RESEARCH ARTICLE OPEN ACCESS

Prediction of Burned Calories using Random Forest Regression Algorithm

Thalakola Syamsundararao

Associate professor, KKR & KSR Institute of Technology and Sciences, Department of CSE-DS

ABSTRACT

To be healthy and fit, you must engage in regular physical activity. An equation and MET charts are used to estimate how many calories each person burns during exercise. In order to get more precise findings, this study uses a regression model as one of the machine learning algorithms to predict the number of calories burned. Prior to being fed into regression models, data must first be prepared, cleaned, and analysed. Based on the findings of model testing exercises, the efficacy and performance of linear regression models were assessed. The average accuracy was calculated, and the result demonstrates that the best model for the study is linear regression, with an accuracy of 96.75%. The ability of the algorithm to predict the value of the target variable may be impacted by the visualisation and analysis of the relationships between the variables in the data. In this study, we used the Random Forest regression approach to enhance calorie prediction.

Keywords: Cleaning, Random Forest, Computational Modelling, Data Visualization, Predictive Models, and Machine Learning Algorithms.

I. INTRODUCTION

Body fat is also necessary for storing energy, safeguarding internal organs, and supporting vital processes including growth, immunity, hormone synthesis, reproduction, and metabolism, according to the National Institute of General Medical Sciences. Men need between 2 and 5 percent of body fat, while women need between 10 and 13%, according to the American Council on Exercise (ACE). [1] However, these bare minimums might not be adequate. A healthy range for persons under the age of 40 is 8 to 20 percent for males and 21 to 33 percent for women, according to the most often referenced study on the subject, even though there is no official recommendation for the ideal body fat percentage.

Nevertheless, the connection between health and body fat is nuanced and poorly understood. Focus on the things you can control rather than worrying about how your birth sex influences your calorie burn. [2] The final line, according to Gonzalez, is that both men and women should concentrate on increasing their muscle mass and cardiovascular fitness through a well-balanced cardio and strength-training regimen. If you've ever taken a group exercise class where each participant's heart rate and projected calorie burn are shown on a screen, you are aware of how much these numbers fluctuate from person to person. Additionally, you may have noticed that men often burn more calories than women. But have you ever questioned why, even during the same workout, different people burn calories at such disparate rates?

There are six variables that can influence your calorie burned:

i) Body Mass: Kyle Gonzalez, a qualified strength and conditioning specialist and performance coach at Future with a San Francisco basis, states that in general, the more weight you have, the more calories you'll burn every session. The more weight you have, the more energy it takes to move your body because calories are merely a measure of energy. In other words, if two persons have differing weights, the person who is heavier will burn more calories since they expend more energy when they move. Larger internal organs (such as the heart, liver, kidneys, and lungs) are more common in people with larger bodies, which has an impact on how many calories are used during exercise and at rest. These organs and their processes demand energy. [3] According to one study, variations in internal organ size can account for up to 43% of the variation in total calories burned among individuals. One of the many reasons why losing weight is difficult is that your body burns fewer calories as you lose weight, which can result in a weight loss plateau or even weight gain. However, it's not the only explanation.

According to a recent review, losing weight can also result in other physiological changes, such as the body's propensity to burn stored fat for energy (a process known as fat oxidation), increased ghrelin levels, which increase appetite, and decreased leptin levels, which decrease satiety. Consider working with a registered dietician who specialises in weight reduction to help you reach your goal in a healthy and sustainable way if you're trying to lose weight but have reached a plateau. Visit the Academy of Nutrition & Dietetics to locate one. Additionally, remember that exercise is beneficial for your health

in general whether or not you lose weight. According to an analysis that appeared in iScience in October 2021, regardless of weight, higher cardiorespiratory fitness is linked to better health outcomes and a lower risk of early death, even if increased exercise rarely results in long-term weight loss.

(ii) Muscle Mass: Here's where things get a little trickier. If two people have the same weight but have different amounts of muscle, the person with greater muscle will burn more calories. According to Jenaed Brodell, RD, a specialist sports nutritionist with a private practise in London, "Muscle tissue burns more calories than fat tissue." However, estimates of how many calories a pound of muscle will burn are sometimes grossly exaggerated.

According to data, a pound of muscle burns roughly five calories per day, compared to two calories per day for a pound of fat. Because your body needs more energy to support the greater pace at which your muscles are contracting during exercise, having more muscle mass will raise your total calorie burn. [4] To cut a long tale short, think about boosting up your strength-training routine if you want to increase your calorie burn. According to Brodell, "Evidence shows that lifting weights burns more fat [than cardio exercise] and has more promising long-term results." However, she adds that because everyone has different goals and physical limitations, it is ultimately up to you to decide how you exercise.

(iii) Birth Sex: According to Gonzalez, males generally burn more calories both at rest and during exercising than women do. But there's nothing mystical about the reason for this; it's just that men are typically bigger and have more muscle mass than women of the same age and weight. At rest, men typically burn 5 to 10 percent more calories than women do, and this difference typically rises with exercise, according to Gonzalez. Additionally, despite the fact that women can definitely increase their muscular mass through strength training, physiological distinctions prevent them from generally being as thin as males.

According to Brodell, women are genetically prone to store more fat in order to promote hormone synthesis and pregnancy. [5] Body fat is also necessary for storing energy, safeguarding internal organs, and supporting vital processes including growth, immunity, hormone synthesis, reproduction, and metabolism, according to the National Institute of General Medical Sciences. Men need between 2 and 5 percent of body fat, while women need between 10 and 13 percent, according to the American Council on Exercise (ACE). However, these bare minimums might not be adequate. A healthy range for persons under the age of 40 is 8 to 20 percent for males and 21 to 33 percent for women, according to the most often referenced study on the subject, even though there is no official recommendation for the ideal body fat percentage. Nevertheless, the connection between health and body fat is nuanced and poorly understood. Focus on the things you can control rather than worrying about how your birth sex influences your calorie burn. The final line, according to Gonzalez, is that both men and women should concentrate on increasing their muscle mass and cardiovascular fitness through a well-balanced cardio and strength-training regimen.

(iv) Age: According to Brodell, "as we age, we tend to lose muscular mass." You can start losing up to 3 to 5 percent of your muscle mass every decade after the age of 30. its unclear why this happens, but a review article from July 2017 in Ageing Research Reviews suggests that it's probably because your body becomes less sensitive to hormones that support protein synthesis, which is essential for maintaining muscle mass. Your metabolic rate, or how quickly you burn calories both at rest and while exercising, decreases as a result of this loss of muscle mass. According to a study on human metabolism, which was published in the August 2021 issue of Science, metabolic rate may not fall throughout adulthood instead plateauing between the ages of 20 and 60 before starting to decline.

Using the doubly labelled water approach, the industry standard for this kind of assessment, researchers determined the energy expenditure of 6,421 men and women aged 8 days to 95 years. However, this does not always imply that everyone's calorie expenditure remains constant as they age. According to Rosen, who also notes that it's impossible to precisely quantify metabolism, the study on daily energy expenditure over the span of a human life is "interesting, but it is not necessarily definitive proof that our concept of metabolism is erroneous." Additionally, ageing brings about a number of physiological changes in humans, not all of which may have been taken into account, even by trained eyes. While you can't stop your body from ageing, Gonzalez claims that regular strength training can help you maintain or even grow your muscle mass. Strength training can help you raise your resting metabolic rate, which gradually increases the amount of calories you burn when at rest.

(v) Fitness Level: A particular workout becomes easier the more often you perform it. That's not in your imagination, Gonzalez says; given time, your body does adjust to perform tasks more readily. This is a wonderful thing all around. It implies that

with practise, you may run farther or faster, and that with the right kind of exercise, your muscles would be able to lift more weight. But it also influences how many calories you burn. Gonzalez asserts that as your body adjusts to exercise, you will burn fewer calories while still exercising. Your body becomes more efficient as you get fitter, from your lungs to your muscles to your heart to your brain. Because of this, a beginner may burn much more calories than someone who has been performing the same exercise for years. Changing your exercise regimen can also improve your fitness level and calorie burn.

(vi) Training Intensity: It's also feasible that two people who perform the same exercise differently from one another in terms of calorie expenditure. According to Brodell, someone exercising at a high intensity—which means they are breathing deeply and are unable to carry on a conversation—can burn twice as many calories as someone exercising at a low level in the same length of time. Additionally, just because you are travelling the same distance or performing the same exercises as someone else does not imply that you are exerting yourself to the same degree.

According to the Department of Health and Human Services (DHHS), running and walking both lower blood pressure and lower your chance of developing chronic diseases including heart disease and type 2 diabetes. However, a prior study indicated that walking one mile for adults burnt about 89 calories. According to the DHHS, 150 minutes of low-intensity exercise each week is sufficient to produce a variety of health advantages, such as lowered anxiety, improved sleep, lowered blood pressure, greater cardiovascular fitness, and a decreased risk or slower progression of several chronic illnesses. Adding more intense exercise to your schedule will increase your calorie burn and enhance these advantages. The ACE advises raising your speed, range of motion, or weight when performing strength-training exercises to increase the intensity of your workouts.

II. LITERATURE SURVEY

According to Daniel Bubnis [6], the variety of energy consumed during the day has a direct impact on whether a person maintains, gains, or loses weight. A calorie deficit results when people consume less calories than they burn. However, they are curious about their daily caloric expenditure. The majority of people believe that eating and weight loss are most closely related to calories. Different definitions of calories exist as units of energy or heat. Knowing how many calories one is taking in each day is crucial for anyone wanting to gain, reduce, or maintain weight.

Sánchez Camacho (9 No country has yet been able to reverse the global obesity issue, which has been steadily getting worse. The primary contributing factor to obesity, according to the World Health Organization, is an energy imbalance between calories consumed and calories expended. However, accumulating evidence suggests that the concept of calorie imbalance may not be sufficient to curb and end the obesity epidemic.

In order to highlight the need for an updated theory regarding the causes of obesity, it is necessary to investigate the calorie imbalance hypothesis and its components as a weight-management tool, as well as any potential downsides and implications for public health. By preventing weight gain or promoting weight loss, this change may help public health programmes to combat obesity.

World Health Organization. Obesity Study. [Online] (2011, October). [10] Someone who wants to change their diet or exercise regimen may find it helpful to understand the elements that affect calorie burning. Numerous studies have been published in the literature that use data mining and machine learning to identify these issues. Some publications written two to three years ago have a poorer accuracy for estimating calories burnt issues when compared to today's study.

III. PROPOSED METHOD

This study focuses on acquiring the appropriate data set to train our machine learning models in order to estimate how many calories an individual would burn. Prior to the statistics feeding operation, the records must be pre-processed. Following the completion of data processing, the data is organised as plots and graphs using a variety of visualisation approaches. Here, we compare and evaluate these models using a linear regressor model from machine learning.

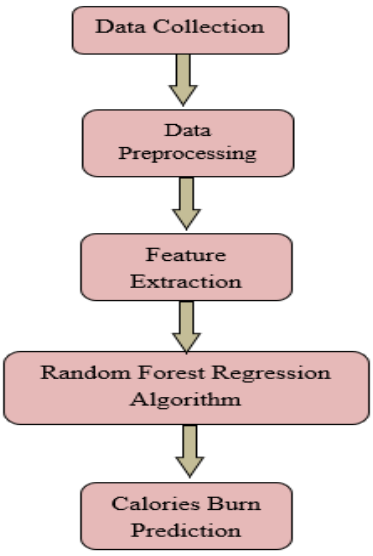

Fig 1: Block Diagram of proposed method

Random Forest is an ensemble methodology that can handle both regression and classification tasks, using a number of decision trees and a technique called Bootstrap and Aggregation, commonly known as bagging. The key tenet of this approach is to combine multiple decision trees to obtain the outcome rather than relying simply on one decision tree. The core learning models of Random Forest are multiple decision trees. By selecting rows and attributes from the dataset at random, we produce sample datasets for each model. This part is referred to as Bootstrap.

Similar to other machine learning techniques, the Random Forest regression method must be handled. Create a specific query or group of data, then request the required information from the source. Make that the data is in an accessible format; if not, convert it to the required format. List any glaring anomalies and any information that might be required in order to achieve the appropriate data. Create a machine learning model. Choose the baseline model you want to achieve. Utilize the data to train the machine learning model. To understand the model, use the test data. Now compare the performance indicators between the test data and the projected data from the model. If it falls short of your expectations, you can try changing your model accordingly, dating your data, or utilising another data modelling technique.

Code:

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

exercise = pd.read_csv(r'C:\Users\Lenovo\Desktop\KK\exercise.csv')
calories = pd.read_csv(r'C:\Users\Lenovo\Desktop\KK\calories.csv')
exercise.head()
```

	User_ID	Gender	Age	Height	Weight	Duration
0	14733363	male	68	190.0	94.0	29.0
1	14861698	female	20	166.0	60.0	14.0
2	11179863	male	69	179.0	79.0	5.0
3	16180408	female	34	179.0	71.0	13.0
4	17771927	female	27	154.0	58.0	10.0

```
print(calories)

       User_ID  Calories
0      14733363     231.0
1      14861698      66.0
2      11179863      26.0
3      16180408      71.0
4      17771927      35.0
...         ...       ...
14995  15644082      45.0
14996  17212577      23.0
14997  17271188      75.0
14998  18643037      11.0
14999  11751526      98.0

[15000 rows x 2 columns]
df = pd.merge(exercise, calories, on = 'User_ID')
print(df)

       User_ID  Gender  Age
Height  Weight  Duration
Heart_Rate  Body_Temp  \
0      14733363    male   68
190.0    94.0      29.0
105.0       40.8
1      14861698  female   20
166.0    60.0      14.0
94.0       40.3
2      11179863    male   69
179.0    79.0       5.0
88.0       38.7
3      16180408  female   34
179.0    71.0      13.0
100.0       40.5
4      17771927  female   27
154.0    58.0      10.0
81.0       39.8
...         ...     ...  ...
...     ...       ...          ...
...
14995  15644082  female   20
193.0    86.0      11.0
92.0       40.4
14996  17212577  female   27
165.0    65.0       6.0
85.0       39.2
14997  17271188  female   43
159.0    58.0      16.0
90.0       40.1
14998  18643037    male   78
193.0    97.0       2.0
84.0       38.3
14999  11751526    male   63
173.0    79.0      18.0
92.0       40.5

       Calories
0         231.0
1          66.0
2          26.0
3          71.0
4          35.0
...         ...
```

```
14995       45.0
14996       23.0
14997       75.0
14998       11.0
14999       98.0

[15000 rows x 9 columns]
```

```
x = df.drop(columns=['User_ID','Calories'])
y = df['Calories']

from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test = train_test_split(x,y,test_size=0.2,random_state=5)

from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
X_train['Gender'] = le.fit_transform(X_train['Gender'])
X_test['Gender'] = le.fit_transform(X_test['Gender'])
X_test
```

Gender	Age	Hgt	Wgt	Dura	Heart_ Rate	Body Temp
	1	64	180.0	82.0	12.0	96.0
13951	1	39	194.0	93.0	18.0	89.0
8082	0	25	147.0	47.0	25.0	108.0
10284	0	72	146.0	49.0	22.0	103.0
13473	0	45	162.0	67.0	17.0	99.0
...	...	...	...	...	...	...
1040	0	53	164.0	65.0	13.0	98.0
669	0	28	172.0	68.0	3.0	85.0
2729	1	20	179.0	75.0	17.0	109.0
9858	1	26	174.0	77.0	9.0	88.0
6090	0	61	150.0	55.0	2.0	87.0

3000 rows × 7 columns

```
from sklearn.linear_model import LinearRegression
lr = LinearRegression()
lr.fit(X_train,y_train)
lr.score(X_test,y_test)

0.9674975501970082

from sklearn.ensemble import RandomForestRegressor
rf = RandomForestRegressor()
rf.fit(X_train,y_train)
rf.score(X_test,y_test)

0.997911666203642

import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
dataset = pd.read_csv(r'C:\Users\Lenovo\Desktop\KK\exer.csv')
X = dataset.iloc[:, 1:2].values
y = dataset.iloc[:, 2].values
from sklearn.ensemble import RandomForestRegressor
regressor = RandomForestRegressor(n_estimators=100, random_state=0)
regressor.fit(X,y)
X_grid = np.arange(min(X), max(X), 0.01)
X_grid = X_grid.reshape((len(X_grid), 1))
plt.scatter(X, y, color = 'red')
plt.plot(X_grid, regressor.predict(X_grid), color = 'blue')
plt.title('Random Forest Regression Model')
plt.xlabel('Duration')
plt.ylabel('Calories Burned')
plt.show()
```

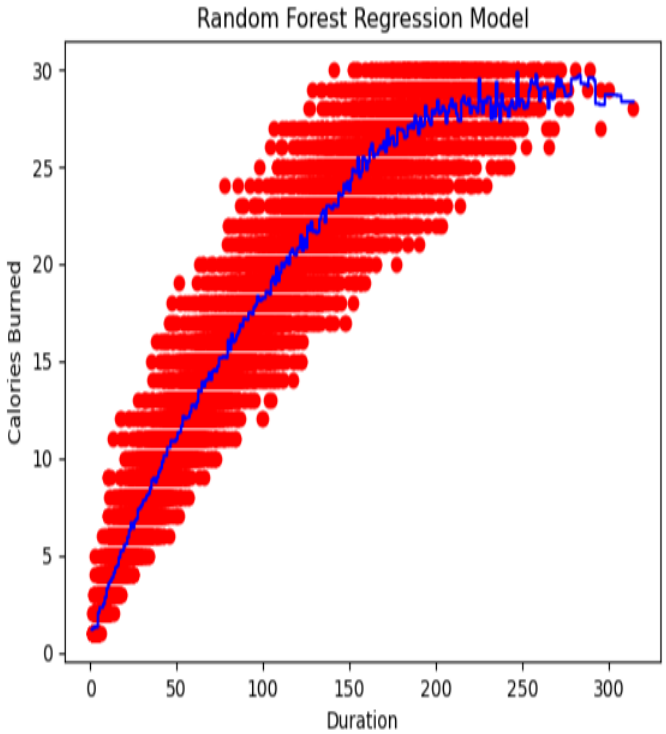

Fig 2: Duration Vs Calories Burned

IV. CONCLUSION

Random Forest is an ensemble methodology that can handle both regression and

classification tasks, using a number of decision trees and a technique called Bootstrap and Aggregation, commonly known as bagging. The key tenet of this approach is to combine multiple decision trees to obtain the outcome rather than relying simply on one decision tree. The core learning models of Random Forest are multiple decision trees. By selecting rows and attributes from the dataset at random, we produce sample datasets for each model. This part is referred to as Bootstrap.

Similar to other machine learning techniques, the Random Forest regression method must be handled. Create a specific query or group of data, then request the required information from the source. Make that the data is in an accessible format; if not, convert it to the required format. List any glaring anomalies and any information that might be required in order to achieve the appropriate data. Create a machine learning model. Choose the baseline model you want to achieve. Utilize the data to train the machine learning model. To understand the model, use the test data. Now compare the performance indicators between the test data and the projected data from the model. If it falls short of your expectations, you can try changing your model accordingly, dating your data, or utilising another data modelling technique.

V. REFERENCES

[1] https://www.kaggle.com/fmendes/fmendesdat263xdemos
[2] https://machinelearningmastery.com/linear-regression- for-machine-learning/
[3]https://machinelearningmastery.com/xgboost-for-regression/
[4]https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5496172
[5] https://devhadvani.github.io/calorie.html
[6] https://www.medicalnewstoday.com/articles/319731#factors-influencing-daily-calorie-burn-and-weight-loss
[7] https://zenodo.org/record/6365018
[8] https://devhadvani.github.io/calorie.htm
[9] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5496172
[10] https://www.jetpac.com/
[11] World Health Organization. (2011, October) Obesity Study. [Online] http://www.who.int/mediacentre/factsheets /fs311/en/index.html.