**Machine Learning-Based Caloric Expenditure Prediction for Personalized Fitness Assessment during Physical Activity**

V Karthick, Associate Professor  
*Department of CSE*  
*Rajalakshmi Engineering College*Chennai, India  
*vkarthick86@gmail.com*

Sudharsan S, UG Student  
 *Department of CSE*  
 *Rajalakshmi Engineering College*Chennai, India  
*210701266*@rajalakshmi.edu.in

Swetha Adluru, UG Student  
 *Department of CSE*  
 *Rajalakshmi Engineering College*Chennai, India  
*210701276*@rajalakshmi.edu.in

***Abstract—* The goal of this project is to use machine learning for parsing how our body works when burning calories in different activities. Data was taken consisting of the variables, age, gender, height, weight to record body metrics also how long participants had engaged in exercise. The heart rate and temperature were taken as well. Once this data is preprocessed and arranged, we use this data to train multiple machine learning models such as Linear Regression, Decision Trees etc. The metrics used to evaluate these models — Mean Squared Error, R2 Score etc. indicate that the Random Forest Regressor is the best model in predicting caloric expenditure with minimum error of accuracy. To achieve this, we create a Streamlined UI application (below) that allows the users to input their data age, gender, height and weight in metric units of [kg], type of exercise hour duration — walking or running— heart rate during training and temperature in Celsius unit. The application then uses the learned machine learning model to predict how many calories were burned and produce the output with the accuracy level of 93.77%.**

***Keywords— Caloric Expenditure Prediction, Machine Learning, Personalized Fitness Assessment, Random Forest Regressor, Data Preprocessing, Health and Fitness, Model Evaluation, Streamlit UI Application.***

1. INTRODUCTION

Among the recent fitness paradigms, accurate prediction of caloric throughput during physical exertion is important for shaping personalized Fitness Assessment. Many given equations in traditional approaches might be based on commonly generalized formulas that fail to consider an individual’s differences in metabolism and physical activity. This project uses machine learning (ML) to increase the predictive accuracy, while relying on a comprehensive dataset in which most important factors are provided: age, gender, height and weight of people who exercise, duration of their workouts and information about body properties like heart rate during entire workout session and temperature at that time as well as actual burnt calories. The ultimate goal is to produce an ML model that can properly predict caloric expenditure in physical activity. The project then proceeds through a structured process of data preprocessing. This involves merging data and transforming certain traits to numerical form using the ‘train\_test\_split’ function from the sklearn library for creating training and testing datasets.

Do experiments with different machine learning models like Linear Regression, Ridge Regression, Lasso Regression Decision Tree Regressor and Random Forest Regressor to see which model working the best in order to predict calorie burn while doing exercise After evaluating the model using metrics such as Mean Squared Error and R2 Score we found that, Random Forest Regressor is perform the best among all other models in term of predicting power. It is essentially smart about how to factor in the multiple components that contribute to calorie-burning. We saved the model using the Python pickle module so that this technology is much more accessible. We also created a user-friendly Streamlit UI application as well. This is an application where a user can enter his age, gender, Height, Weight and receive along with heart rate of the person ( beats/min) throughout the workout hours where he will be able to obtain approx. amount of calories burnt by using trained ML mode

With that said, the organization also makes it clear how imperative utilizing modern techniques to improve personalized workout routines and health evaluations is as well. When researchers applied machine learning approaches, like the Random Forest Regressor, this allowed for accounting those variations in metabolic rates and exercise intensity at an individual level to generate more accurate predictions. According to the approach, fitness people and professionals can take better food and activity decision. Streamlit UI tool is really easy to estimate caloric expenditure with more convenience. The end result is “user-friendly” software that enables users to input their own personal data and receive in return a rapid assessment of how many calories they have torched. Therefore, the use of machine learning for maximizing fitness and health is very vast.

Its improved predictive ability aside, the project currently aims to solve a shortfall of current systems—many rely on only limited substantive information, and as such can yield static or impersonal forecasts. These traditional methods don’t consider the intensity of each toddling step or running stride (or any other kind of physical activity) — or varying individual fitness levels. This project is more reactive or responsive, based on machine learning techniques. With its ability to capture the complex patterns present in the dataset, Random Forest Regressor comes out to be a useful tool for predicting caloric expenditure. The project also underlines the importance of user-friendly interfaces as a way to democratize access for this kind of high-tech features. The above Streamlit UI application is a perfect example of how one can leverage user interface, interact with that and receive the results as machine learning insights about your health and fitness.

1. MATERIALS AND METHODS

The materials and methods applied on this studies project are geared toward predicting caloric expenditure at some point of physical hobby through device getting to know algorithms. The challenge leverages a comprehensive dataset comprising vital metrics along with age, gender, peak, weight, workout duration, coronary heart fee, temperature throughout workout, and actual energy burnt. Another dataset that links user IDs with total energy burned is included to assessment processes.

First, we start by combining different datasets to make one whole dataset that we will use for assessment. We change specific such things as gender into numbers so that our pc programs can understand them higher. To make sure our laptop applications study nicely, we split our records into parts: one for coaching them and one for checking out how right they have emerged. We set aside 20% of the facts for

checking out. We then use exceptional methods, like Decision Tree and Linear Regression, to train these applications. They analyze from the patterns in our prepared facts. After they have found out, we check how accurate they may be by means of the use of measures like R2 Score and Mean Squared Error. These measures help us see if our programs are predicting effectively. These measures help us recognize how effective our applications are at making predictions.

Following evaluation, the Random Forest Regressor emerges as the optimum version, demonstrating advanced performance with minimal Mean Squared Error and excessive R2 Score. The Python pickle module is hired for model serialization, allowing the very last calibrated Random Forest Regressor model to be stored and deployed for real-time predictions.

Additionally, a consumer-pleasant net application is advanced the usage of the Streamlit framework. This software lets in users to input parameters together with age, gender, top, weight, exercise length, coronary heart charge, and temperature, sooner or later producing predictions of caloric expenditure based totally on the chosen Random Forest Regressor model. The techniques and tools we're using in this project help us figure out how many calories a person burns when they exercise, and they also make our health assessments more Personalized Fitness Assessment using machine learning. But right now, the system isn't very good at adjusting its calculations for different kinds of workouts, how long they last, or the surroundings where they happen.Because it doesn't adapt or use advanced analysis, it might give us wrong predictions sometimes, which can lead to less effective fitness plans. This shows why we need more advanced methods, like machine learning algorithms, to make better predictions about calorie burning and improve personalized health assessments for physical activity planning.

1. EXISTING SYSTEM

The contemporary systems for estimating calorie expenditure throughout workout rely on conventional methods and guide guesses the usage of preset equations or algorithms. However, this method lacks depth and accuracy because it would not don't forget person body sorts and interest ranges.

One predominant downside of the existing device is its incapacity to recollect the various factors that impact calorie burn. Everyone's metabolism and health tiers range, but the cutting-edge programs treat all users the same. This ends in conventional estimates that might not mirror a person's real calorie burn accurately, affecting fitness planning and strategies.

Moreover, the consumer interfaces of current structures are frequently fundamental and lack actual-time comments or interactive functions. Users enter their statistics and obtain static effects without the potential to visualize tendencies, examine situations, or get personalized suggestions. This static nature makes it challenging for the machine to evolve to convert user desires and choices, restricting its usefulness for health fans and specialists.

The current system has trouble adjusting its calculations for different levels of exercise intensity, duration, and the environment. It lacks the ability to make dynamic changes or use advanced analytics, which can lead to inconsistent or inaccurate predictions. This can result in less effective fitness planning outcomes. Overall, these challenges show the importance of using more advanced methods like machine learning algorithms to improve predictions of calorie burn, personalized health assessments, and fitness optimization strategies.

1. PROPOSED SYSTEM

The new exercise calorie burning predictor system. This was designed to solve the problems with other methods out there. It is an accurate and personalized fitness assessment using cutting-edge machine learning methods. We will develop a model that predicts accurately how many calories a person burns when they exercise. The random forest regressor approach is what we are concentrating on because it can handle complex data and give precise predictions.

Initial step encompasses collection of different data points like age, gender, height, weight, duration of exercise, heart rate as well as temperature during exercise. It allows us to have an inclusive set of data which takes into account various factors influencing the burning of calories. We cleanse and arrange this so that it’s ready for analysis.

The subsequent thing we do is preprocess the dataset by converting categorical variables such as gender into numerical values which can be understood by machine learning models; after that, we divide our dataset into two parts – one used for training the model and another reserved for testing it. This helps us evaluate how well the model performs in predicting new observations: allocating 20% of all available records ensures high accuracy during evaluation.

In order to democratize this technology, we create a straightforward application using the Streamlit framework. With our app, users can input details about themselves like their age group, sex, time spent working out (in minutes), average heart rate while exercising and temperature during the activity period. Then this application uses a model trained on the above information to estimate how many calories a person burnt during the exercise.

1. METHODOLOGY

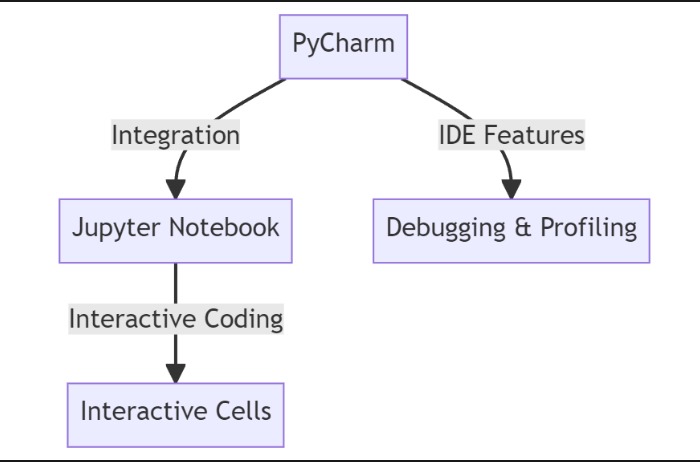


Fig 1. Architecture Diagram

The method used in this project is to employ training sets (Learning Machines) that will improve forecasts on how much energy a person uses each day and individual evaluations for physical fitness. It starts with collecting a lot of information which include age, sex, height in centimeters or meters; weight in kilograms, pounds or stones; duration of exercises in minutes per session; typical heart rates during workouts measured as beats per minute (bpm); ambient temperatures while working out known as “exercise temp” in degrees Celsius (°C), Fahrenheit (°F) etcetera followed by the real amount of calories burned.

Later, we trained several machine learning models like linear regression, ridge regression, lasso regression, decision tree regressor, and random forest regressor to find patterns in the data. Performance evaluation metrics such as mean squared error (MSE) and r2 score are used to check the models’ predictive capability. It is worth noting that Random forest regressor stands out as the best model because it can handle complex relationships between variables well and accurately predict calorie expenditure. Additionally, I have created a pickle module serialized random forest regressor model which can be deployed for use in real time without any difficulty.

Afterward, we trained several machine models for deep learning on the data, including linear regression, ridge regression, lasso regression, decision tree regressor, and random forest regressor, and used mean squared error (MSE) and r2 score as performance evaluation metrics to verify the models’ predictive ability. It is important to note that the random forest regressor is the best model among them all since it can handle complex relationships between variables effectively and predict energy expenditure accurately based on this fact. Furthermore, I have saved a serialized version of the random forest regressor model using a pickle module which can be deployed easily for real-time usage.

1. RESULTS AND CONCLUSION

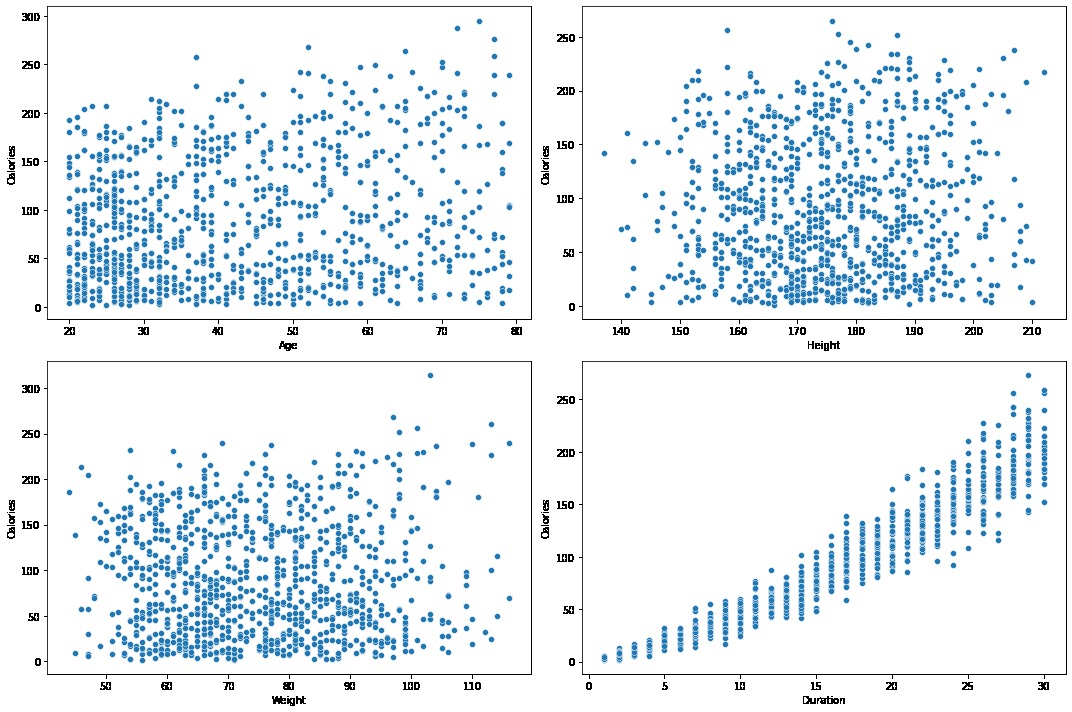


Fig 2 Graphical representation of trained data

Our research indicates the Random Forest Regressor is the most effective model among machine learning models in estimating the number of calories burned. The Random Forest Regressor was found to be the most accurate and reliable after a profound evaluation using metrics such as Mean Squared Error and R2 Score. It easily accommodates various aspects including age, gender, time spent exercising and weather conditions thereby leading to accurate predictions. This system has been incorporated into a user-friendly Streamlite UI app where people can feed in their details then get real time precise calorie burn estimations. The application’s simplicity coupled with its ability to provide feedback instantly has made it quite easy to use thus becoming important among people who love keeping fit.

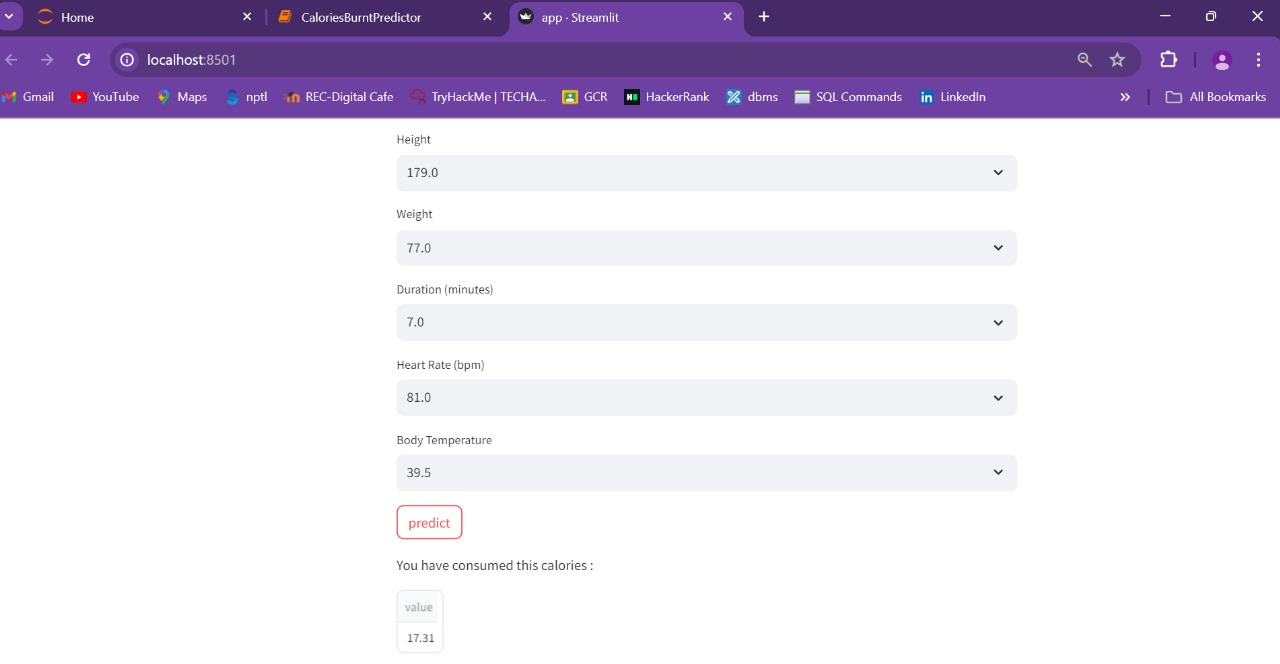
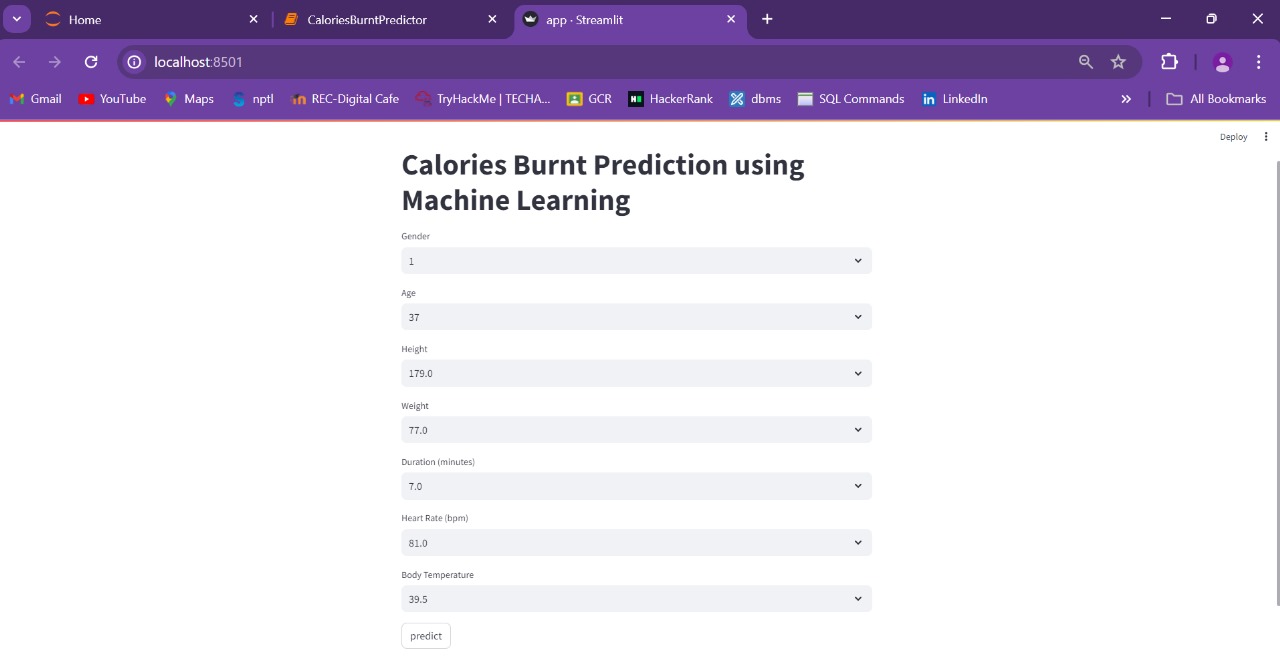


Fig 3, 4 Output Screenshots

It was concluded in this project that machine learning has great potential. This particularly lies in the Random Forest Regressor which enables the accuracy of 93.77% in predicting caloric expenditure for personalized fitness assessments to be improved. We made sure that our predictions were reliable and accurate by combining extensive data preprocessing with strong model evaluation. The development of Streamlit UI application has made such advanced technologies available to wider audiences thus enabling informed decision making on health and fitness matters. This research emphasizes how much impact can be brought about by machine learning in the healthcare industry concerning physical exercises whereby it can provide grounds for future developments related to personalized health diagnosis as well as optimization of workouts. Integration of artificial intelligence into daily routines related to keeping fit is seen as a big stride towards using technology for general well-being improvement.

1. REFERENCE

1. Aziz, Mohammad Tarek, et al. "Calories Burnt Prediction Using Machine Learning Approach." Rangamati Science and Technology University, Bangladesh, October 2023. Available at: [(PDF) Calories Burnt Prediction Using Machine Learning Approach](https://www.researchgate.net/publication/375147744_Calories_Burnt_Prediction_Using_Machine_Learning_Approach)

2. Panwar, Punita, et al. "A Study on Calories Burnt Prediction Using Machine Learning." Jaipur Engineering College & Research Centre, Computer Science Department, Jaipur, Rajasthan, India, July 2023. Available at: [(PDF) A Study on Calories Burnt Prediction Using Machine Learning](https://www.researchgate.net/publication/372104285_A_Study_on_Calories_Burnt_Prediction_Using_Machine_Learning)

3. Kadam, A. "Calories Burned Prediction Using Machine Learning." IEEE Xplore, 2023. Available at: [Calories Burned Prediction Using Machine Learning | IEEE Conference Publication](https://ieeexplore.ieee.org/document/10397623)

4. Meenigea, Niharikareddy. "Calorie Burn Prediction: A Machine Learning Approach using Physiological and Environmental Factors." JETIR, 2014. Available at: [Calorie Burn Prediction: A Machine Learning Approach using Physiological and Environmental Factors](https://www.jetir.org/papers/JETIR1701933.pdf)

5. A Study on Calories Burnt Prediction Using Machine Learning." ITM Web of Conferences 54, 01010 (2023) I3CS-2023. Available at: [A Study on Calories Burnt Prediction Using Machine Learning | ITM Web of Conferences](https://www.itm-conferences.org/articles/itmconf/abs/2023/04/itmconf_I3cs2023_01010/itmconf_I3cs2023_01010.html)

6. Nipas, Marte, et al. "Burned Calories Prediction using Supervised Machine Learning: Regression Algorithm." 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA).

7. Wesly, U. John, et al. "Calorie and Body Mass Index Prediction Using Machine Learning." International Journal of Research Publication and Reviews, Vol 4, no 4, pp 1154-1158, April 2023. Available at: [Calorie and Body Mass Index Prediction Using Machine Learning](https://ijrpr.com/uploads/V4ISSUE4/IJRPR11482.pdf)

8. Ratnakar, Suvarna Shreyas, and Vidya S. "Calorie Burn Prediction using Machine Learning." International Advanced Research Journal in Science, Engineering and Technology, Vol. 9, Issue 6, June 2022. Available at: [Calorie Burn Prediction using Machine Learning](https://iarjset.com/wp-content/uploads/2022/07/IARJSET.2022.961125.pdf)

9. Vinoy, Sona P., and Binumon Joseph. "Calorie Burn Prediction Analysis Using XGBoost Regressor and Linear Regression Algorithms." Proceedings of the National Conference on Emerging Computer Applications

10. Singh, Rachit Kumar, and Vaibhav Gupta. "Calories Burnt Prediction Using Machine Learning." Student, Information Technology, Maharaja Agrasen Institute of Technology, Rohini, Delhi. Available at: [Calories Burnt PredictionUsing Machine Learning](https://ijarcce.com/wp-content/uploads/2022/06/IJARCCE.2022.115105.pdf)