International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 6, pp 147-152, June-2023

https://doi.org/10.47001/IRJIET/2023.706023

Diet Recommendation and Fitness Using Machine Learning

[1]Neeraj Nalawade, [2]Himanshu Deore, [3]Lancer Lobo, [4]Dr. Shailaja Patil

[1,2,3]Student, Electronics and Telecommunication Engineering, JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

[4]Professor, Electronics and Telecommunication Engineering, JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

***Abstract* - The fitness industry market has become one of the most important social transformations of the 21st century. After COVID-19 the importance of maintaining good health and stamina has increased. The number of people interested in sport has not stopped to increase, so we can speak about a field with dual demographic and economic issues. As the popularity of gym and fitness clubs has increased it has been difficult to maintain the data using the traditional method of pen and paper or a normal excel sheet. We observed a need to maintain the database of the gym members in order keep track of the members attending the gym, to maintain their personal details and subscription package. As we know for maintaining and improving the health of the Gym members along with regular usage of the gym eating the right food in specific quantity is required to achieve the required goals of the gym member. Most of the gym going people especially in India neglect the diet part of the workout or spend hefty amounts of money to acquire the same which is not cost effective.**

Keywords: Diet, Fitness, Machine Learning, AI, Gym, health.

I. INTRODUCTION

Most often, when individuals think of calories, they only think of food or weight reduction. A calorie, however, is often a measure of heat energy. Calories are the units of energy needed to elevate 1 gramme (g) of water by 1°C. The measurement may be used to assess a variety of energy-releasing systems unrelated to the human body. The amount of energy needed by the body to carry out a task is how many calories are considered from the perspective of the human body. There are calories in food. Each and every item has a distinct quantity of energy included in it since various foods have varying calorie counts.

The temperature of the body and the heartbeat will start rising up when we perform exercise or some heavy workout. The carbohydrates or carbs are broken down into glucose which is further converted/broken down into energy using O_2 (oxygen). The variables used here are the timescale the person is training, the average heart rate per minute, and the temperature. Then get more height, weight, gender, and age of the person to predict the tonnage of energy that the person burns. Parameters that can be considered for input are the duration of exercise, average heart rate per minute, temperature, height, weight, and gender. A machine learning XG Boost regressor algorithm is used to predict calories burned depending on exercise time, temperature, height, weight, and age.

II. LITERATURE SURVEY

Manal et. al., proposed a machine learning based pipelined approach for predicting the calories from food images. The system takes an image of the food item and passes it through Math works Image Processing which extracts the raw features and improves the quality. The image is passed through a compression phase which helps to reduce the number of features using the Principal Component Analysis (PCA) method and scale the subsequent learning phases. The food type classification is done by inputting the compressed image to the classifier. The food size prediction is done by passing the compressed image to a regressor. Calories are predicted by passing the compressed image and predicted values to another regressor. This is based on supervised learning model [1].

Kohila et. al., proposed a calorific value prediction mechanism using image processing and machine learning. The image of the food is transmitted through a mobile device and it initially undergoes image processing and final output is displayed. The mathematical morphology is utilized as a tool for extracting the image components and the region shape description such as erosion, dilation, opening and closing. Feature extraction is performed to retrieve interesting parts of the image and then calorie measurement is done [2].

Kiran et al., proposed a method for measuring the calories and nutrition from food images using machine learning techniques. The images got from the mobile device

© 2023-2017 IRJIET All Rights Reserved

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 6, pp 147-152, June-2023

https://doi.org/10.47001/IRJIET/2023.706023

are pre-processed followed by the segmentation step to extract the colour and texture features through K Means clustering. The food portion volume measurement is done by superimposing a grid of squares onto the image segment which matches the irregular shape of the food images easily. The calorie measurement is done based on the food mass and nutritional tables. The system has limited cuisine varieties mixed food images have not been considered [3].

Chang et al., proposed a food image recognition for computer aided dietary assessment based on deep learning techniques. The proposed approach utilized two real-world food image datasets namely (UEC – 256 and Food – 101). The food image recognition is done by a new Convolutional Neural Network (CNN) method based on supervised learning algorithms. The CNN consisted of 3 convolutional layers, 2 sub sampling layers and a fully connected layer. The model was trained for a nonstop period of 2 to 3 days using a server with NVidia K40 GPU. After the training, the model classified the image in less than a minute. The proposed system lacked real world data and the accuracy of the measurements have to be improve [4].

Yanchao et al., proposed a calorie estimation model based on deep learning approach which increases the detection accuracy and reduces the error of volume estimation. The image attainment is done by obtaining the food image using a smartphone. The object detection is done by using Faster Region based Convolutional Neural Networks (Faster R-CNN), which includes Region Proposal Network (RPN) and an Object Detection Network. Grab Cut, an image-based segmentation algorithm which depends on optimization by graph cuts is used for image segmentation. Volume Estimation requires calculation of side and top views scaling factor using equations based on the shape types. Finally, calorie estimation is obtained by using the volume and density value of the food mapped using ECUSTFD dataset [5].

Liang et al., proposed a "Calorie estimation method" which was designed for obese patients to check their food intake per day. This method is based on computer vision technique which requires the top and side view of food to estimate calorie from it. For the detection of food items, one yuan coin is taken as a calibration object. The Faster R-CNN algorithm is used and the contour of each food is detected using grab cut algorithm whereas the volume of food is estimated using volume estimation formulas and at last calories of each food are estimated as the output [6].

Podutwas et al. proposed a system named "food portion recognition system" which measures the calorie and nutrition values by taking the picture of food and then detect and classify the food portion using SVM. Segmentation and food portion recognition are performed using skull stripping and classification using SVM to calculate calorie and nutrients present in food [7].

III. METHODOLOGY

We take input parameters like duration of exercise, heart rate, height, weight, gender, age to calculate calories burnt. Then we take these calories burnt into the dataset for diet recommendations and other parameters and recommend appropriate diet for breakfast, lunch and dinner for the users.

All this is done using the ML algorithms like XGBoost, Random Forest and K Mean Clustering.

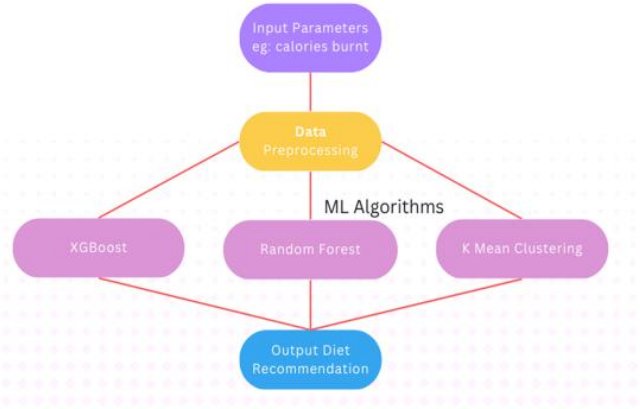

Figure 1: System Architecture

IV. WORKING PRINCIPLE

- Dataset procurement from open-source websites.
- We looked at various datasets on sites like Kaggle, then selected datasets based on number of data points, structuredness and variety of data.
- Machine learning algorithms selection was based on dataset analysis and research papers. We used the XGBoost Regression algorithm for the classification of the calorie dataset. Trained the model for predicting the calorie burnt during the exercises.
- The foods in datasets are categorical data, hence we decided on using classification based algorithms like K Mean Clustering and Random Forest Classifiers.
- Trained the machine learning models viz. K Mean Clustering and Random Forest Classifier. We used Google Collaband python modules to train our models.
- Tested the model for accuracy and recall. We tested the model's accuracy by using Confusion Matrices, accuracy scores, etc.
- We will then export the model using Pickle module.
- Then we will export the model to integrate it with our website UI. We will then create a Preliminary UI to test the model interface.

© 2023-2017 IRJIET All Rights Reserved

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 6, pp 147-152, June-2023

https://doi.org/10.47001/IRJIET/2023.706023

- We will then integrate the model with the UI.
- After creating a simple UI, we will integrate it with our ML model and test it by giving sample calorie burnt numbers.

V. ALGORITHMS USED

A) XGBoost Algorithm

XgBoost stands for Extreme Gradient Boosting, which was proposed by the researchers at the University of Washington. It is a library written in C++ which optimizes the training for Gradient Boosting.

Before understanding the XGBoost, we first need to understand the trees especially the decision tree:

Decision Tree:

A Decision tree is a flowchart-like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label.

A tree can be "learned" by splitting the source set into subsets based on an attribute value test. This process is repeated on each derived subset in a recursive manner called recursive partitioning. The recursion is completed when the subset at a node all has the same value of the target variable, or when splitting no longer adds value to the predictions.

Bagging:

A Bagging classifier is an ensemble meta-estimator that fits base classifiers each on random subsets of the original dataset and then aggregate their individual predictions (either by voting or by averaging) to form a final prediction. Such a meta-estimator can typically be used as a way to reduce the variance of a black-box estimator (e.g., a decision tree), by introducing randomization into its construction procedure and then making an ensemble out of it.

Each base classifier is trained in parallel with a training set which is generated by randomly drawing, with replacement, N examples (or data) from the original training dataset, where N is the size of the original training set. The training set for each of the base classifiers is independent of each other. Many of the original data may be repeated in the resulting training set while others may be left out.

Bagging reduces over fitting (variance) by averaging or voting, however, this leads to an increase in bias, which is compensated by the reduction in variance though.

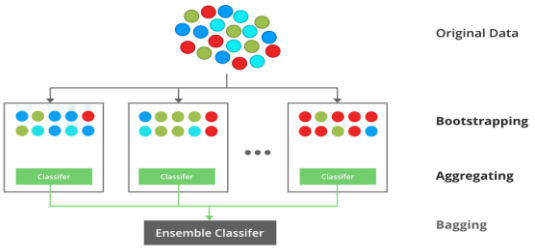

Figure 2: Bagging classifier

Random Forest:

Every decision tree has high variance, but when we combine all of them together in parallel then the resultant variance is low as each decision tree gets perfectly trained on that particular sample data and hence the output doesn't depend on one decision tree but multiple decision trees. In the case of a classification problem, the final output is taken by using the majority voting classifier. In the case of a regression problem, the final output is the mean of all the outputs. This part is Aggregation.

The basic idea behind this is to combine multiple decision trees in determining the final output rather than relying on individual decision trees.

Random Forest has multiple decision trees as base learning models. We randomly perform row sampling and feature sampling from the dataset forming sample datasets for every model. This part is called Bootstrap.

Boosting:

Boosting is an ensemble modelling, technique that attempts to build a strong classifier from the number of weak classifiers. It is done by building a model by using weak models in series. Firstly, a model is built from the training data. Then the second model is built which tries to correct the errors present in the first model. This procedure is continued and models are added until either the complete training data set is predicted correctly or the maximum number of models is added.

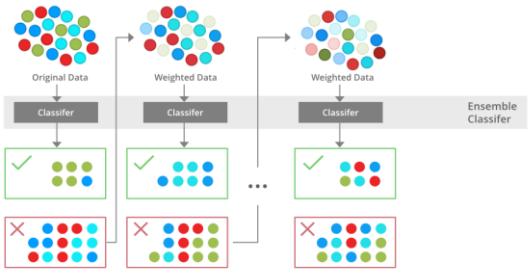

Figure 3: Boosting

© 2023-2017 IRJIET All Rights Reserved

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 6, pp 147-152, June-2023

https://doi.org/10.47001/IRJIET/2023.706023

Gradient Boosting:

Gradient Boosting is a popular boosting algorithm. In gradient boosting, each predictor corrects its predecessor's error. In contrast to Adaboost, the weights of the training instances are not tweaked, instead, each predictor is trained using the residual errors of predecessor as labels.

There is a technique called the Gradient Boosted Trees whose base learner is CART (Classification and Regression Trees).

XGBoost:

XGBoost is an implementation of Gradient Boosted decision trees. XGBoost models majorly dominate in many Kaggle Competitions.

In this algorithm, decision trees are created in sequential form. Weights play an important role in XGBoost. Weights are assigned to all the independent variables which are then fed into the decision tree which predicts results. The weight of variables predicted wrong by the tree is increased and these variables are then fed to the second decision tree. These individual classifiers/predictors then ensemble to give a strong and more precise model. It can work on regression, classification, ranking, and user-defined prediction problems.

B) Random Forest Algorithm

Random forest is a type of algorithm which is used for classification and regression. As we know the random forest is a one of supervised learning algorithm which means random forest algorithm uses the technique based on supervised learning if we talk about supervised learning in simple word supervised means a supervisor which gives the instruction, i.e. training data which gives input and output and based on the input and output of training data we are going to prepare a model and we will give new input to that model and check the output whether the valid output is coming.

The random forest is a type of ensemble classifier which is using the decision tree algorithm in a randomized fashion. It consists of many trees which are called decision trees and these trees are of different structures and to make a decision tree we are choosing features and samples randomly from the training dataset and that's how we construct many decision trees and combined all the decision trees makes a random forest.

How random forest works?

- Initially, we should have training data which consist of various attributes and target attribute. Now, we have to make a decision tree, to make a decision tree we have to generate BD (Bootstrap dataset) and make BD we have to do sampling which means we have to pick any sample randomly from the training dataset and put it into Bootstrap dataset. Duplication is allowed with less frequency. Using BD, we have to plot a decision tree in a randomized fashion and calculate how we can choose the root nod from the BD which is producing the best split of samples.
- Again do the splitting of features for child node and provide the leaf node to the child node after splitting of features.
- Repeat steps 2–4 and makes as many decision trees as we can.
- Take the test tuple and let the model classify and predict the output of the given test tuple.
- Now, calculate the votes produced by various decision trees.
- Consider the majority of votes produced for the target attribute of test tuple and that will be a final prediction.

C) K-Mean Clustering Algorithm

We are given a data set of items, with certain features, and values for these features (like a vector). The task is to categorize those items into groups. To achieve this, we will use the K-Means algorithm; an unsupervised learning algorithm. 'K' in the name of the algorithm represents the number of groups/clusters we want to classify our items into.

The algorithm will categorize the items into k groups or clusters of similarity. To calculate that similarity, we will use the Euclidean distance as measurement.

The algorithm works as follows:

- First, we initialize k points, called means or cluster centroids, randomly.
- We categorize each item to its closest mean and we update the mean's coordinates, which are the averages of the items categorized in that cluster so far.
- We repeat the process for a given number of iterations and at the end, we have our clusters.
- The "points" mentioned above are called means because they are the mean values of the items categorized in them. To initialize these means, we have a lot of options. An intuitive method is to initialize the means at random items in the data set. Another method is to initialize the means at random values between the boundaries of the data set (if for a feature x the items have values in [0,3], we will initialize the means with values for x at [0,3]).

VI. RESULTS

The some of the result screenshots are as follows:

© 2023-2017 IRJIET All Rights Reserved

ISSN (online): 2581-3048
Volume 7, Issue 6, pp 147-152, June-2023
https://doi.org/10.47001/IRJIET/2023.706023

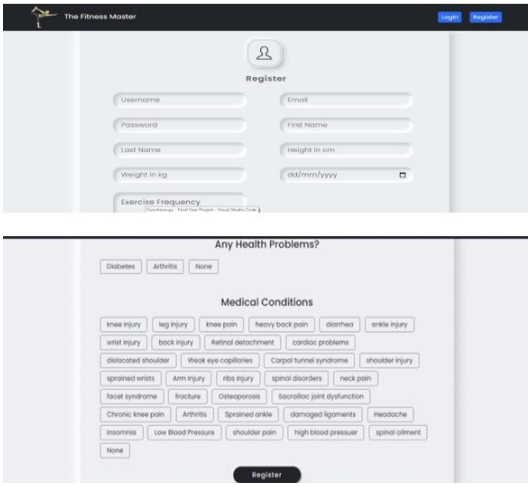

Figure 4: Result

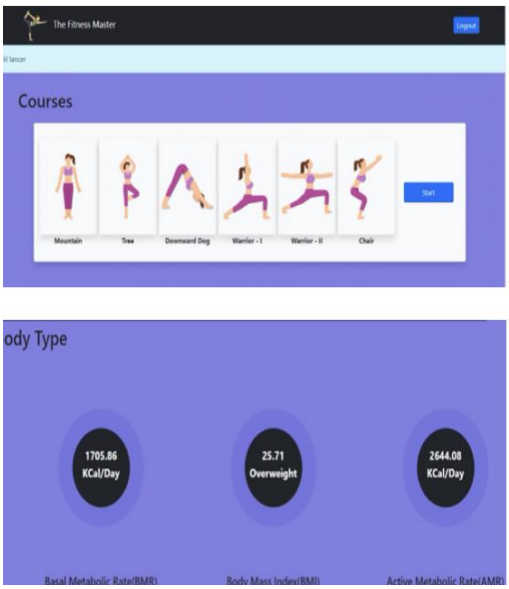

Figure 5: Result

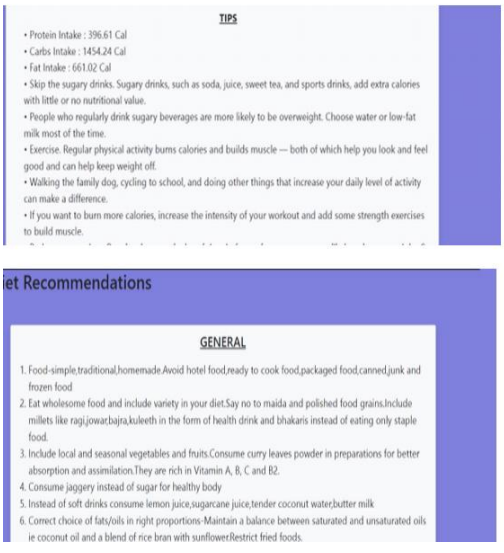

Figure 6: Result

VII. CONCLUSION

We deduced from the analysis that the XGB Regressor produces more accurate findings. Mean absolute error suggests that absolute error should be as minimal as possible. It is nothing more than the discrepancy between values that were seen and those that were predicted by models. 2.71 is a good value for the mean absolute value that the XGB Regressor gives us. The mistake rates are quite low. Therefore, we can say that XG Boost Regressor is the best model for predicting calorie burn. The flexibility of the suggested technique can also be improved with variations. In this study, we have concentrated on the seven primary factors that influence how many calories our body burns, but there are other factors that also play a role. It's also crucial to understand how many calories we are consuming if we want to stay healthy and fit. Additionally, ML may be used to construct this (machine learning). A UI (user interface) is also required so that users may input their values and obtain results that show how many calories they have burned. Additionally, we are able to create a completely functional app with all of these features and our recommended diet and exercise regimen.

A Diet Recommendation System is implemented with the working functionalities like: Desired food list prediction, Weight category prediction, BMI Calculation, Health is vital for an individual and can be achieved with this working module. Thus making life healthy.

REFERENCES

[1] Manal Chokr and Shady Elbassuoni, “Calories Prediction from Food Images”, Innovative Applications of Artificial Intelligence Twenty - Ninth IAAI Conference, 2017.

[2] Kohila R and Meenakumari R, “Predicting calorific value for mixed food using image processing”, In Proceedings of the 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017, pp. 1–4.

[3] Kiran Abhore and Dawande N A, “Measuring Calories and Nutrition from Food Image”, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5, Issue 6, 2016.

[4] Liu C, Cao Y, Luo Y, Chen G, Vokkarane V and Ma Y, “Deepfood: Deep learning-based food image recognition for computer aided dietary assessment”, International Conference on Smart Homes and Health Telematics, 2016, pp. 37-48. doi: https://doi.org/10.1007/978-3-319-39601-9_4.

[5] Yanchao Liang and Jianhua Li, “Deep Learning-Based Food Calorie Estimation Method in Dietary Assessment”, 2018, arXiv preprint arXiv: 1706.04062.

© 2023-2017 IRJIET All Rights Reserved

ISSN (online): 2581-3048
Volume 7, Issue 6, pp 147-152, June-2023
https://doi.org/10.47001/IRJIET/2023.706023

[6] Liang, Y., & Li, J. (2017). Deep Learning-Based Food Calorie Estimation Method in Dietary Assessment. ArXiv, abs/1706.04062.
[7] Podutwar, M. A., Pawar, P. P., & Shinde, P. A. (2017). A Food Recognition System for Calorie Measurement.

AUTHORS BIOGRAPHY

Neeraj Nalawade, Student, Electronics and Telecommunication Engineering, JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India.

Himanshu Deore, Student, Electronics and Telecommunication Engineering, JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India.

Lancer Lobo, Student, Electronics and Telecommunication Engineering, JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India.

Dr. Shailaja Patil, Professor, Electronics and Telecommunication Engineering, JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India.

Citation of this Article:

Neeraj Nalawade, Himanshu Deore, Lancer Lobo, Dr. Shailaja Patil, "Diet Recommendation and Fitness Using Machine Learning" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 6, pp 147-152, June 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.706023

****

© 2023-2017 IRJIET All Rights Reserved