dataset, the decision tree model attained an accuracy of 75.07%. Using the Soil Nutrient Data Collection, the authors of [13] developed a model to predict soil quality and integrated soil composition using the C4.5 decision tree method; the model achieved an accuracy of 92.71%. The Health Monitoring system and the Diet & Exercise Recommendation system were both solved independently by the aforementioned articles. Based on the research, it seems that just one study, [10], generated its own data collection, while the others relied on pre-existing datasets. None of them were condition-specific, unlike diabetes, thyroid, and blood pressure all at once. For the reason that C4.5 is a model for making predictions and not suggestions [13]. Using C4.5 for dietary and exercise recommendation, our system fills in all the aforementioned research gaps to create a unified health monitoring system that also includes personalized separate data sets for patients with diabetes, hypertension, and thyroid issues.

III. C4.5 DECISION TREE ALGORITHM

A. Algorithm:

TABLE II C4.5 DECISION TREE ALGORITHM

Steps	Steps to generate C4.5 Decision Tree
1	Finding Global Entropy - probability of occurrence of a value in the output $Entropy(output) = \sum -p(I).\log_2 p(I)$ $where,\ p(I) = \lvert Xj/X \rvert$ where, Xj is the number of occurrence of true value in that particular attribute and X is the total no of value in that attribute
2	Again we will use the same formula to find entropy of each attribute with respect to output in step 1.
3	Now finding the information gain of every attribute with every possible values in the attribute with respect to output. **InfoGain (attribute)=** $Entropy(output) - \sum (p(output \mid attribute).Entropy(output \mid attribute))$
4	**SplittingInfogain(attribute)=** $-\sum p(I) * \log_2 p(I)$
5	**InfoGainRatio (attribute)=** $Infottain(attribute) / SplittingInfogain(attribute)$
6	To find out the attribute that have highest information gain ratio
7	The root of the decision tree is the attribute that have highest information gain ratio.
8	Splitting the dataset based on the newly created root in step 7.
9	For all the sub-dataset in step 7, calling function C4.5 recursively (i.e .all steps from step 1 to 8).
10	Attaching the tree found in step 9.i.e. the subtree to the root found in step 7.
11	Return Tree.

This algorithm Machine learning makes use of TABLE II for classification, which is a tree structure with an internal node representing a feature or attribute's value and a leaf node representing a decision based on that feature or attribute. Data categorization in C4.5 is based on entropy. Following on from the ID3, we have the C4.5. [16] According to TABLE I, C4.5 is quicker and more accurate than ID3 because it uses the information gain ratio as its splitting criterion, while ID3 uses the information gain as its criterion. We may learn about the amount of information attached to an attribute value and the likelihood of its occurrence in relation to the output by looking at the Information Gain. The entropy, or likelihood of an occurrence, is computed once the qualities have been chosen. One way to quantify the degree of surprise, unpredictability, or uncertainty in a dataset is by calculating its entropy. Data with both

International journal of application or innovation in engineering and management

Ijaiem.com/June 2025/ Volume 14/Issue 1/Article No-1/893-901

ISSN: 2319-4847

numerical and category values may be processed using this technique. Missing values may also be handled by it. Additionally, C4.5 does not use these missing variables when calculating gains. C4.5 algorithm fixes the issues of ID3 method, such as overfitting, temporal delay, and difficult to handle continuous characteristics. [16] See Table I.

IV. SYSTEM DESIGN

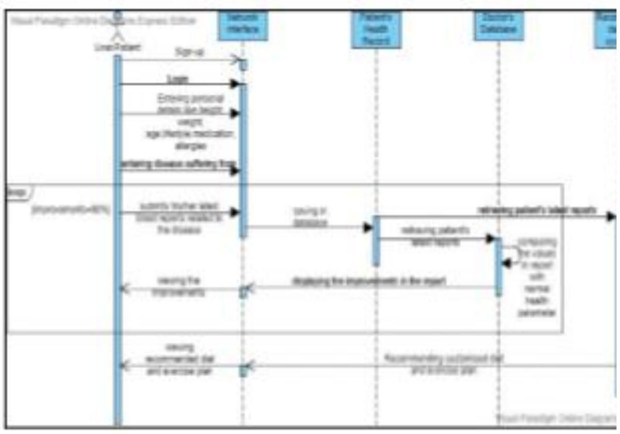

Fig. 1. Sequence Diagram

In Figure 1 you can see our model's system design/sequence. 1) The user or actor who will interact with this system is a patient or doctor. 2) The user interface allows users to input personal information such as height, weight, age, gender, lifestyle, exercise level, as well as medical conditions (such as diabetes, hypertension, or thyroid) and recent reports pertaining to such conditions. Thirdly, the patient's medical history will include all of the information from POINT 2. Doctor's Database - This database stores the fixed parameter used to identify illness categories, such as whether a person is normal, has prediabetes, or diabetes, based on the reports of that particular condition. To be saved in the database, the following parameters will be used. a) The predetermined criteria for identifying illness categories, which are utilized for tracking purposes and are checked by medical professionals and reviewed on various websites. references [20, 21, 22]:

i) Diabetes

Blood Sugar Classification	Blood Sugar level		
	HB A1C (%)	Fasting (mg dl)	2 hours Post-prandial (mg dl)
Normal	<5.7	70-99	100-139
Prediabetes	5.7-6.4	100-125	140-199
Diabetes	>=6.5	>125	>200

Fig. 2. Diabetes Parameters used

ii) Blood Pressure

Blood Pressure Category	SYSTOLIC (mmHg)		DIASTOLIC(mmHg)
Low Blood Pressure - Hypotension	<80	or	<60
Normal	80-120	and	60-80
Prehypertension	121-139	or	81-89
Stage 1 high blood pressure Hypertension	140-159	or	90-99
Stage 1 high blood pressure Hypertension	160-179	or	100-109
Hypertensive crisis (where emergency care is required)	180=<	or	110=<

Fig. 3. Blood pressure Parameters used

iii) Thyroid

Gender	Age	Normal	Hypothyroidism	Hyperthyroidism
Male	18-30	0.45-4.15	<0.45	>4.15
Male	31-50	0.45-4.15	<0.45	>4.15
Male	51-70	0.45-4.5	<0.45	>4.5
Male	71-90	0.4-4	<0.4	>4
Female	18-30	0.4-4.68	<0.4	>4.68
Female	31-50	0.4-4	<0.4	>4
Female	51-80	0.46-2.38	<0.46	>2.38

Fig. 4. Thyroid Parameters used

FIGURE 2,3,4 are the disease parameters store in Doctor's Database POINT 4.

Improvements in reports	Threshold
Yes	>90%
Less	65-89%
No	<65%

International Journal of application or innovation in engineering and management

Ijaiem.com/June 2025/ Volume 14/Issue 1/Article No-1/893-901

ISSN: 2319-4847

Fig. 5. Threshold Parameters used

b) In FIGURE 5, we can see the anticipated threshold values for comparing health monitoring improvements. 5) Recommendation System—Using a machine learning approach, this system will assist physicians in recommending exercise and food regimens for patients.
A fitness trainer, a doctor, and numerous websites all contribute to the preparation of personalized exercise and diet data sets. the references [24] [25] (26, 27, 28) [29] a) Exercise data-set [20]—Comprising 107 exercises, it has 1261 records altogether and inquires about the user's age, activity level, and preference for gym or yoga. b) Diet data-set [20]—102 records altogether, including calorie intakes and portions for 102 different foods, broken down by whether the user is a vegetarian or not.

6) BMI (Body Mass Index) [18] and Calories Requirement Calculation [17]

$$BMI = \frac{\text{body weight in kg}}{\text{Square of body height in m}} = kg/m^2$$

where, Underweight ≤ 18.5
Normal weight = 18.5 – 24.9
Overweight = 25 – 29.9
Obesity ≥ 30

$$Calories =$$

$$\textbf{For Men}: 66.5 + 13.8(W) + 5.0(H) - 6.8(A))$$

$$\textbf{For Women}: 65.51 + 9.6(W) + 1.9(H) - 4.7(A$$

where,
W = Weight in lbs.
H = Height in inches.
A = Age in years

V. METHODOLOGY ADAPTED

Here is the block schematic of the whole system, as seen in FIGURE 6. The user's profile is the primary factor in a customized healthcare recommendation system's dietary and activity recommendations. For various disorders, the dietary item changes. However, we aim for user-friendliness and adaptability in our diet and exercise recommendations. First, there is the health monitoring system, which keeps tabs on the patient's most recent data on their ailment (such as diabetes, hypertension, or thyroid) and records any improvements. a) First, the patient's demographic information (age, gender, height, weight, etc.), followed by his medical history (diabetes, hypertension, thyroid, etc.). b) Next, the patient's measured values for the condition as it appears in his most recent report.

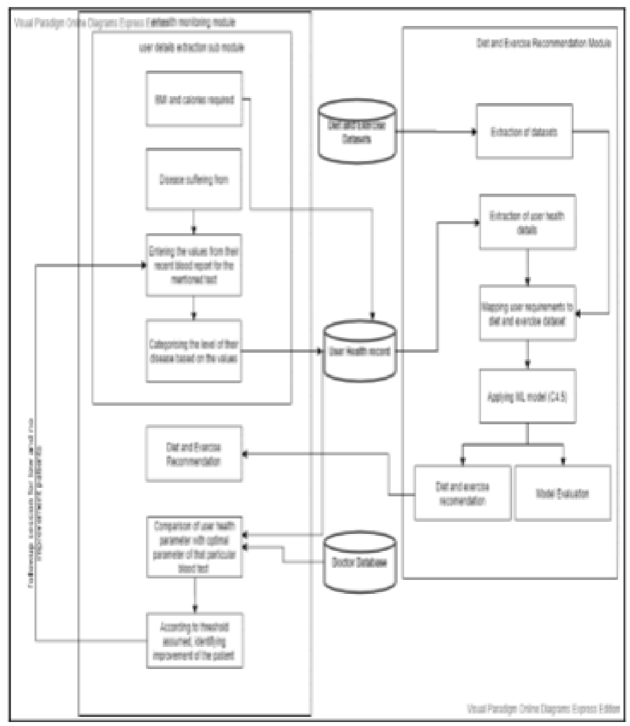

Fig. 6. System Block Diagram

c) Utilizing the formula from the previous part, we can calculate his body mass index (BMI) and the number of calories he needs by entering his personal information at the back end. d) The User Health Record keeps all of the user's information that was indicated in POINT a. e) In order to classify their ailment, it is compared to predetermined parameters of that particular illness that are maintained in the Doctor's Database. f) The Diet & Exercise Recommendation Module has suggested a personalized exercise and diet plan to the user. g) The process continues by evaluating the reports for any improvements (i.e., how similar they are to the

International journal of application or innovation in engineering and management

Ijaiem.com/June 2025/ Volume 14/Issue 1/Article No-1/893-901

ISSN: 2319-4847

normal parameters of the disease). If the reports show an improvement of 90%, the system suggests continuing with the same diet and exercise plan and the process ends. If the reports show an improvement of 65-89% (LESS) or below 65% (NO), the patient is advised to come back at a specific interval with updated reports about their disease. h) In order to proceed with the procedure from point (a), the following steps will be taken during the follow-up session: comparing the values in the reports with the fixed parameters of the disease that are stored in the doctor's database, but this time, we will also compare the reports from previous sessions with the current report. i) The patient's report of progress is considered complete when the number of follow-up sessions reaches 90%. 2) Recommendations for Diet and Exercise—Generating better eating plans and exercise regimens using a Machine Learning algorithm that takes into account the monitoring data mentioned in POINT 1 and other disease-related health metrics. a) First, have a look at the User Health Record and extract and map all the relevant information. Then, import the custom diet and exercise data set from the Diet and Exercise Database [20]. b) With a 30% training set and a 25% test set, the dataset is divided into two halves. c) To quickly determine the most relevant variables and the relationships between them, the C4.5 decision tree based classifier is used. We can improve our ability to anticipate target variables by developing new features and variables. d) Decision tree generation—choosing the best characteristics from all of them using characteristics Selection Measures (ASM) such information gain ratio and splitting the dataset into smaller subsets—occurs during training. e) Each child node in the decision tree will carry on the operation of generating the tree. f) Diet and exercise regimen suggestions are provided as a result of this approach. g) For Model Evaluation, we merge the training and testing results. h) After that, the procedure is ended after checking the model's correctness using many Performance Evaluation Measures such as the Confusion Matrix, Precision, and Recall. We evaluate C4.5 and ID3 based on our model's accuracy to see which one works better.

VI. RESULTS

Fig. 7. Screenshot of User Record suffering from hyperthyroidism after putting all the details

Fig. 8. Screenshot of Diet and exercise recommended to the patient whose details are mentioned in FIGURE 7

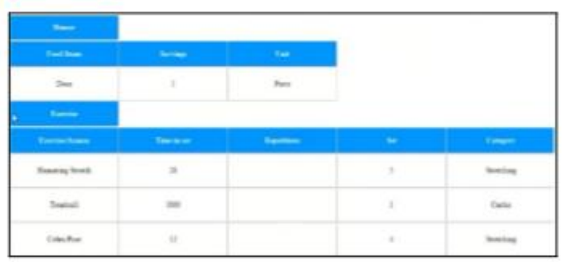

Fig. 9. Screenshot of Diet and exercise recommended to the patient whose details are mentioned in FIGURE 7

The user record entered for the female patient with thyroid disease is shown in Figure 7. Her body mass index (BMI) and caloric consumption have been computed in the background using her age, weight, height, and activity level. Her TSH level was 0.2, which puts her condition into the hyperthyroidism category shown in Figure 4. Since this number is

International journal of application or innovation in engineering and management

Ijaiem.com/June 2025/ Volume 14/Issue 1/Article No-1/893-901

ISSN: 2319-4847

considered abnormal, a "Assess again" button will display, indicating that the patient needs to return for a follow-up appointment to get further treatment. Their diet and exercise routine will be adjusted based on her updated reports in the follow-up. The hyperthyroidism treatment regimen shown in Figure 7 is shown in Figures 8, 9.

Fig. 10. Screenshot of followup session of the user suffering from hyperthy- roidism mentioned in FIGURE 7

Fig. 11. Screenshot of Diet and exercise recommended to the patient whose details are mentioned in FIGURE 10

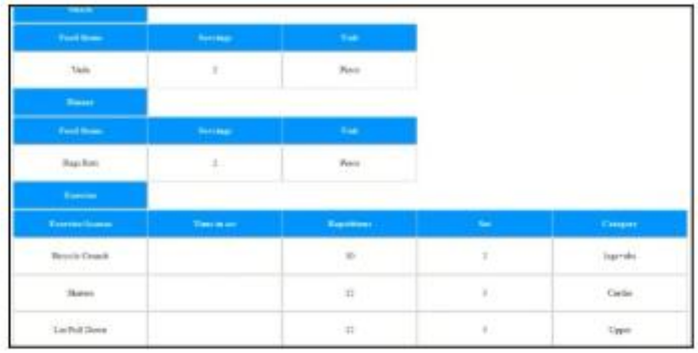

Fig. 12. Screenshot of Diet and exercise recommended to the patient whose details are mentioned in FIGURE 10

Figure 10 displays the patient's value during the follow-up session seen in Figure 7. As seen in Figure 4, the patient's fresh TSH result showed normal levels of 0.42. The 'Assess again' button vanishes when the category is determined to be normal; this means that the patient does not need a follow-up session and is considered normal.In the follow-up session stated in Figure 10, the user's modified food and activity plan is shown in FIGURE 11, 12.

VII. PERFORMANCE ANALYSIS

We compared our model using the different types of diseases. against determine how effective the C4.5 method is, we also compared our model, which used C4.5, against ID3. The Diet Data-set [20] has 102 data points about illnesses, while the Exercise Data-set [20] has 1261 data points about diseases, as shown in FIGURES 13, 14. From the graphs in Figures 13 and 14, it is evident that C4.5 outperforms ID3 in terms of accuracy, and this holds true for both of the customized datasets we used.

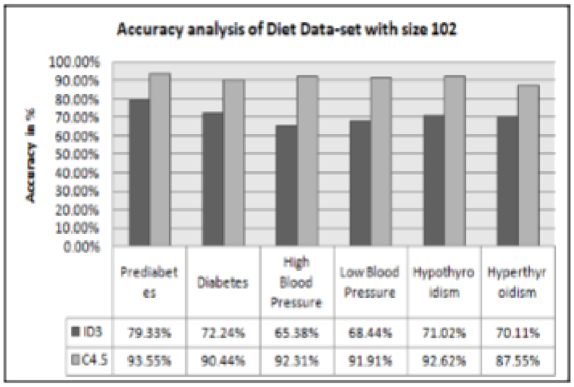

Fig. 13. Comparison of Accuracy for ID3 and C4.5 Decision tree Algorithm with Diet Data-set with 102 size

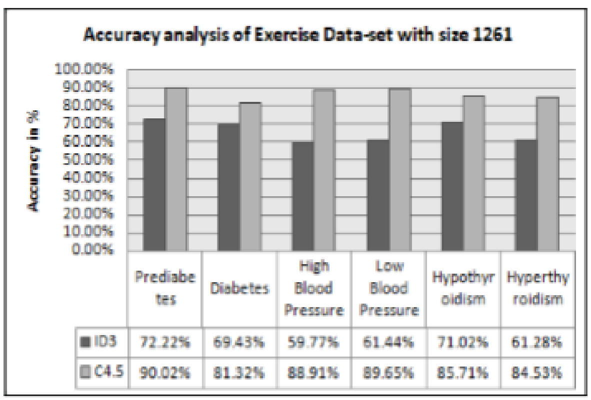

Fig. 14. Comparison of Accuracy for ID3 and C4.5 Decision tree Algorithm with Diet Data-set with 1261 size

VIII. CONCLUSION

This presentation covers ground in the area of health care machine learning. We built a system that may assist physicians in making dietary and exercise recommendations to their patients. Using the C4.5 decision tree algorithm, a machine learning technique, it deals specifically with health monitoring of diseases like diabetes, hypertension, and thyroid based on the patient's most recent report, looking for improvements in each follow-up session and recommending an appropriate and updated exercise and dietary plan based on the reports and other credentials like height, weight, age, and activity level. Between the two datasets we tested, it is easy to see that C4.5 outperforms the ID3 method. Although C4.5 is a prediction model, we find that it may be utilized for recommendations with further enhancements.

IX. FUTURE WORK

Future work will include an exercise and diet tracking system that, in the event that the user's preferences change, will offer alternative options based on their health concerns. Additionally, we will create an alert system that will remind the user before each follow-up session and notify them in the event of extreme reports.

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] Shurouq Hijazi, Alex Page, Burak Kantarci, Tolga Soyata, ‘Machine Learning in Cardiac Health Monitoring and Decision Support’ Com- puter , vol. 49, Issue. 11, pp. 38 - 48, Nov. 2016.

[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, USA, 4-9 April 2019.

[4] Yurong Zhong, ‘The analysis of cases based on decision tree,’ 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, Aug. 2016,

[5] He Zhang, Runjing Zhou, ‘ The analysis and optimization of decision tree based on ID3 algorithm’, 2017 9th International Conference on Modelling, Identification and Control (ICMIC),Kunming, China, 10-12 July 2017.

[6] Nawras Georgi, Regine Le Bouquin Jeanne`s, 'Proposal of a remote monitoring system for elderly health prevention', 2017 International Conference on Smart, Monitored and Controlled Cities (SM2C), Sfax, Tunisia, 17-19 Feb. 2017.

[7] Akash Maurya, Rahul Wable, Rasika Shinde, Sebin John, Rahul Jadhav, R Dakshayan 'Chronic Kidney Disease Prediction and Recommendation of Suitable Diet Plan by using Machine Learning', 2019 International Conference on Nascent Technologies in Engineering (ICNTE), Navi Mumbai, India, 4-5 Jan. 2019.

[8] Romeshwar Sookrah, Jaysree Devee Dhowtal, Soulakshmee Devi Nagowah, 'A DASH Diet Recommendation System for Hypertensive Patients Using Machine Learning,' 2019 7th International Conference on Information and Communication Technology (ICoICT), Kuala Lumpur, Malaysia,24-26 July 2019.

[9] Esther Gyimah, Delali Kwasi Dake , 'Using Decision Tree Classification Algorithm to Predict Learner Typologies for Project-Based Learning', 2019 International Conference on Computing, Computational Modelling and Applications (ICCMA), Cape Coast, Ghana, 27-29 March 2019.

[10] Vivek Kumar, Roman Zinovyev, Abhishek Verma, Prayag Tiwari, 'Per- formance Evaluation of Lazy and Decision Tree Classifier: A Data Mining Approach for Global Celebrity's Death Analysis,' 2018 In- ternational Conference on Research in Intelligent and Computing in Engineering (RICE), San Salvador, El Salvador,22-24 Aug. 2018.

[11] G. Agapito, B. Calabrese, P. H. Guzzi, M. Cannataro, M. Simeoni, I. Care´, T. Lamprinoudi, G. Fuiano,A. Pujia, 'DIETOS: A recommender system for adaptive diet monitoring and personalized food suggestion,' 2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), New York, NY, USA,17-19 Oct. 2016.

[12] Mihnea Donciu, Madalina Ionita, Mihai Dascalu, Stefan Trausan-Matu, 'The Runner – Recommender System of Workout and Nutrition for Runners', 2011 13th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, Timisoara, Romania,26-29 Sept. 2011

[13] Li Dongming, Li Yan, Yuan Chao, Li Chaoran, Liu Huan, Zhang Lijuan, 'The Application of Decision Tree C4.5 Algorithm to Soil Quality Grade Forecasting Model ', 2016 First IEEE International Conference on Computer Communication and the Internet (ICCCI), Wuhan, China, 13- 15 Oct. 2016.

[14] Izhar Alam, 'Diet Recommendation' [Data set] 2020. Available: https://www.kaggle.com/izharalam150/diet-recommendation.

[15] Joe89, '721 Weight Training Workouts' [Data set] 2018. Available: https://www.kaggle.com/joep89/weightlifting.