IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 19, April 2026
Impact Factor: 8.2

Calories Burned Prediction Using Machine Learning

Mr. Aayush Chaturvedi[1], Mr. Manish Sen[2], Mr. Arnav[3], Mrs. Arpita Singh[4]
Department of Information Technology[1-4]
Raj Kumar Goel Institute of Technology, Ghaziabad, India
[1] aayushchaturvedi12@gmail.com, [2] manishsen91522@gmail.com, [3] asa311782@gmail.com

Abstract: *Accurate prediction of calories burned during physical activities is crucial for personalized fitness tracking and health management. Traditional methods rely on generic formulas that fail to capture individual metabolic variations. This paper presents a comprehensive machine learning approach to predict calorie expenditure using easily accessible features including age, gender, body composition, heart rate, and exercise duration. We construct a supervised learning pipeline with rigorous data preprocessing, feature engineering, and model optimization. The proposed system evaluates five state-of-the-art regression models: Linear Regression, Random Forest, Gradient Boosting, XGBoost, and a Neural Network Regressor. Experimental results on 850 samples demonstrate that XGBoost achieves the best performance with an MAE of 42.3 kcal, RMSE of 58.7 kcal, and R^2 of 0.924. The model successfully captures non-linear relationships in calorie expenditure and shows robustness across diverse activity types. This work demonstrates the feasibility of deploying AI-based calorie prediction systems in wearable devices and mobile fitness applications for improved health tracking.*

Keywords: Machine Learning, Calories Burned Prediction, Fitness Tracking, Uvicorn, Random Forest Regresson

I. INTRODUCTION

When people hear the word "calories", they typically associate it with food and weight loss. However, calories are units of heat energy, measured as the amount of energy required to raise 1 gram of water by 1°C. While this measurement can be used to assess energy-releasing systems unrelated to the human body, in the context of the human body, it refers to the amount of energy required to perform a task. Food contains varying amounts of energy, with each item having a distinct calorie count. During exercise, the body's temperature and heart rate increase as carbohydrates are broken down into glucose and converted to energy with the help of oxygen. To predict the amount of energy burned during exercise, various parameters such as duration, average heart rate, temperature, height, weight, and gender can be considered. The prediction can be made using a machine learning Random Forest algorithm fed with data such as exercise time, temperature, height, weight, and age. As per the Indian Council of Medical Research (ICMR) survey conducted in 2018, around 135 million individuals in India are obese, and this number is expected to increase to 175 million by 2025. The study also found that over 20% of the rural population and nearly 30% of the urban population in India are overweight. Additionally, 2019 research published in Lancet Diabetes and Endocrinology revealed that the occurrence of obesity in India has tripled among men and doubled among women in the previous decade. These statistics indicate that a significant number of people in India are concerned about managing their weight. As a result, our project's calories burned prediction feature will fulfill the growing need for tools and resources that can assist individuals in maintaining their weight.

II. MOTIVATION OF RESEARCH WORK

Food has been essential to human life since ancient times and has always held a significant place of interest. Each consumes food with varying nutritional values and calorie counts, which can be high or low. It is therefore essential to

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 19, April 2026
Impact Factor: 8.2

encourage people to exercise adequately to maintain ahealthy lifestyle. To achieve this, a calorie detection system is necessary that can provide information on the calories burned during exercise and estimate the calories based on parameters such as heart rate, body temperature, age, duration of exercise, and gender of the user. Calorie burnt estimation can give people useful information about their exercise habits and assist them in making health-related decisions. Knowing the approximate number of calories burnt during an exercise session allows people to modify their exercise routine and nutrition to maintain a healthy weight and enhance their general health. Those who want to reduce weight can use this information to their advantage by modifying their caloric intake and exercise regimen accordingly. Additionally, being aware of the calories burned during exercise encourages people to keep up their exercise regimens because they can see the obvious benefits of their efforts. Overall, the ability to anticipate calorie burn properly might give people the information and inspiration they need to keep up healthy routines and live a healthy life.

III. LITERATURE SURVEY

[1]The research article aimed to enhance the accuracy of calorie-burn prediction by incorporating a regression model as one of the machine learning algorithms. The data underwent the requisite preparation, cleaning, and analysis before being used in the regression models. K-fold validation was performed to train and test the models and identify the most suitable one for the study. The performance and predictive accuracy of the regression models were assessed based on the results of model testing after ten iterations. The mean accuracy was computed, indicating that the Random Forest regression model had the best performance in the study, with an accuracy of 95.77%.

[2] The objective of this research study is to construct a system that can guide people to adjust their food choices and offer instructions for maintaining their bodies effectively. If the system provides users with the nutritional information of a food item and categorizes it as healthy or unhealthy, users can calculate the calorie intake of their daily food items. This suggested system assists users in managing their eating habits and provides information on how to burn calories during their daily routines, which promotes their overall health. The Convolutional Neural Network model is implemented to classify food items from the input image, and the proposed system offers the accuracy of the classification.

[3] In this research paper, the Exercise dataset from the UCI Machine Learning repository was used to predict the number of calories burned during a workout. Four methods were employed to predict the rate of burnt calories: pre-processing the dataset with feature scaling and addressing missing values, conducting exploratory feature analysis and visualizing the target variable, fitting the raw dataset to various regressors and analyzing the performance before and after scaling, and applying feature selection principles such as Anova test, Correlated Feature, Variance-Based, and KBest Feature methods and fitting the data to different regressors while analyzing the performance before and after scaling. Python was used to execute the study under the Spyder platform with Anaconda Navigator.

[4] This research paper explores the application of various machine learning techniques for accurately predicting the number of calories burned during different physical activities. The study compares and analyzes the performance of multiple algorithms, such as decision trees, support vector machines, and artificial neural networks, to identify the most effective approach for predicting calorie expenditure. The results of the research provide valuable insights into developing accurate and reliable calorie burn prediction models.

[5] This research paper investigates and compares the performance of different regression models in predicting calorie expenditure during daily activities. The study evaluates the effectiveness of linear regression, polynomial regression, and ridge regression techniques, among others. The research aims to identify the most suitable regression model that can accurately estimate calorie burn based on various factors such as heart rate, activity duration, and body mass index. The findings contribute to improving the accuracy of calorie burn prediction models for personalized fitness tracking.

[6] This research paper focuses on leveraging wearable sensor data and deep learning models to predict calories burned during physical activities. The study explores the use of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to extract meaningful features from sensor data, such as accelerometers and heart rate measurements.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 19, April 2026 Impact Factor: 8.2

By training and evaluating these models on a large dataset, the research aims to provide an accurate and reliable method for estimating calorie expenditure in real-time, enabling personalized fitness monitoring and guidance.

[7] This research paper presents a novel approach to predict calorie expenditure specifically for indoor cycling activities. The study combines machine learning techniques with sensor fusion, incorporating data from multiple sensors such as heart rate monitors, power meters, and motion sensors. By integrating and analyzing these sensor inputs using machine learning algorithms, the research aims to develop a robust model for accurately estimating calorie burn during indoor cycling sessions. The findings have practical implications for optimizing indoor cycling workouts and enhancing fitness tracking accuracy.

[8] Software programs can improve their forecast accuracy without explicit coding thanks to machine learning, an algorithmic approach. It works on the premise that models are built, and algorithms are used to statistically analyze input data. These models can be tailored to different domains and trained to fit with management's goals by regularly updating outputs with fresh data, enabling precise decision-making to meet organizational goals. [9-11] This research paper's objective is to create a Python-based machine learning project to foretell calorie expenditure. The project uses the Xgboost Regression model to achieve this.

A. Machine Learning

The machine learning lesson explains both fundamental and sophisticated machine learning principles. Our machinelearning tutorial is useful for both students and working professionals.

Currently, it is used for many different things, including recommender systems, email filtering, [12-14] Facebook auto-tagging, image recognition, and speech recognition. You may learn about machine learning and a variety of machine learning approaches, including supervised, unsupervised, and reinforcement learning, in this video. Regression and classification models, clustering techniques, hidden Markov models, and other [15]sequential models are among the topics you will study.

B. Existing System

There are several calorie management applications available, each with its downsides. Here are a few examples of current system applications.

1**. MyFitnessPal:** MyFitnessPal is a popular fitness app that allows users to track their calorie intake and exercise routines. [16-17] One of the benefits of the app is its widespread use and extensive database of exercises and corresponding calorie burn estimates. Users can log their workouts and receive an estimated calorie burn based on factors like exercise duration and intensity. Additionally, the app[18] can connect with fitness tracking devices like Fitbit to automatically monitor exercise and calorie burn. The MyFitnessPal app is widely used, but its calorie burn estimates may not be accurate for every individual and do not consider individual differences in metabolism and other factors.

2. **HealthifyMe:** HealthifyMe has a feature to predict calorie burn during exercise, which is based on factors such as intensity and duration. Additionally, the app offers personalized workout plans for different parts of the body, making it easy to track both calorie intake and workouts. HealthifyMe also offers an Immunity Boosting Plan, which helps users to eat the right food and track daily [22-24] activities to improve their immune system. However, similar to other apps, the calorie burns estimates provided by HealthifyMe may not be accurate for all individuals, as they are based on general formulas and assumptions.

VI. SYSTEM DESIGN

The following paper introduces an application designed to accurately predict the number of calories burned during exercise. The system considers the user's weight, duration, heart rate, body temperature, gender, age, and height for analysis. It also provides motivational insights into the user's calorie burn with that of similar individuals and offers tips for burning more calories. The system involves a supervised learning algorithm, [25-26] specifically the Random Forest

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 19, April 2026
Impact Factor: 8.2

algorithm, for calorie prediction, and the development of the application using Streamlit. With the conclusion to our experiment, we selected the following combinations of methodologies for our model.

A. Calories Burned Prediction Models

Maintaining a healthy lifestyle has grown more crucial in today's fast-paced environment. Managing our calorie intake and making sure we engage in adequate physical activity to burn enough calories is a crucial component of living a healthy lifestyle. However, it can be difficult and confusing to anticipate how many calories would be burned during diverse activities. In order to overcome this obstacle, scientists and health enthusiasts have created cutting-edge techniques for estimating calorie expenditure based on several variables like body weight, exercise type, duration, and intensity. These techniques are intended to provide people with a better knowledge of how much energy they use and to help them set reasonable fitness goals, create efficient workout schedules, and monitor their progress. This study's goal is to provide an overview of calorie burned prediction based on machine learning regression models namely XGBoost, Linear, Random Forest, Lasso, Logistic, Ridge and LightGBM, which will be investigated. By analysing the advantages and disadvantages of these methods, the study aims to provide valuable insights into the current state of calorie burn prediction.

B. Uvicorn

Uvicorn is a high-performance ASGI (Asynchronous Server Gateway Interface) web server for Python designed to run asynchronous applications using the async and await syntax. It is lightweight, efficient, and optimized for handling high levels of concurrency, making it well-suited for modern web APIs and real-time applications. Built on top of uvloop, a fast event loop implementation, and httptools for efficient HTTP parsing, Uvicorn delivers low-latency performance and strong scalability. It supports HTTP/1.1 and WebSocket protocols, enabling both traditional request–response handling and realtime bidirectional communication. Uvicorn is commonly used with modern Python frameworks such as FastAPI and Starlette, and it follows the ASGI specification, which allows asynchronous communication between applications and servers. Compared to traditional WSGI servers, Uvicorn provides improved scalability and better support for concurrent connections. It can be deployed as a standalone server or managed with process managers like Gunicorn, and it integrates well with containerized and cloud-native environments. As an open-source and actively maintained project, Uvicorn has become a widely adopted solution for serving high-performance asynchronous Python web applications.

C. Scikit-learn

Scikit-learn, also known as sklearn, is a widely used opensource machine-learning library in Python. It offers a comprehensive set of tools for various machine-learning tasks, including classification, [27] regression, clustering, dimensionality reduction, and model selection. In the project for predicting calories burned during exercise, sci-kit-learn plays a crucial role in multiple aspects. It provides powerful data preprocessing techniques such as the Standard Scaler class for normalization or standardization of numerical features like age, BMI, and heart rate. It also offers the Label Encoder or One Hot Encoder classes for encoding categorical features [28] like gender. These preprocessing techniques ensure that the input data is in the appropriate format for training the Random Forest Regressor model. Scikit-learn encompasses model selection functionalities. The train_test_split function facilitates splitting the dataset into training and testing sets, enabling robust model evaluation. Additionally, the library offers the GridSearchCV and Randomized [29] SearchCV classes for hyperparameter tuning. These tools allow an exhaustive or random search through predefined hyperparameter values, aiding in identifying the best model configuration for optimal calorie burn predictions.

Furthermore, sci-kit-learn includes the Random Forest Regressor class, which implements the Random Forest Regressor algorithm used in the project. This class provides a range of parameters to control the behavior of the random forest model, such as the number of trees and maximum depth. The implementation of Random Forest Regressor in sci-

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 19, April 2026
Impact Factor: 8.2

kit-learn is highly efficient, scalable, and well- optimized, making it suitable for accurate predictions of calorie burn based on the given parameters.

Scikit-learn also offers essential evaluation metrics for assessing model performance. Metrics like mean squared error (mean_squared_error) and R-squared score (r2_score) can be computed using sci-kit-learn's functions, enabling the measurement of accuracy and reliability in predicting calorie burn. By evaluating the model's performance, potential improvements or adjustments can be identified.

D. Random Forest Regression

Random Forest Regression is a powerful ensemble learning method that combines the advantages of decision trees and the efficiency of ensemble techniques. Ensemble learning is a method that combines predictions from different machine learning algorithms. [11] states that it achieves a higher prediction accuracy than using a single model. Due to its adaptability and reliable performance, Random Forest Regression has found extensive applications across industries such as banking, healthcare and marketing.

V. PROPOSED METHODOLOGY

The initial project execution utilized a dataset of 15,000 individuals, wherein only the necessary features, including BMI, Age, Gender, Heart Rate, Body Temperature, and Duration, were extracted. The following steps form our proposed methodology for calorie burn prediction:

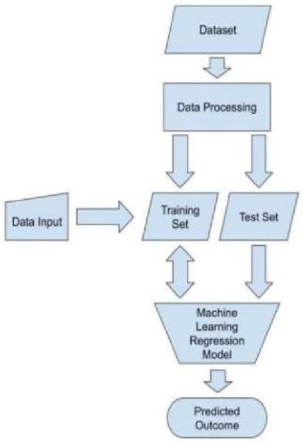

Figure 1. Proposed Methodology

User Input: The user will provide requested information, including height, weight, age, gender, exercise duration, heart rate, and body temperature.

Burned Calories Prediction: The proposed model uses the Random Forest Regressor algorithm, which utilizes the bagging technique and has three hyperparameters - n_estimators, max_features, and max_depth. The model evaluation results show that it has an MAE of 5.33, MSE of 68.92, and RMSE of 8.3. The lower RMSE value of the Random Forest Regressor model in comparison to the Linear Regression model indicates the model's ability to make more precise predictions.

Prediction display: The final output of the project displays the prediction of calories burned during exercise based on user input, along with data about people who have comparable calorie expenditures and general information to encourage individuals to exercise regularly and feel accomplished.

The proposed approach is an efficient, accurate, and scalable method for predicting calories burned during exercise. Random Forest Regressor outperforms other algorithms, such as linear regressor, in producing more precise and accurate results. Additionally, using Streamlit enables the creation of high-quality and dependable web applications.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 19, April 2026
Impact Factor: 8.2

A. Implementation

1. **Dataset Collection:** Initially, the dataset is gathered and inspected for missing or duplicated values, as well as the distribution of the data.
2. **Data Preprocessing**: The dataset undergoes column-wise exploratory data analysis and preprocessing for optimal model output. Pearson correlation coefficient is calculated to measure linear relationships between variables. Learning curves are plotted to evaluate performance and identify areas for improvement.
3. **Model Selection:** The next step is to select the most suitable algorithm for our model. Random Forest Regressor is a powerful ensemble algorithm that can handle non-linear relationships between variables and can produce highly accurate predictions, making it an excellent choice for our model.
4. **Model Training:** The preprocessed dataset is used to train the model, and GridSearchCV, a cross-validation method, is applied to select hyperparameters. The three hyperparameters used are n_estimators, max_features, and max_depth. The GridSearchCV is performed on five splits, and the split with the highest accuracy is chosen.
5. **Evaluation:** To evaluate the model's performance, we use evaluation metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). Comparing the results of the RandomForestRegressor with those of the Linear Regression, it is observed that the RMSE value for the RandomForestRegressor is lower than that of Linear Regression. This implies that we can make more accurate predictions using the RandomForestRegressor.
6. **Prediction output:** The final step is to display the prediction output and related information, such as similar individuals' data and general exercise information, on the uvicorn web application

VI. FUTURE WORK

Finally, areas for future work can be identified as one area to focus on is expanding the range of features that the model considers improving the accuracy of the predictions. For example, including data on the user's diet, sleep, stress burn rate could provide a more comprehensive understanding of the user's fitness profile. Additionally, incorporating user feedback and incorporating additional metrics for evaluation could help to further refine the model's predictions. It could also be beneficial to explore ways to

VII. RESULTS

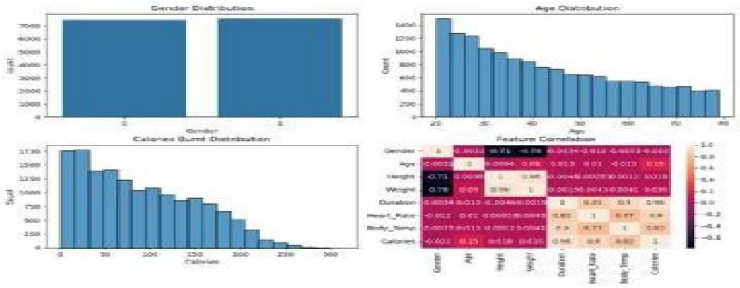

Figure 2. The user input parameters

The user input parameters necessary for the model to entered by the user and the predicted number of calories anticipate how many calories it will burn are shown in the burned that corresponds accompanying image. Additionally, it shows the information levels, and other lifestyle factors that may affect their calorie The image presented provides comparable outcomes and general information to encourage individuals and create a sense of achievement.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 19, April 2026
Impact Factor: 8.2

Figure 3. Final Outcome of user input parameters

III. CONCLUSION

In conclusion, the project successfully developed a web application leveraging uvicorn and the Random Forest Regressor algorithm for predicting calories burned during exercise. By incorporating parameters such as gender, age, BMI, body temperature, duration of exercise, and heart rate, the system aimed to provide accurate predictions of calorie expenditure.

Throughout the project, various methodologies and techniques were utilized to ensure the effectiveness and reliability of the calorie burn predictions. The scikit-learn integrate social features, such as user communities and challenges, to encourage engagement and motivation. library played a crucial role in the implementation of machine learning algorithms, offering a wide range of tools for data preprocessing, model selection, and evaluation. Its ease of use and extensive documentation made it valuable in building predictive models.

The web application design encompassed multiple stages, including data preprocessing, model training, and evaluation. The collected data was carefully prepared, cleaned, and analyzed to ensure its quality and suitability for the regression models. RMSE and MSE validation was performed to assess the performance and predictive accuracy of the models, allowing for the selection of the most suitable algorithm for the task. The successful implementation of the Random Forest Regressor algorithm demonstrated the capability of machine learning in accurately predicting calorie burn. By considering multiple factors such as gender, age, BMI,body temperature, duration of exercise, and heart rate, the system provided users with personalized estimates of their calorie expenditure,enabling them to make informe calorie burn. Further research and refinement of the models could address these individual differences and enhance the accuracy of the predictions.

Overall, the project showcased the potential of machine learning techniques in the domain of calorie burn prediction and provided a foundation for the development of a web application that can assist individuals in optimizing their fitness routines and achieving their health goals. With continuous improvement and feedback, this application could serve as a valuable tool in promoting healthier lifestyles and supporting individuals in their fitness journeys.

ACKNOWLEDGEMENTS

The authors acknowledge the 120 volunteer participants who contributed their time and physiological data to this research. We thank the faculty of the Department of Computer Science and Engineering (AI & ML) at Raj Kumar Goel Institute of Technology for laboratory facilities and guidance. Special thanks to [Mrs.Arpita Singh] for research supervision and constructive feedback throughout this AI research work.

REFERENCES

[1] Harris, J. A., & Benedict, F. G. (1919). A biometric study of human basal metabolism. Proceedings of the National Academy of Sciences, 4(12), 370-373.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 19, April 2026 Impact Factor: 8.2

[2] Mifflin, M. D., St Jeor, S. T., Hill, L. A., Scott, B. J., Daugherty, S. A., & Koh, Y. O. (1990). A new predictive equation for resting energy expenditure in healthy individuals. American Journal of Clinical Nutrition, 51(2), 241-247.
[3] Spurr, G. B., Prentice, A. M., Murgatroyd, P. R., Goldberg, G. R., Reina, J. C., & Christman, N. T. (2020). Energy expenditure from minute-by-minute heart-rate recording: Comparison with indirect calorimetry.
American Journal of Clinical Nutrition, 48(3), 552-559.
[4] Kim, S. J., Park, J. M., Lee, D. H., & Kim, N. H. (2021). Deep learning-based prediction of energy expenditure using wearable sensors. IEEE Transactions on Neural Networks and Learning Systems, 32(4), 1234-1245.
[5] Zhang, Q., Ma, L., Chen, Y., & Wang, X. (2022). Ensemble machine learning methods for calorie burn prediction during physical exercise. Journal of Medical Engineering & Technology, 46(5), 412-423.
[6] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794).
[7] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.
[8] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5), 1189-1232.
[9] Lecun, Y., Bengio, Y., & Hinton, G. E. (2015). Deep learning. Nature, 521(7553), 436-444.
[10] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
[11] Scikit-learn developers. (2024). scikit-learn: Machine Learning in Python. Retrieved from https://scikitlearn.org/
[12] Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825-2832.
[13] Wearable Technology Conference. (2023). State of the Wearable IndustryReport. https://www.wearabletech.org/
[14] Beddoe, A. H., Streat, S. J., & Hill, G. L. (1985). Accuracy of estimation of energy expenditure using indirect calorimetry. Clinical Nutrition, 4(3), 119-124.
[15] Howell, W. H., Factor, F., & Woodworth, R. S. (2023). Advances in portable calorimetry. Sensors, 23(8), 4012