ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 9| Issue: 12| December 2023|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2023: 8.224 || ISI Value: 1.188

MACHINE LEARNING APPROACHES FOR CALORIE BURN ANALYSIS AND PREDICTION

Neha Valmiki[1], Ambili P S[2*]

[1]*PG Student, School of Computer Science and Application, REVA University, Bangalore, India*
[2]*Associate. Professor, School of Computer Science and Application, REVA University, Bangalore, India*

[*]***Corresponding Author***

Article DOI: https://doi.org/10.36713/epra15281
DOI No: 10.36713/epra15281

ABSTRACT

Aim: *It is widely acknowledged that running is the most effective activity for burning maximum calories per hour. Additionally, fixed bicycling, swimming, and HIIT (High-Intensity Interval Training) exercises are also excellent choices for calorie burning. HIIT exercises, in particular, have been shown to continue calorie burning for up to 24 hours after the workout. However, accurately predicting the number of calories burned during a specific exercise remains a challenge due to variations in individual physiology and fitness levels. To address this challenge, we utilized an exercise dataset obtained from the UCI Machine Learning repository to forecast the calories burned during different exercises.*

Methodology: *Many individuals are curious about the calories burned during their workouts and the effectiveness of their weight loss plans. To tackle this issue, we employed machine learning algorithms such as XGBoost regressor and Linear Regression.*

Results: *The results demonstrate that XGBoost regressor and Linear Regression outperform other approaches in predicting calorie expenditure during exercise.*

KEYWORDS— *Machine learning, precision, accuracy, classification, XGBoost regressor, Linear Regression.*

INTRODUCTION

As of now, Machine education is being utilized in vivid regions in the wellbeing industry, including the improvement of cutting-edge clinical cycles, patient records and information activity, and the treatment of constant circumstances [1]. Medical care ventures might utilize machine education to satisfy rising need, cut expenses, and improve activities. The activity of machine education in medical care isn't new, as digitization has assumed control over each business. Taking an interest, conveying, and obtaining information is getting decreasingly significant. Man-made consciousness, large information, and machine proficiency can all assist with the issues that gigantic volumes of information can bring forth [2,3,4]. Then again, machine education can assist translators with treating conditions all the more successfully by feting them previously and outfitting additionally customized care.

In this quick and occupied plan life, individuals aren't giving importance to the nature of food they're eating. They will generally disregard their eating examples and propensities. The cheap food utilization rate is alarmingly high, prompting unfortunate food contributions. This prompts brilliant medical problems comparable as stoutness, diabetes, an expansion in pulse and so forth. Subsequently, it has become genuinely fundamental for individuals to have a decent adjusted nutritive solid eating routine. Various tasks are thundering to help individuals so they can have command over their eating regimen and subsequently can lessen weight or assist them with keeping fit and sound. This plan centers around the calories consumed in concurrence with the length gave and pulse during the activity time frame. It presents the substance of direct retrogression and its forecasting ability with the adequacy from the information gave. This investigation helps in outfitting the advantages of an AI calculation over anticipating the calories consumed [5,8,10].

Calories in the food sources we eat give energy as intensity so our bodies can serve. This implies that we really want to eat a specific number of calories just to support life. However, on the off chance that we take by too various calories, additionally, we risk putting on weight. Thus, there's need to consume Calories, for consuming calories we doing practices and further. for know-how significant calories we've consume Moment we will make a machine education model that forecast calories grounded on certain information.

The daily energy expenditure plays a crucial role in weight loss, weight gain, or weight maintenance. To achieve weight loss, individuals must consume fewer calories than they intake, creating a calorie deficit. However, accurately determining the number of calories burned each day is essential. Many people associate calories primarily with food and weight loss. It is important to note that a calorie is a unit of energy or intensity.

LITERATURE SURVEY

The food recognition system initially divided the food images into different categories to create a feature vector based on size, shape, texture, normalized RGB values, and other contextual features. This feature vector was then augmented with Gabor

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 9| Issue: 12| December 2023|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2023: 8.224 || ISI Value: 1.188

filter responses (texture), pixel intensity, and variation factors. The system exhibited good performance for food clones but showed less effectiveness with real images [6]. The variation in image sizes and capturing methods may contribute to the decline in performance. Interestingly, the system achieved better performance when dealing with a lower number of classes, despite having more images per class. The images contain common visual patterns that aid in determining food categories, reducing the complexity associated with direct image-matching approaches [7, 9, 13]. The system was evaluated using a real-world database of fast food images.

Hitherto the investigation has focused on gathering the sorts of datasets and proposed a new dataset to appraise calculations to fete food which assists with covering eats less carbs. The information base has been made for an additional number of instances of food pictures. A medium-measured dataset has been made to foster a portable grounded log framework. Profound Convolutional Neural Networks have been utilized for food acknowledgment recently which utilizes a blend of birth point birth and brain network adjusting [12,14]. Convolutional Neural Networks alongside a Maximum Pooling subcaste produces Activation Charts (heat graphs of food likelihood). Adjusting is finished Activation Chart age, which incorporates adding a convolutional subcaste with step, and setting a delicate greatest subcaste. Likewise, by means of edge, it is produced to bound boxes. The current work is to join above procedures together, that makes a food section framework that predicts the class of food the picture is in, and furthermore gives the calorie esteem grounded on the food weight or count given by the stoner [13,15,16]. This origination has a high compass in the wellbeing area, as individuals need to monitor what and how significant they eat and improving on the cycle into this type of execution expands activity and care of wellbeing related factors.

EXISTING METHOD

Presently a day, individuals are having authentically bustling timetables because of changes in their day-to-day existence and work responsibilities. Individuals aren't gathering in their food propensities, henceforth, it prompts breadth. Breadth is getting a typical issue in second's ultramodern life. So, we want a framework which can pursue changes to the food decisions of individuals and furnishes them with guidance that prompts viable consequences of keeping up with their bodies. However, likewise individuals are appropriate to distinguish their diurnal contribution of calorie worth of their food specifics. If a framework advises the nutritive data regarding a food thing and characterizes it as sound or non-beneficial to the stoner. This proposed framework assists the stoner with controlling their food propensity framework and furthermore gives data on the most proficient method to consume calories in diurnal schedules that makes stoners solid. Convolutional Neural Network model is applied to arrange food things from the info picture and the proposed framework gives the delicacy of 91.65.

PROPOSED METHODOLOGY

This paper focuses on developing a machine learning model to accurately predict the calorie intake of an individual. Before applying the learning algorithms, the data undergoes preprocessing to ensure its quality and compatibility. The data is then analyzed using visualization techniques to gain insights and patterns. The dataset is divided into training and test sets for model evaluation. XGBoost regressor and linear regression are employed as machine learning models for analysis and prediction [11]. The tool used for this study is Google Collaboratory (Google Collab), which is a web-based platform and cloud service.

Fig. 1: Proposed Method Block Diagram

Implementation

Then we're utilizing two sort of calculations on provided informational collection all together make out a valuable model for foreseeing the calories consumed during the drill grounded on the drill span, age, orientation, level and weight of the individual.

Algorithms

RGBoost Regressor

RGBoost Regressor is a regression algorithm used to analyze the relationship between a dependent (target) variable and independent (predictor) variables with one or more independent variables. In the field of artificial intelligence, the XGBoost algorithm is highly effective due to its ability to handle various types of data, relationships, distributions, and the availability of multiple hyperparameters for fine-tuning. XGBoost regressor can be employed for regression, classification for both binary and multiclass problems, and ranking tasks.

Linear Regression

Linear Regression is a supervised learning algorithm in Machine Learning. It establishes a linear relationship between one or more independent variables and a dependent variable by fitting a specific set of data values (x) to the predicted output (y) using a linear equation of the form Y = a +bX.

Steps

A. Dataset Collection
B. Data Preprocessing
C. Data Analysis
D. Machine Learning Model
E. Evaluation

Data Source

The dataset used for this study was obtained from Kaggle. It consists of two CSV files, comprising 15,000 instances and 7 attributes. The dataset obtained from the Kaggle repository includes various features related to each individual's characteristics, such as gender, age, exercise duration, heart rate, body temperature, height, and weight. This dataset serves as the training data. Additionally, the supplementary calories

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 9| Issue: 12| December 2023|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2023: 8.224 || ISI Value: 1.188

dataset contains the target class, which represents the corresponding calories consumed by each individual.

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 the exercising in minutes
heart_rate	Average heart rate during the workout (more than normal rate 75 beats/min)
body_temp	Body temperature during the workout (greater than 37 degree Celsius)
calories	Total calories burnt during the exercise.

There are two dataset csv lines which ought to be transferred to Collab which is utilized for handling.

We use information outlines for investigation and handling. It gain a few factual measures about the information.

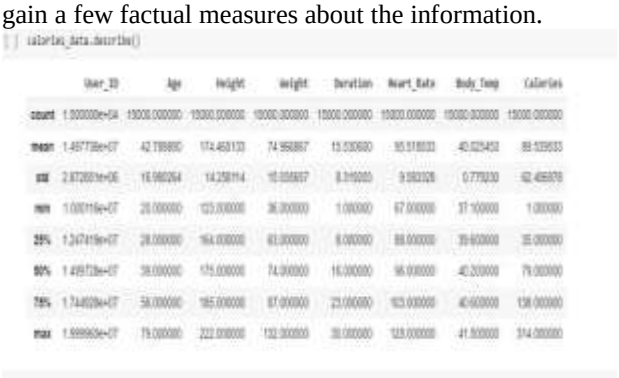


Fig. 2: Dataset

The average age of the individuals in the dataset is 42.7 years, and we have information on the standard deviation and percentiles. The body temperature is approximately 40 degrees. Individuals who engage in more physical activity tend to have higher body temperatures. The key variables of interest for this analysis are the heart rate and temperature. In order to better understand the data, we will visualize it using various plots and charts.

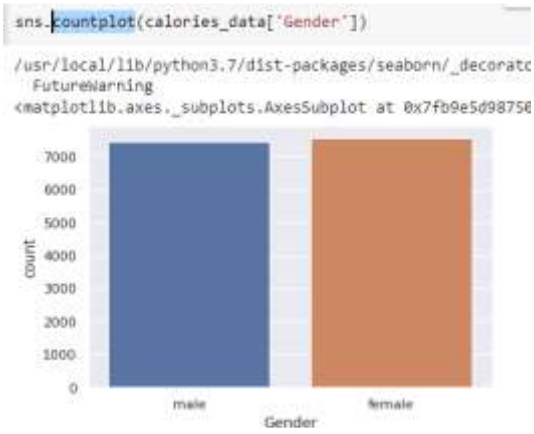


Fig. 3: Gender Plot

we have almost great indistinguishable conveyance of masculine and womanish pieces of information

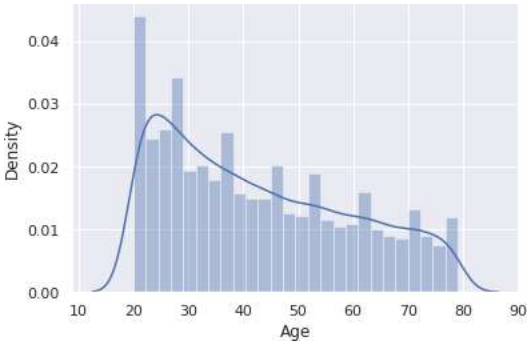


Fig. 4: Age plot

The average age of the individuals in the dataset is 42.7 years, and we have information on the standard deviation and percentiles. The body temperature is approximately 40 degrees. Individuals who engage in more physical activity tend to have higher body temperatures. The key variables of interest for this analysis are the heart rate and temperature. In order to better understand the data, we will visualize it using various plots and charts.

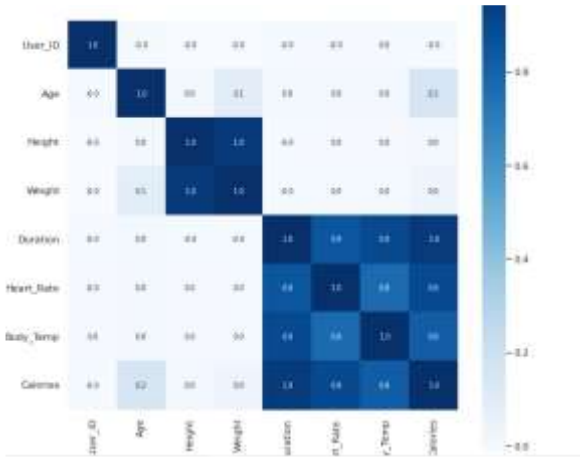
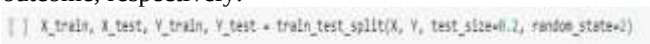


Fig. 5: Correlation Heatmap

This implies that we can discern the endless weight with gratitude, and there is a thankful correlation between the terms pulse, internal heat level, and calories. Now, let's organize the information into training and testing data by separating the factors X and Y, which represent variables and the desired outcome, respectively.



Moreover, the XGB Regressor model is loaded to perform predictions on the test data. This model assesses the test data and estimates the calories consumed, utilizing the values from X_test. Furthermore, a comparison is made between the predicted values generated by our model and the actual values. The mean absolute error metric is employed to evaluate the accuracy of the predictions, which measures the average magnitude of errors. The obtained mean absolute error is 2.71.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 9| Issue: 12| December 2023|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2023: 8.224 || ISI Value: 1.188

Fig:6- XGBregressor output

Next, we do this in linear regression model. lin_reg_model=LinearRegression ()
The mean absolute error getting is 8.38.

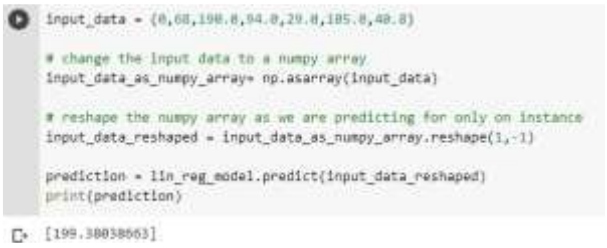


Fig. 7: Linear Regression Output

RESULT

The purpose of analyzing this dataset was to estimate the calories consumed based on the exercise duration and factors such as gender, age, internal heat level, and pulse during the activity. We used advanced AI algorithms to find a machine learning model with a lower mean absolute error, indicating more accurate results. Comparing the XGB Regressor and Linear Regression models, we found that the XGB Regressor provides more precise calorie predictions with a mean absolute error of 2.71, surpassing the accuracy of the direct regression model.

Table 2: Result of Mean Absolute Error

Machine Learning Method	Mean Absolute Error
XGB Regressor	2.71
Linear Regression	8.38

Table 3: Absolute Result

Machine Learning Model	Input Data	Predicted Calorie	Expected Calorie
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	199.38	231.0

CONCLUSION

Based on the analysis, the authors have determined that the XGB Regressor provides more accurate results compared to the Linear Regression model. The mean absolute error, which measures the discrepancy between the actual and predicted values, should ideally be kept as low as possible. In the case of the XGB Regressor, the mean absolute error value obtained is 2.71, which is considered good. The error values are significantly lower, indicating that the XGBoost Regressor is a reliable model for predicting calorie consumption. Therefore, the study can confidently conclude that the XGBoost Regressor is the preferred model for predicting calorie consumption.

ACKNOWLEDGEMENT

The authors acknowledge the support from REVA University for the facilities provided to carry out the research.

REFERENCES

1. *F. Mendes, "fmendesdat263xdemos," Kaggle, (2023). [Online].Available:https://www.kaggle.com/fmendes/fmendesdat263xdemos.*
2. *J. Brownlee, "Linear Regression for Machine Learning," Machine Learning Mastery, (2023). [Online]. Available: https://machinelearningmastery.com/linear-regression-for-machine-learning/.*
3. *J. Brownlee, "XGBoost for Regression," Machine Learning Mastery, (2023). [Online]. Available: https://machinelearningmastery.com/xgboost-for-regression/.*
4. *Y. Liu, G. Xiang, Y. Lu, Y. Cao, Y. Li, and L. Lv. (2016), "Calibration of MEMS Accelerometer Based on Kalman Filter And the Improved Six Point Position Method," Journal of Communication, vol. 11, no. 5.*
5. *Ambili, P S and Biku Abraham. (2022), "A Predictive Model for Student Employability using Deep Learning Techniques", ECS Transactions 107 (1):10149,(2022), DOI: https://doi.org/10.1149/10701.10149ecst.*
6. *E. A. Finkelstein. (2012), "Obesity and severe obesity forecasts through 2030," American Journal of Preventive Medicine, vol. 42, no. 6, pp. 563-570.*
7. *L. Vinh, S. Lee, H. Le, H. Ngo, H. Kim, M. Han, et al.(2011) , "Semimarkov conditional random fields for accelerometer-based activity recognition," Applied Intelligence, vol. 35, pp. 226-241.*
8. *Y. Kawahara, N. Ryu, and T. Asami.(2009), "Monitoring daily energy expenditure using a 3-axis accelerometer with a low-power microprocessor," International Journal on Human-Computer Interaction, vol. 1, no. 5, pp. 145-154.*
9. *P. G, D. PS, S. Suresh and P. K. Pareek.(2023), "Scheduling IoT Application Tasks using Flamingo Search Algorithm in Cloud Computing," 2023 International Conference on Network, Multimedia and Information Technology (NMITCON), pp. 1-6, doi: 10.1109/NMITCON58196.2023.10276133.*
10. *Abraham, B., Ambili, P.S. (2023)," An Enhanced Career Prospect Prediction System for Non-computer Stream Students in Software Companies", Lecture Notes in Electrical Engineering, vol 984. Springer, Singapore, doi:.10.1007/978-981-19-8493-8_60*
11. *I. Janssen.(2005), "Comparison of overweight and obesity prevalence in school-aged youth from 34 countries and their relationships with physical activity and dietary patterns," Obesity Reviews, vol. 6, no. 2, pp. 123-132.*
12. *S. Dreiseitl and L. Ohno-Machado.(2002), "Logistic Regression and Artificial Neural Network Classification Models: A Methodology Review," Journal of Biomedical Informatics, Elsevier, vol. 35, pp. 352-259.*
13. *Pillai, A.P.S. (2023), "AIoMT-Assisted Telemedicine A Case Study of eSanjeevani Telemedicine Service in India", Handbook of Security and Privacy of AI-Enabled Healthcare*

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 9| Issue: 12| December 2023|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2023: 8.224 || ISI Value: 1.188

Systems and Internet of Medical Things, pp. 445–464, DOI: 10.1201/9781003370321-19.

14. *Ambili P S, Agnesh L, & Arun K V. (2023), "Siamese Neural Network Model for Recognizing Optically Processed Devanagari Hindi Script", International Journal of Computational Learning & Intelligence, 2(3), 107–113, https://doi.org/10.5281/zenodo.8210372*
15. *D. Harrison, P. Marshall, N. Bianchi-Berthouze, and J. Bird.(2015), "Activity tracking: barriers workarounds and customisation," Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 617-621.*
16. *G. S. Morrison.(2013), "Tutorial on logistic-regression calibration and fusion: converting a score to a likelihood ratio," Australian Journal of Forensic Sciences, vol. 45, no. 2, pp. 173-197.*