

DR. AMBEDKAR INSTITUTE OF TECHNOLOGY

**(An Autonomous Institution, Afiliated to VTU, Belgaum and
Aided by Government of Karnataka)**

Near Jnana Bharathi Campus, Bangalore-560056



Department of Information Science & Engineering

PROJECT SYNOPSIS

ON

“Calories Burnt Prediction Using Machine Learning”

BACHELOR OF ENGINEERING

IN

INFORMATION SCIENCE AND ENGINEERING

2023-2024

SUBMITTED BY

ROOPITHA G NAYAK [USN: 1DA21IS040]

SHREYAS D [USN: 1DA21IS048]

UNDER THE GUIDANCE OF:

Dr. SATISH B BASAPUR

Asst. Professor, Dept. of ISE

Dr. Ambedkar Institute of Technology.

ABSTRACT

The definition of a calorie is the amount of energy necessary to raise the temperature of 1 gram (g) of water through 1° Celsius.

The type and amount of food people eat determine how many calories they consume. If a person consistently consumes more energy than they require, they will gain weight. Similarly, if they consume less, they will lose weight, fat, and eventually muscle mass. According to the 2020-2025 Dietary Guidelines for Americans, adult females are likely to require between 1,600–2,400 calories a day, and adult males from 2,200–3,000.

However, this depends on their age, size, height, lifestyle, overall health, and activity level. As such, many more factors than just how many calories a person consumes plays a role in maintaining weight.

Calorie burnt prediction by machine learning algorithm” aim to predict the number of calories burnt by an individual during physical activity using machine learning techniques. We collected a dataset that includes features such as heart rate, body temperature, and duration of activity. We used various machine learning models, including XGBoost, linear regression, SVM and random forest, to predict calorie burn based on 15,000 records with seven features. The results we obtain from various models are compared and the most accurate one is chosen for further prediction.

INTRODUCTION

Understanding the concept of calories burned is essential for anyone interested in fitness, weight management, or overall health. The term "calories burned" refers to the energy expenditure of the body during various activities, which includes everything from resting and sleeping to exercising and engaging in daily chores. The estimation of calories burnt during physical activities is crucial for individuals aiming to manage their weight, improve fitness, and maintain a healthy lifestyle. Traditional methods of estimating caloric expenditure often rely on generalized formulas or specific devices like heart rate monitors and pedometers. However, these methods can sometimes lack accuracy and personalization. With the advancement of technology, machine learning (ML) has emerged as a powerful tool to enhance the precision of these estimates by leveraging vast amounts of data and sophisticated algorithms.

Machine learning and deep learning strategies are performed using the python library to predict the total calories burnt extensively using personal information like height, gender, weight, age also the exercise duration etc. and predicting based on the dataset. This will allow us to analyse how various methods that are used to predict, varies in accuracy of the predicted result and help us to choose the best among them i.e most accurate method to find out the calories burnt.

In this project we are using various regression methods which includes decision tree regression, random forest, XG boost, linear regression and ADA boost regression models to predict the burnt calories with respect to the dataset.

We also use python libraries namely pandas, seaborn or matplotlib, numpy, sklearn, xgboost and warnings.

LITERATURE SURVEY

The World Health Organization (WHO) there are many factors that affect the calories burned, but anyone can be modified their diet chart or activity level to get the desired results. There is a study in the literature that used ml and data mining to diagnose problems. When we compare from today's scenario some articles are published earlier with low accuracy of calories burned prediction problems.

Suvarna Shreyas Ratnakar discussed how to predict calories burnt from physical activities. They used the XGB boost Machine learning algorithm to predict it including 15,000 raw dataset and their mean absolute error value is 2.7 and model accuracy is not mentioned. Rachit Kumar Singh et al. illustrated their method to predict calorie burn using machine learning techniques. In their work, logistic regression, linear regression and lasso regression models were used but they didn't mention about mean error absolute value, dataset and model accuracy.

Marte Nipas et al. discussed how to predict burned calories using a supervised learning algorithm. They used a random forest algorithm and gained 95.77% model accuracy. They also used the iterative method to find out the appropriate output from an input. Their work is almost better than other recent work.

Akshit Rajesh Tayade et al. used logistics regression algorithm for diet recommendation system to support mental fitness and physical fitness and accuracy of the proposed model was 85.96.

In summary, these studies demonstrate the potential for machine learning algorithms to predict energy expenditure accurately during physical activity. However, there is still a need for models that can accurately predict energy expenditure across various physical activities and individuals.

OBJECTIVE

- The goal of this model is to predict calories burnt using various models and finding accuracy with respect to each model.
- This model could be used by dieticians to prepare diet plan for their clients based on their scale of fitness.
- This model could be useful for hospitals to track their patient fitness records and diagnose any complications.
- This model could be extremely beneficial to fitness freaks who concentrate mainly on their fitness and help them burn calories as per their plan.

EXISTING SYSTEM

The study on calories burnt prediction using linear regression, decision tree regressor, random forest regressor to predict calories burnt.

Techniques to predict calorie from input images. They used some digital image processing techniques such as image acquisition, RGB conversion, feature extraction and image enhancement so on.

Logistics regression algorithm for diet recommendation system to support mental fitness and physical fitness.

Machine learning algorithms such as XGBboost regressor and Linear regression models to find out calorie burnt in physical activities.

PROPOSED SYSTEM

XGBoost is an optimized distributed gradient boosting library designed for efficient and scalable training of machine learning models. It is an ensemble learning method that combines the predictions of multiple weak models to produce a stronger prediction. XGBoost stands for “Extreme Gradient Boosting” and it has become one of the most popular and widely used machine learning algorithms due to its ability to handle large datasets and its ability to achieve state-of-the-art performance in many machine learning tasks such as classification and regression.

AdaBoost, short for Adaptive Boosting, is an ensemble machine learning algorithm that can be used in a wide variety of classification and regression tasks. It is a supervised learning algorithm that is used to classify data by combining multiple weak or base learners (e.g., decision trees) into a strong learner.

XGBoost, ADABOOST models are used to enhance the accuracy levels and accurately predict Calories burnt.

To attain more accurate prediction, we will use more than 15000 records and try to avoid overfitting of data.

We will compare and show different models with varied accuracy and finally choose the most accurate model for further prediction.

SYSTEM REQUIREMENTS SPECIFICATION

Hardware required:

- ✓ Computer system(laptop)
- ✓ AMD Ryzen 5 5000 series with 16 GB of Ram and 3.3 GHz processor speed
- ✓ Power supply
- ✓ Software required: Anaconda3, Jupyter Notebook, Pycharm.

Operating system:

- ✓ Windows
- ✓ Anaconda3
- ✓ Python modules like Pandas, Flask framework, numpy, seaborn, matplotlib, warnings, xgboost, numpy and Sklearn
- ✓ Dataset collected from GitHub, Kaggle and some added based on our activities.