­CALORIES BURNED PREDICTION

Team Data:

Alle Mohan Venkateswara Rao

Kalyana Tirupathi Rao

MummidiVarapu Krishna Vamsi

Veera Venkata naga sai Prakash

Acahnta Harsha

Kondeti Sai Bhaskar

STATEMENT:

The amount of calories burnt depends on internal and external factors, it is subjective and different for everyone

depending on their height, weight, and fitness level

The amount of calories burnt depends on internal and external factors, it is subjective and different for everyone

depending on their height, weight, and fitness level

The amount of calories burnt depends on internal and external factors, it is subjective and different for everyone

depending on their height, weight, and fitness level

The amount of calories burnt depends on internal and external factors, it is subjective and different for everyone

depending on their height, weight, and fitness level

The amount of calories burnt depends on internal and external factors, it is subjective and different for everyone depending on their height, Weight, and fitness level.

INTRODUCTION:

Usually, when people think of calories, they only think of food or weight loss. However, a calorie is usually a measure of heat energy. Calories are the units of energy required to raise 1 gram (g) of water by 1°C. The measurement can be used to evaluate many energy-releasing systems unrelated to the human body The amount of energy required by the body to perform a task is the number of calories considered from the point of view of the human body. There are calories in food. Each dish contains a distinct amount of energy. Body temperature and heart rate will start to rise when we exercise or exercise hard. Carbohydrates or carbohydrates are broken down into glucose which is then converted/broken down into energy using O2 (oxygen). The variables used here are the time scale a person exercises, average heart rate per minute, and temperature. Then add the person's height, weight, gender, and age to predict how much energy that person is burning. Parameters that can be taken into account are exercise time, average heart rate per minute, temperature, height, weight and gender. The XG Boost machine learning regression algorithm is used to predict calories burned based on exercise time, temperature, height, weight, and age.

ALGORITHM:

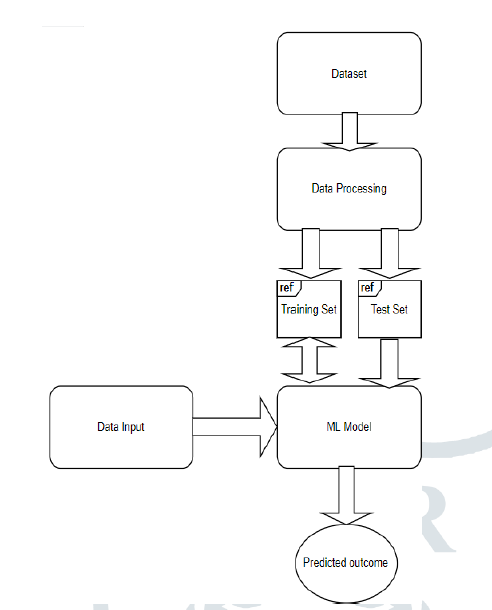
**Linear Regression**

Linear regression is a fundamental machine learning algorithm used for predicting a continuous target variable based on one or more input features. It assumes that there is a linear relationship between the input features and the target variable. In other words, it tries to find the best-fitting linear equation that describes the relationship between the variables

**METHODOLOGY**

To determine how many calories an individual will burn, this study involved gathering the right data set to train our machine learning models. Pre-processing of records is necessary before performing the operation that provides statistics. After that, the data processing is complete and the data is organized into diagrams and graphs using a number of visualization techniques. Here, we use the XG Boost regressor as the ML (machine learning) model to compare and then evaluate these models.

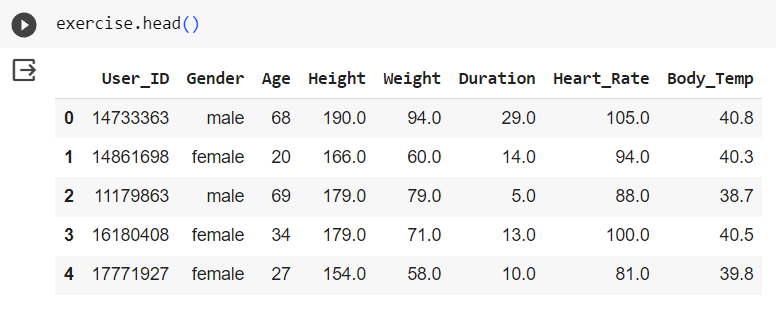
**Work Flow:**

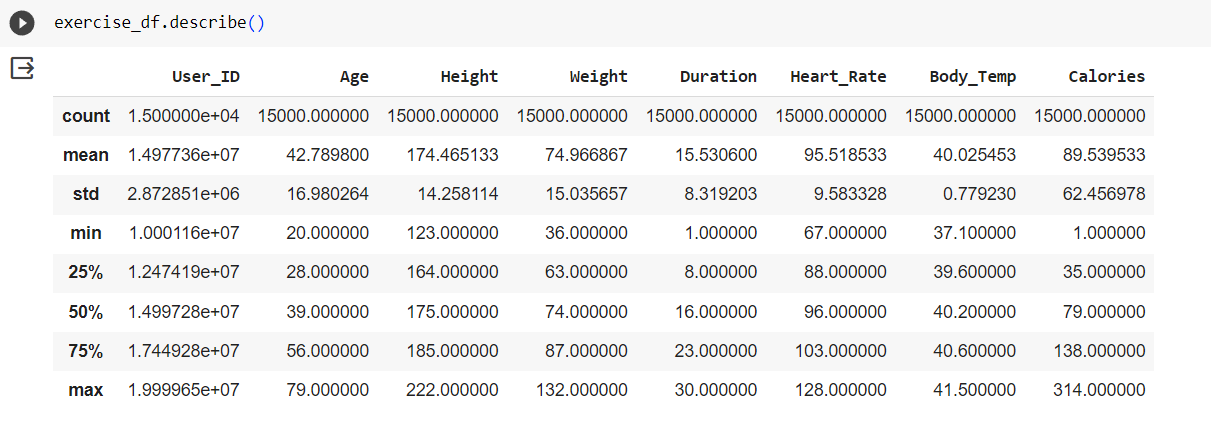


**Data sourse:**

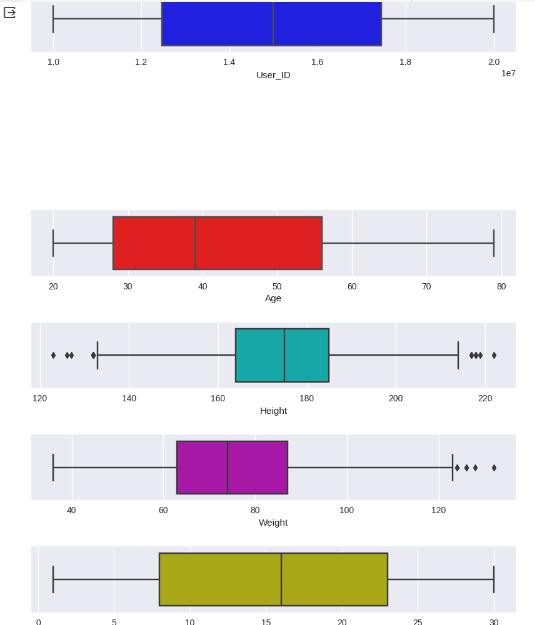
There are a total of 15,000 instances and 7 data attributes in 2 CSV files. The "Kaggle" archive dataset includes information about a variety of people, including their height, weight, gender, age, exercise intensity, heart rate, and body temperature. Exercise data is obtained from the "exercise.csv" and "calories.csv" datasets. In addition, the target class mapped by the user ID from the second calorie dataset includes the calories that person burned in the exercise dataset.

|  |  |
| --- | --- |
| **Attribute** | **Function** |
| Gender\_ individual | Gender (female : 1, male : 0) |
| Age \_Individual | Age in years |
| Height\_ Individual | Height of a person |
| Weight \_Individual | Weight of a person |
| Heart\_ rate \_Individual | Average heart rate of an individual during exercise (normal heart rate 75 beats/min) |
| Body \_temp \_individual | Average body temperature recorded over the entire course workout (above 37 degrees Celsius) |
| Duration \_individual | Training time in minutes. |
| Calories \_individual | The total amount of calories burned while workout. |

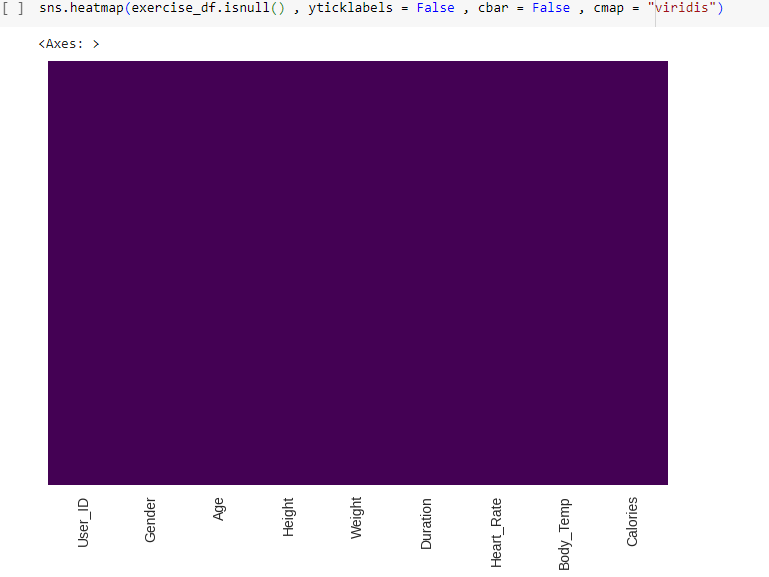




* As we can see, the table above shows the Descriptive Statistics(for example  centeral tendency) of each column or feature.
* For example for Age column.%25 of the data lie between **20** and **28**, anohter %25 lie between **28** and **39**, and so on .The box plot shows the exact concept that I just mentioned.
* The outliers are shown with dots in box plots, which we will discuss about them in the next section.

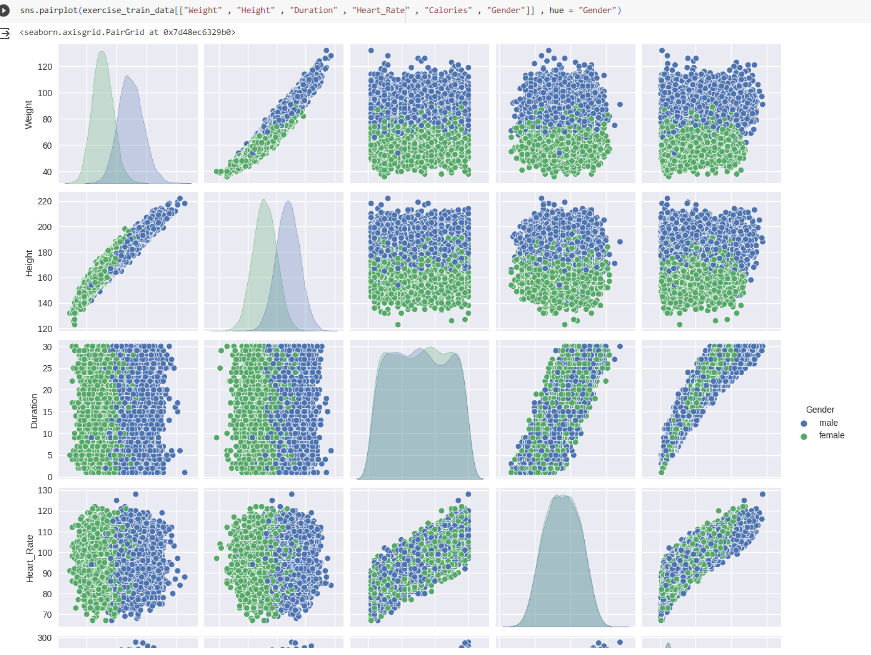


Null Values



### Dataset's Distribution

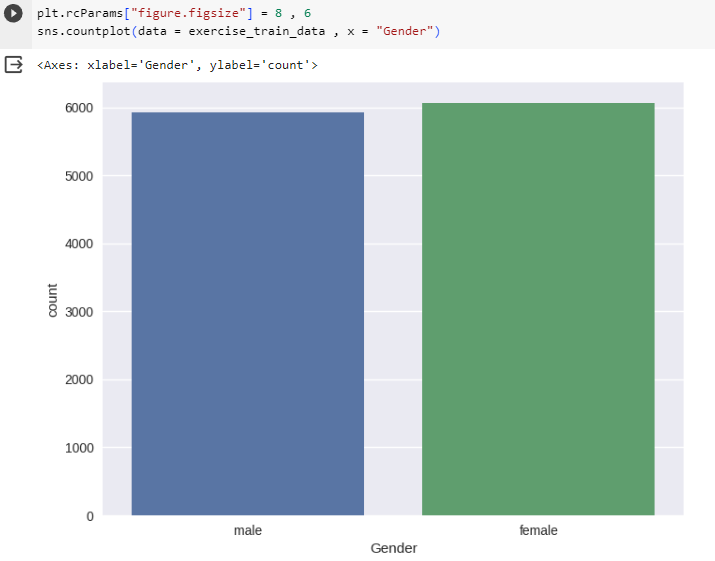
### 



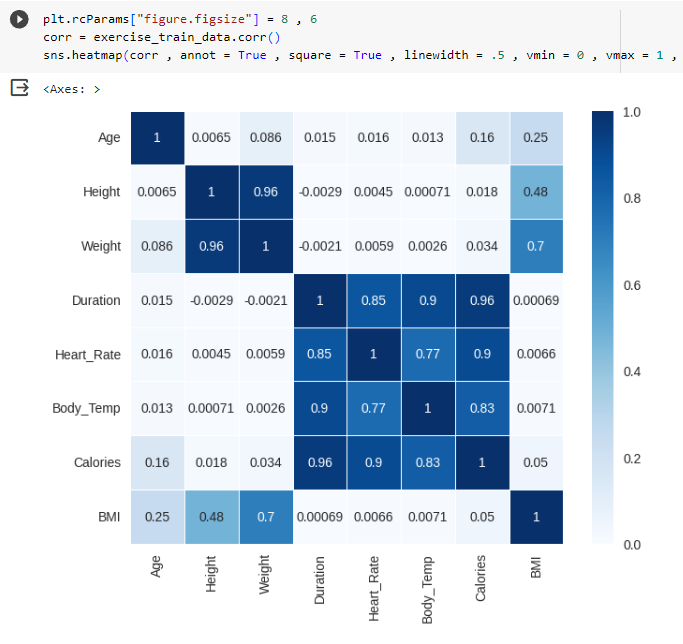
### Exploratory Data Analysis(EDA)

### 

GENDER



CORELATION



A.Data collection

Data retrieval is the first step. Kaggle is the data store we use. It is loaded into the Collab program. The information collected is both categorical and numerical.

B.Data Preprocessing

15,000 instances and 7 data attributes are contained in two csv files ("exercise.csv" and "calorie.csv."

Each person's attributes are included in the Kaggle data collection.

Including their size , weight, gender, age, exercise duration, heart rate, and body temperature.

Data preprocessing is an important step in the machine learning process because the quality of the data and the insights that can be extracted from that data directly affect the trainability of the model ours. It is important that we preprocess our data before providing it to our model as output.

C. Data Analysis

Colab, the platform used for the processing, requires the upload of two dataset csv files ("exercise.csv" and "calorie.csv"). The average body temperature is 40. People who exercise will have a higher body temperature. Heart rate and coronary temperature were the most important results of this analysis. The data is then visualized using a few tables and charts. Two types of correlation, positive and negative, were then studied between different records. Then load the XGB Regressor model and evaluate the prediction using the test data. This test data and the calories burned for the X test are run in the model. Similarly, compare our model's expected values with the original values.

D. Machine Learning model

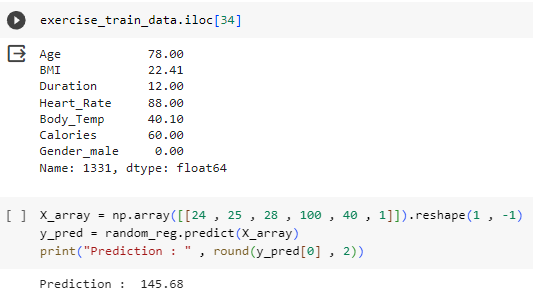
This is the step where we apply our chosen algorithm (in this case, the XGBoost regressor) to determine the mean absolute error. The XGB regression procedure was used and the results obtained. For this, we use indicators that indicate the level of errors of the version

The XGBoost regression algorithm has been proven to be an effective and efficient method for predicting calories burned.

E. Evaluation

This dataset was analyzed to make predictions about how many calories were burned based on exercise duration as well as factors such as age, gender, body temperature and heart rate at different time points. different points during exercise. We are looking for a machine learning model with lower mean absolute error that produces more accurate results using these machine learning methods.

RESULT:



**CONCLUSION:**

The XG Boost Regressor model performed exceptionally well, with a mean absolute error of 2.71, indicating highly accurate predictions of calories burned. The model considered seven main factors influencing calorie expenditure but acknowledged the presence of other influencing variables. The study emphasized the importance of understanding both calorie consumption and expenditure for maintaining health. The suggestion to develop a user-friendly interface for input and result display was made, with potential plans for a comprehensive app integrating diet and exercise recommendations based on the predictions