

1. Calories Burnt Prediction Using Machine Learning

Author: **Rachit Kumar Singh, Vaibhav Gupta**

http://www.ijaresm.com/uploaded_files/document_file/Rachit_Kumar_Singh822z.pdf

ABSTRACT

- Machine Learning is a category of algorithms that allows software applications to become more accurate in predicting outcomes without being explicitly programmed.
- The object of this research paper is to create a project that can be used to predict calories burnt using Machine Learning with Python.
- The Xgboost Regression model is used in this project.
- The muscles that are doing exercise need more oxygen and as the body requires more oxygen the heart beat will increase a lot, so increased heart beat means increased blood flow which in turn will give more oxygen to the muscle which is used to break those glucose molecules, so the energy from these glucose molecules is used and when it is used only a part is used and rest is converted into heat, so body temperature would increase and our body will sweat, so the parameters which we would be taking into consideration for input are :- Duration for which the exercise is done, Average heart beat per minute, Body temperature, Height, weight and gender of the person.
- All these would be used to create a prediction model also for calories burnt.

XGBOOST REGRESSOR

- The Xgboost regressor has two parameters λ which denotes Regulation parameter , more the regulator parameter more is the tuning of the decision tree and the other parameter is γ which is threshold.
- If $\text{Similarity Weight(SM)} = \frac{\sum(\text{Residue})^2}{(\text{No of Residues} + \lambda)}$ Then $\text{Gain} = \text{Similarity Weight(left decision tree)} + \text{Similarity Weight(right decision tree)} - \text{Similarity Weight(root)}$. If $(\text{Gain} > \gamma)$ then decision tree bifurcation takes place for further levels else not takes place, this makes the xgboost algorithm efficient as compared to others.
- Features of XGBOOST :-
 1. It is platform free.
 2. Xgboost has a high speed of processing.
 3. Xgboost uses parallelization, and uses the maximum available computational power of the system.

METHODOLOGY

- The data is analyzed to show that the heart rate and body temperature would be more when the person is doing exercise.
- We will find the correlation in the dataset by constructing a heat map.
- Then we will train the model using the xgboost regressor

Data Visualization :- find distribution of gender column for example how many males in this data point and how many females in this data point in count-plot training the model with X_train, my xgboost regressor will learn from the model automatically prediction on training data

Multiple instance parameters and various factors can be used to make the calories prediction more innovative and successful. Accuracy, which plays a key role in prediction-based systems, can be significantly increased. One of the main highlights is more expressive regression outputs, which are more understandable and bounded with some accuracy. Moreover, the flexibility of the proposed approach can be increased with variants .

2. Calorie Burn Prediction Analysis Using XGBoost Regressor and Linear Regression Algorithms

Author: **Sona P Vinoy , Binumon Joseph**

<https://zenodo.org/record/6365018>

- The machine learning XGBOOST algorithm and linear regression is used to predict calorie burned which depends on height , weight , workout duration, body temperature and age of a person
- This document predicts the calorie burned during the workout of different people and compare the two algorithms through datasets

- Calorie is a unit of energy or heat variously defined . Defined as the amount of energy that is vital to increase the temperature of 1 gm of water by 1 degree celsius in the case of the human body, calories are the measure of how much the body requires to function.
- Many factors affect calorie burn some of these factors are under direct control of us while others cant be controlled these are
 - Age: Older person burns lesser calories
 - Gender: men burn greater calories than women
 - Body Composition: Those with more muscular tissues burn more calories
 - Body Size : Larger people burn more calories even at relaxation
 - Thermogenesis: Amount of strength the body uses to break down meals.
- So the parameters which we would be taking into consideration for input are :-
 1. Duration for which the exercise is done,
 2. Average heart beat per minute
 3. Body temperature
 4. Height
 5. weight
 6. gender of the person
- Steps
 - Collecting data
 - Data preprocessing
 - Data analysis

- Machine learning model
 - Evaluation
- Data taken from kaggle has 1500 instances and 7 attributes
 - We are looking for a model with less absolute mean error which gives more accurate results. Hence XGBOOST regressor gives more accurate results of calories burned over linear regression.

In conclusion XGBOOST has more accurate results than linear regression model

3. Calorie Burn Prediction using Machine Learning

Author: **Suvarna Shreyas Ratnakar , Vidya S**

<https://iarjset.com/wp-content/uploads/2022/07/IARJSET.2022.961125.pdf>

Daniel Bubnis The variety of burned energy in day-to-day life is directly related to weight maintenance, weight gain, or weight loss.

1. People need to burn more calories than they consume, causing a calorie deficiency.
2. But they want to know how many calories they burn every day.
3. Most people think that calories are most effectively associated with food and weight loss. Calories are variously defined units of energy or heat.
4. For men or women trying to gain, lose, or maintain weight, it is essential to know how many calories they are consuming each day.
5. Salvador Camacho the global obesity crisis has been continuously increasing, and thus far no nation has been able to turn it around.

6. The World Health Organization identifies an energy imbalance between calories ingested and calories expended as the root cause of obesity.
7. But mounting data indicates that the idea of calorie imbalance might not be enough to control and stop the obesity pandemic.
8. To examine the calorie imbalance idea and its components as a weight-management tool as well as any potential drawbacks and implications for public health, with the goal of highlighting the need for an updated theory about the origins of obesity.
9. This revision could better direct public health initiatives to control obesity by avoiding weight increase or encouraging weight reduction. World Health Organization.

DATA SOURCE

We use "Kaggle" as our dataset store. A total of 15000 instances and 7 attributes of data are present throughout 2 CSV files. The "Kaggle" repository's dataset comprises information about a variety of people, including their height, weight, gender, age, workout intensity, heart rate, and body temperature. The training data is taken from the "exercise.csv" and "calories.csv" datasets. Additionally, the user id-mapped target class from the second calorie dataset comprises the calories burned by the person in the exercise dataset.

This is the stage in which we apply the algorithm we've chosen (in this case, XGBoost regressor) to determine the mean absolute error. The XGB regression procedure is used, and the results are obtained. For this, we employ metrics that indicate the magnitude of errors the version is committing. The XGBoost regressor algorithm was proven to be an effective and efficient method in predicting the calories burnt prediction.

XGBOOST REGRESSOR

In machine learning the XGBoost Algorithm performs well since it has robust handling of many variety of data types, relationships, distributions, and the many hyperparameters that you can fine tune

XGBoost regressor can be used for regression, classification for both binary and Multiclass and ranking problems

- XGBoost is an efficient implementation of gradient boosting that can be used for regression predictive modeling.
- Xgboost keeps all intermediate calculations in cache so that we don't have to do the same calculation again and again.

We deduced from the analytics that XGBOOST produces more accurate findings

In conclusion even linear regression can be used to predict the relationship between independent and dependent variables but XG BOOST gives a lesser error compared to linear regression and hence more efficient

4. SURVEY ON APPLICATIONS FOR PREDICTING CALORIES AND NUTRITIONAL VALUES FROM FOOD

Author: **Rachoti Biradar , Rahul Thiru , Ravi Purwar**

Guided By : Latha N

<https://www.citefactor.org/journal/pdf/SURVEY-ON-APPLICATIONS-FOR-PREDICTING-CALORIES-AND-NUTRITIONAL-VALUES-FROM-FOOD.pdf>

Using image segmentation to separate the food from the background image. This technique will increase the effectiveness of food identification.

In this paper, we take some initial steps towards such a system. Our approach utilizes several deep learning algorithms, tailored to run on a conventional mobile phone, trained to recognize food items and predict the nutritional contents of meals from images that are taken in the wild .

Calories Analysis of Food Intake Using Image Recognition In this model, the authors[11] have developed a model that is used to find calories contained in that food. Here they have used 3 techniques: Image segmentation, feature extraction and Classification. This model mainly concentrates on Thai food.

The authors have created a model that takes food images as input and output sequence of cooking instructions using instruction decoder which separates out the ingredients and thereby separates out the recipe .

The newly proposed time-dependent food distribution model and weight optimization algorithm to make personalized classifiers learn the user's data and adapt to users eating habits.

In all the above systems, a method for measuring calories and nutrition of food objects is carried out. The systems help people closely control their daily food intake. They focused on identifying food and extracting ingredients from the food image and predicting the calorie content from the quantity of food.

5. Burned Calories Prediction using Supervised Machine Learning: Regression Algorithm

<https://ieeexplore.ieee.org/document/9776710/authors#authors>

Authors

Marte Nipas | Aimee G. Acoba | Jennalyn N. Mindoro | Mon Arjay F. Malbog | Julie Ann B. Susa | Joshua S. Gulmatico

- The calories burned depends on many factors such as weight, gender, age, height, metabolism, and the type of activity or exercise done.

- The researchers did not modify any parameters in the regression models. It is recommended to alter some of the parameters to see if there are significant changes to the score.
- The study considered three regression models, namely linear regression, ridge regression, and random forest regression. The three models were trained with the dataset using the program developed in python programming. To test the models, K-fold cross-validation with ten iterations was used to reduce the chances of overfitting and to improve the prediction accuracy.
- By taking the average, it can infer that Linear regression has an accuracy of 92.88%, Ridge regression is 92.82%, and Random Forest regression is 95.77%. This shows that Random Forest Regression is the best model for this study because of its capability to discover complex behaviors in the data being understudy

6. PREDICTION OF USER'S CALORIE ROUTINE USING CONVOLUTIONAL NEURAL NETWORK

<https://www.ijeast.com/papers/189-195.Tesma503,IJEAST.pdf>

Authors :

Sathiya T | Surya Prakash B | Thirukkumaran S V | Vijaiarivalagan K

- The CNN model is trained using preprocessed dataset. The image preprocessing techniques are applied to images in dataset.
- paper is to empower the user by a convenient, intelligent system that helps them become sensible about their calorie intake
- CNNs are consuming high computational time, But the feature-based approach is highly appreciable
- Every day calories of the user are stored in the database for further analysis. An optional option is also given to user to input their blood pressure and blood sugar values for providing food suggestion.
- Calculate the burned calorie by increasing their steps, physical exercises and by doing changes in their daily routines for analysis.

7. Feature Selection Intent Machine Learning based Conjecturing Workout Burnt Calories

AUTHORS: N. Manjunathan a, M. Shyamala Devib, S. Sridevi c, Kalyan kumar Bonalad, Ankam Kavithae and Konkala Jayasreef

- The Exercise dataset with 8 independent variables and 1 dependent variable has been used for implementation. The prediction of burnt out calories is done with the following contributions.
- Firstly, the data set is preprocessed with Feature Scaling and missing values.

- Secondly, exploratory feature examination is done and the scattering of target highlight is visualized.
- Thirdly, the raw data set is fitted to all the regressors and the execution is dissected before and after scaling.
- Fourth, the raw data set is subjected to feature selection axioms like Anova test, Correlated Feature, Variance Based and KBest Feature based methods and are fitted to all the regressors and the performance is analyzed before and after feature scaling.
- Fifth, performance analysis is done using metrics like MAE, MSE, EVS, RScore and running time. Fig. 1 shows the overall workflow of this work
- An attempt was made to find the performance analysis of the exercise dataset in forecasting the workout burnt calories by applying various feature selection methods with High correlation, High Variance, Anova Test analysis and KBest Feature analysis.
- Experimental results shows that the Decision Tree and Gradient Boosting regressors tends to retain 99% before and after feature scaling for the Anova test, Correlated Feature, Variance Based and KBest Feature based methods.

8.Multi-Task Learning for Calorie Prediction on a Novel Large-Scale Recipe Dataset Enriched with Nutritional Information

Author: [Robin Ruede](#), Verena Heusser, Lukas Frank, Alina Roitberg, Monica Haurilet, Rainer Stiefelhagen

<https://arxiv.org/pdf/2011.01082v1.pdf>

- In this work, we aim to estimate the calorie amount of a meal directly from an image by learning from recipes people have published on the Internet, thus skipping time-consuming manual data annotation.
- Since there are few large-scale publicly available datasets captured in unconstrained environments, we propose the pic2kcal benchmark comprising 308,000 images from over 70,000 recipes including photographs, ingredients and instructions.
- To obtain nutritional information of the ingredients and automatically determine the ground-truth calorie value, we match the items in the recipes with structured information from a food item database.
- We evaluate various neural networks for regression of the calorie quantity and extend them with the multi-task paradigm.
- Our learning procedure combines the calorie estimation with prediction of proteins, carbohydrates, and fat amounts as well as a multi-label ingredient classification
- Our experiments demonstrate clear benefits of multi-task learning for calorie estimation, surpassing the single-task calorie regression by 9.9%.
- To encourage further research on this task, we make the code for generating the dataset and the models publicly available.
- In this work, even though we focus on calorie estimation, the dataset contains additional attributes that can characterize the meals. In addition to calories, we include the type of the meal

,ingredient quantities, cooking instructions, an average user rating, preparation time, tags and further properties.

9.The Calorie Burning Calculation System in Jogging Using a Thresholding-Based Accelerometer Sensor

Authors: **Finanta Okmayura | Adlian Jefiza | Witri Ramadhani**

<https://www.researchgate.net/scientific-contributions/Witri-Ramadhani-2167805719>

- Researchers hoped that this application can make it easier for athletes to get their calorie burning information while jogging.
- The application of calorie burning calculation in jogging using a thresholding-based accelerometer sensor was called the Jogging Calories Burning Calculation (JCBC). The following was the display of the application when a new user first entered this JCBC application. It could be seen in several menus in the application, namely jogging menu, health info and about us. Each menu has its own function. Before using this menu, the user must first fill in some data, namely name, age, gender, weight, and height.

