e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 7, July 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707246 |

Calorie Burn Prediction using Machine Learning

Sridhar. S[1], Shweatha B M[2] ,Snowlin A[3]

Student, Department of Computer Science and Engineering, R.M.D. Engineering College (Autonomous), India[1,2,3]

ABSTRACT: The overall thought of this exploration project is to make a near investigation of AI calculations to foresee the calories consume during the exercise. In this paper we first form an AI frameworks that can foresee how much calories consumed during exercise. In this day and age many individuals are curious about the exercise that they do and the weight reduction plan that they take and how much calorie do they consume once they exercise. We can use ML algorithms like linear regression and the XGBoost regressor to solve this issue.

KEYWORDS: Colab, XGBoost Regressor, Linearregression, machine learning, accuracy

I. INTRODUCTION

When we exercise or work out, our heart rate and body temperature will rise. The factors that we require some investment scale for which the individual doing the exercise preparing and what is the typical beats each moment and afterward the temperature. Then, we use the person's age, gender, weight, and height to estimate how much energy they may be using.
A linear regression algorithm and a machine learning XGBoost regressor are used to predict calories burned in relation to a person's age, height, weight, workout duration, and body temperature.

DESCRIPTION

The assortment of energy consumed every day is promptly associated with weight reduction, weight gain, or weight upkeep. To shed pounds, an individual should consume more noteworthy calories than they take in, fostering a calorie deficiency. But before they can do that, they want to know how many calories they burn each day. The majority of people consider calories to be the most effective factor when it comes to food and weight loss. A unit of heat or energy that can be referred to in a variety of ways. The amount of energy required to raise one gram (g)of water by one degree Celsius is known as a calorie. This measurement can be carried out to lots of different strength releasing mechanisms outdoor of the human body. Inthe case of human body, calories are measure of how much energy the body requires to function. Any man or woman attempting to preserve, lose, or maintain weight needs to be able to exercise how many calories are burned each day. A person can control their diet or exercise routine to achieve the goal by knowing what factors help burn calories. There are many variables that influence how much calories an individual consumes every day. a portion of the components that impact step by step calorie consume aren't in an individual's oversee simultaneously as others might be changed.

Key facts

- It is now commonly recognized that a diet high in calories and a lack of physical activity lead to obesity or overweight, which in turn can lead to conditions like noninsulin-dependent diabetes, hypertension, cardiovascular disease, endometrial cancer, and gallstones.
- Raised BMI is a significant trepidation component for noncommunicable wiped out comprising of cardiovascular illnesses (extraordinarily coronaryheart ailment and stroke), that have been the rule reason for kicking the bucket in 2012-diabetes, outer muscle disorders(in specific osteoarthritis - an as a substitute debilitating degenerative joint issues.

Machine Learning

One branch of artificial intelligence known as machine learning emerges as software programmes get better at making predictions without being specifically trained to do so. These calculations make use of verifiable information to predict the new outcome values. With machine learning, a user can provide an algorithm with a lot of data, and the algorithm will review it and make data-driven recommendations and judgements based only on the data it has received.

II.PURPOSE OF THE STUDY

The goal of this article is to compare the two machine learning algorithms and forecast the number of calories that different people will burn while exercising using the data sets. The study's dataset consists of 7 elements, 1 objective

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707246 |

variable, and 15,000 occurrences. These data sets are being used to train a dataset, choose the best model, pinpoint the exact methods, and calculatetheir mean absolute error.

III.METHODOLOGY

The goal of this study is to find the best set to teach our machine learning models so that they can calculate the number of calories an individual burns while exercising. Prior to handling methods, measurements made through records should be handled. After that, we analyse the data and plot and graph it using a variety of visualisation tools. Then, separate the data set into a training set and a test set. In this study, we use the XGBoost regressor and straight relapse to analyse and assess AI models. The tool being used, Google Colaboratory, is botha cloud-based service and a web-based tool.Colab. Colab, an acronym for "Colaboratory," is the name of a Google Research product that runs fully on the cloud. Colab is ideal for algorithms, data analysis, and machine learning because it enables us to run Python code in the browser. Technically speaking, Colab is a Jupyter notebook that offers ahosted service that doesn't need to be installed and gives free access to computational resources with GPUs.

Colab can fully utilise well-known Python libraries for dataanalysis and visualisation. In only a few lines of code, you can use Colab to import a dataset of images, train a picture classifier on it, and view the model.Google's servers run code from Colab notebooks.

Cloud servers, which means you have an advantage of Google hardware, as well as GPUs and TPUs, nevertheless ofthe power of your machine. All you need is a browser. A programmer can perform the using Google Colab.We can write and execute code in Python in colab.

- Create, upload, and share notebooks
- Import and save notebooks from or to Google Drive
- Import or Publish notebooks from GitHub
- Import external datasets, such as those from Kaggle
- Integrate PyTorch, TensorFlow, Keras, and OpenCV
- Free Cloud service with free

IV. IMPLEMENTATION

A useful model for predicting the number of calories burned during exercise based on a person's age, gender, height, and weight as well as the duration of the workout has been constructed using two distinct algorithms on the given data set.

Algorithms

RGBoost Regressor - The RGBoost Regressor is an investigation that is a true way to show the relationship between a dependent (target) and free (indicator) factors with at least one neutral element. The XGBoost algorithm works well in machine learning due to its robust handling ofa wide variety of data kinds, relationships, distributions, and the multiple hyperparameters that may be adjusted. The XGBoost regressor can be used for regression, binary and multiclass classification, and ranking problems.

Linear Regression – A supervised learning algorithm in machine learning is linear regression. This condition consolidates a specific arrangement of information values (x)to which is the anticipated outcome for that arrangement of information values (y) and uses the direct condition to identify the relationship between one or more indicator factors and one result variable. The equation for the issue is $Y = a + b X$.

Steps:

A. Collect Dataset
B Data Pre-processingC Data Analysis
D. Machine learning model
E. Evaluation.

Data Source

The dataset was stored in a repository called Kaggle. There are two csv files, each containing 15,000 instances and

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707246 |

7 attributes. The Kaggle repository's data set contains information on each person's gender, age, workout time, heart rate, body temperature, height, and weight. We'll use this dataset as our preparatory data. The second calories dataset also includes a target class for each person's calories burned.

Table 1: Attributes and its values

Input Attributes	Function
gender	gender (male : 0,female : 1)
age	age mentioned in years
height	height of the person
weight	weight of the person
duration	The time taken to complete theexercising in minutes.
heart_rate	Average heart rate during the workout(more than normal rate 75 beats/min)
body_temp	Body temperature in the course of the workout(greater than 37 degree celsius)
calories	Total calories burned during theexercise.

There are two dataset csv files which should be uploaded to colab which is used for processing. We use data frames for analysis and processing. It obtain some statistical measures about the data.

```
[ ] calories_data.describe()
```

	User_ID	Age	Height	Weight	Duration	Heart_Rate	Body_Temp	Calories
count	1.500000e+04	15000.000000	15000.000000	15000.000000	15000.000000	15000.000000	15000.000000	15000.000000
mean	1.497736e+07	42.789800	174.465133	74.966867	15.530600	95.518533	40.025453	89.539533
std	2.872851e+06	16.980264	14.258114	15.035657	8.319203	9.583328	0.779230	62.456978
min	1.000116e+07	20.000000	123.000000	36.000000	1.000000	67.000000	37.100000	1.000000
25%	1.247419e+07	28.000000	164.000000	63.000000	8.000000	88.000000	39.600000	35.000000
50%	1.499728e+07	39.000000	175.000000	74.000000	16.000000	96.000000	40.200000	79.000000
75%	1.744928e+07	56.000000	185.000000	87.000000	23.000000	103.000000	40.600000	138.000000
max	1.999965e+07	79.000000	222.000000	132.000000	30.000000	128.000000	41.500000	314.000000

Figure 1- dataset

The average age in this instance is 42.7, and we also know the standard deviation and percentile range. There is about 40 degrees of internal heat.Those who are exercising will have higher levels of internal heat. The most significant findings from this study relate to heart rate and body temperature. The data must then be displayed using a few charts and graphs.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707246 |

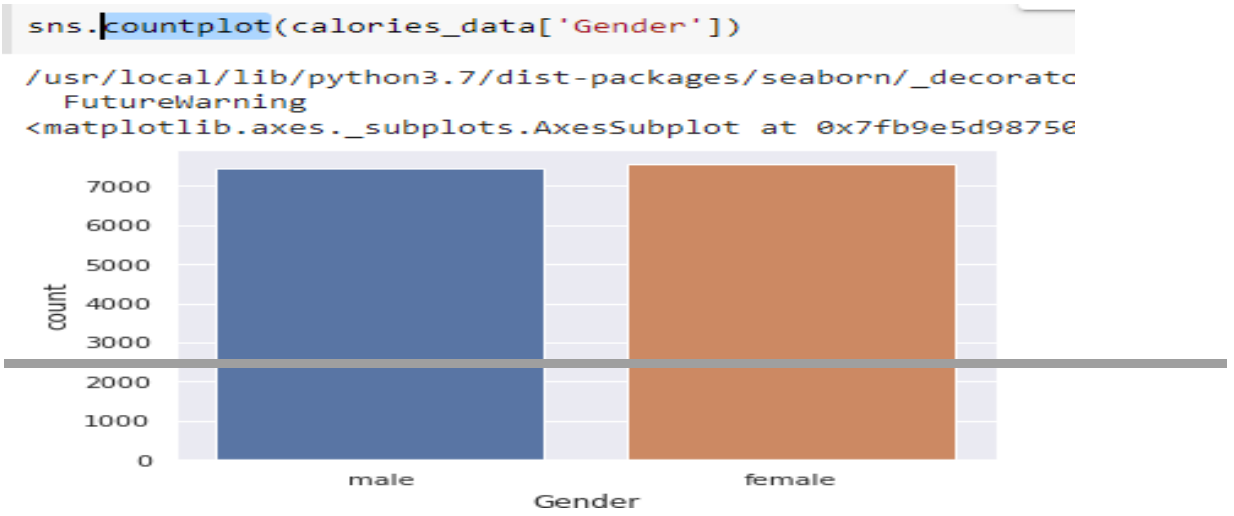

Figure 2- gender_plot

we have almost good identical distribution of maleandfemale data points.

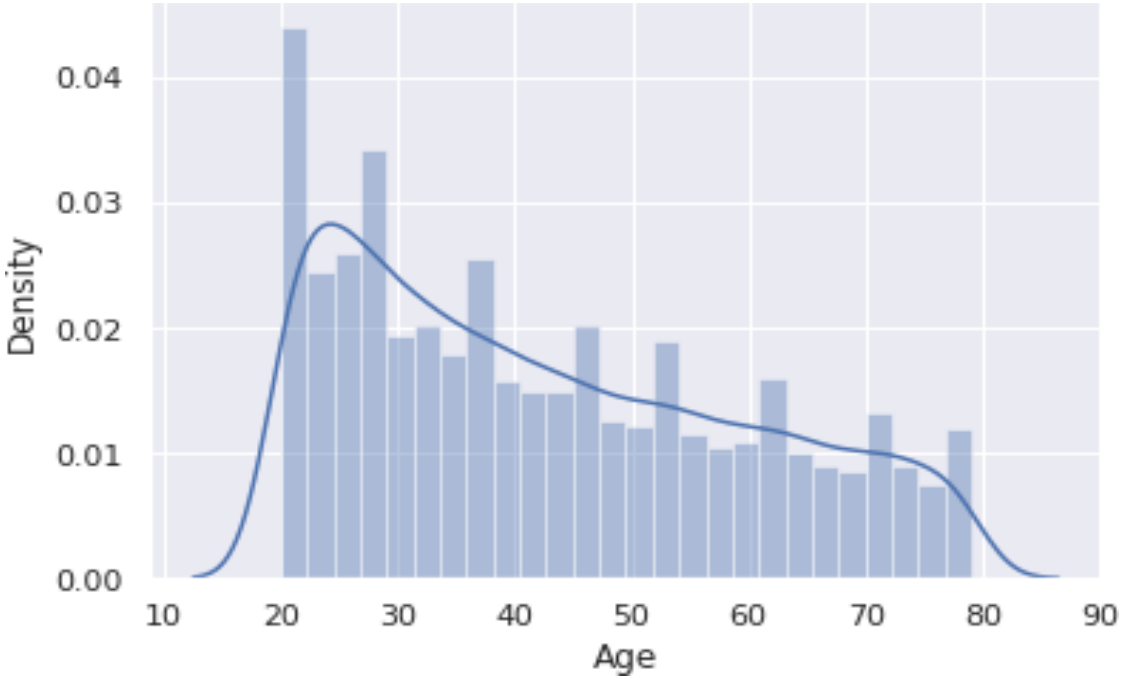

Figure 3 - age_plot

There are 15000 more people in this age group based on the number of values between 20 and 30 where the curve shows a peak. The second factor is a decline in the number of senior citizens who exercise.The following step is to look into the relationships between the different records. Positive and negative correlations are the two types that exist. If the workoutis done for a longer period of time than anticipated, the numberof calories burned will also increase. Therefore, since these values are directly proportional, there is no question that they connected

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707246 |

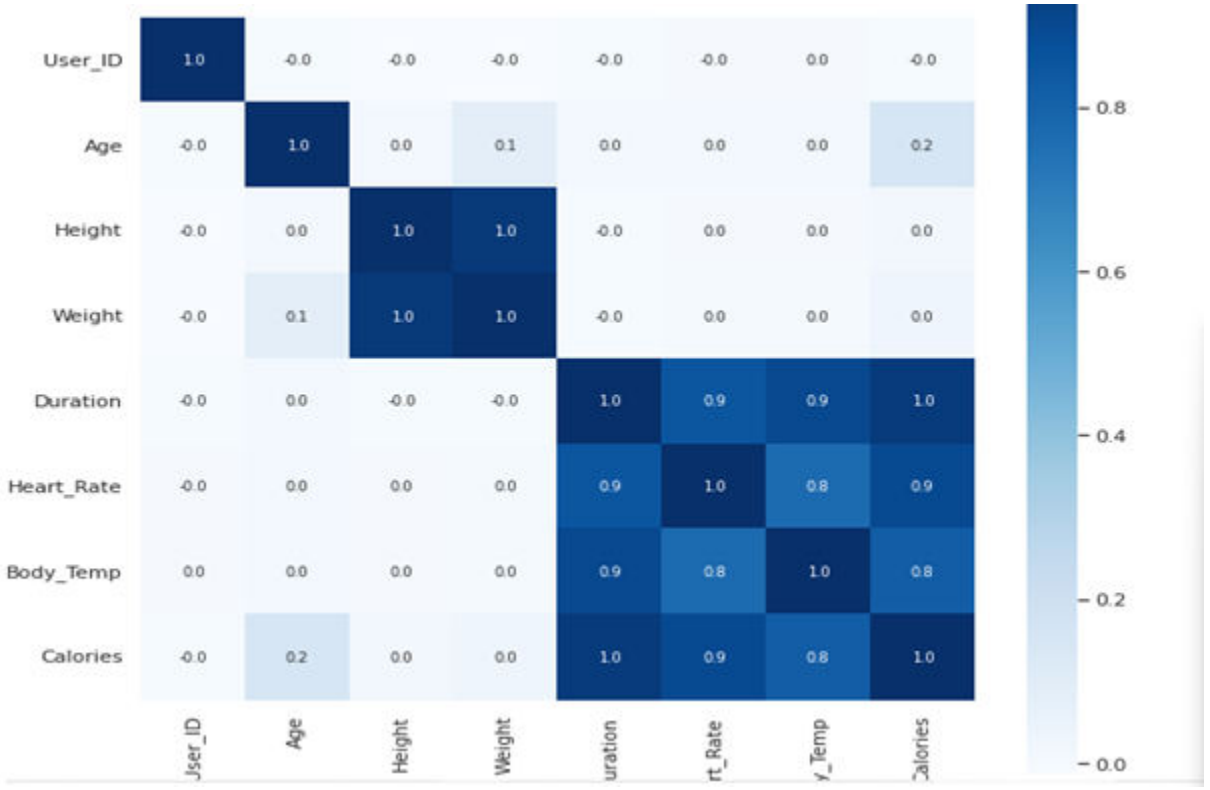

Figure 4- correlation_heatmap

This demonstrates a strong positive correlation between calories, heart rate, and temperature, as well as a positive correlation between weight and duration. Divide the data intotraining and test sets, using the two variables X and Y that must be distinguished for features and the target.

```
[ ] X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=2)
```

Then, use the test data to test the prediction using the loaded XGB Regressor model. The model analyses the test results and calories burned for the X_test. Additionally, have a look at the special features that our model predicts will be present.

The magnitude of the version's errors is calculated using metrics.mean_absolute_error(Y_test, test_data_prediction). The mean absolute error getting is 2.71.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707246 |

```
input_data = (0,68,190.0,94.0,29.0,105.0,40.8)

input_data_as_numpy_array = np.asarray(input_data)

input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)

prediction = model.predict(input_data_reshaped)
print(prediction)

[230.33057]
```

Figure 5 – XGBoost regressor output

Next we do this in linear regression

```
input_data = (0,68,190.0,94.0,29.0,105.0,40.8)

# change the input data to a numpy array
input_data_as_numpy_array= np.asarray(input_data)

# reshape the numpy array as we are predicting for only on instance
input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)

prediction = lin_reg_model.predict(input_data_reshaped)
print(prediction)

[199.38038663]
```

Figure 6 – Linear regression output

lin_reg_model = LinearRegression()The mean absolute error getting is 8.38.

V. RESULT

The analysis of this dataset was utilised to forecast the number of calories burned based on the gender, age, body temperature, and heart rate at different moments during the activity. By using these techniques, we search for a machine learning model that has a smaller mean absolute error and yields more precise results. When comparing the two models—XGB regressor and linear regression—the XGB regressor performs better than the linear regression in terms of accuracy, with a mean absolute error of 2.71.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707246 |

Machine LearningModel	Mean Absolute Error
XGB Regressor	2.71
Linear Regression	8.38

Table – 2

Machine LearningModel	Input data	Predicted Calorie result	Expected Calorie result
XGB Regressor	Male{0},68,190.0,94.0,29.0,105.0,40.8	230.33	231.0
Linear Regression	Male{0},68,190.0,94.0,29.0,105.0,40.8	199.380	231.0

Table - 3

VI.CONCLUSION

We came to the conclusion from the analysis that the XGB Regressor generates more accurate findings than the Linear regression model. The difference between the values obtained from the models' predictions and the actual values is known as mean absolute error. It ought to be as low as is practical. The mean absolute error number for the XGB Regressor, 2.71, is acceptable. The mistake rates are really small. In light of this, we can say that the XGBoost Regressor is the most accurate model for estimating calorie burn.

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

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com