**Machine Learning in Action: Predicting Calorie Burn for Optimal Health and Fitness**

Bethapudi Chinni Raja ,

UG Student

Dept of AI&DS

Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru,

Andhra Pradesh, India

[chinniraja9999@gmail.com](mailto:chinniraja9999@gmail.com)

Buridi Revanth Naga Manikanta,

UG Student

Dept of AI&DS

Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru,

Andhra Pradesh, India

[revanthburidi@gmail.com](mailto:revanthburidi@gmail.com)

Aluri Leela Subhash Chandra,

UG Student

Dept of AI&DS

Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru,

Andhra Pradesh, India

[subhashchowdary09@gmail.com](mailto:subhashchowdary09@gmail.com)

Ashok Reddy Kandula,

Assistant professor

Dept of AI&DS

Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru,

Andhra Pradesh, India

<ashokreddy.gec@gmail.com>

Chinnam Abhishek,

UG Student

Dept of AI&DS

Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru,

Andhra Pradesh, India

[chinnamabhishek09@gmail.com](mailto:chinnamabhishek09@gmail.com)

## **Abstract***—*Efficiently tracking and controlling caloric intake is essential for sustaining a healthy lifestyle and reaching fitness objectives in today's health-conscious culture. The goal of this research is to employ regression machine learning techniques to create a predictive model for calorie burn prediction. To properly predict the amount of calories burnt during a given activity, the model analyses a variety of input data, including the kind, duration, and intensity of physical activity as well as individual factors like age, weight, and gender. With the use of this model's insights, people, fitness enthusiasts, and medical experts may better control their calories and customise their physical activity regimens to achieve their fitness objectives. Additionally, the predictive model may provide tailored advice on how to increase exercise efficacy, which will enhance health outcomes. This strategy not only promotes a more informed approach to fitness but also contributes to the broader goal of promoting healthier lifestyles through data-driven decision-making.

## **Keywords**Calorie Burn Prediction, Machine Learning, Regression Techniques, Personalized Fitness, Physical Activity.

## **I.INTRODUCTION**

In today's world, where health and fitness are increasingly prioritized, understanding and managing calorie expenditure has become essential for many individuals striving to maintain or improve their well-being. Calorie management is at the core of various fitness goals, whether it’s weight loss, muscle gain, or simply maintaining a healthy lifestyle. Traditional methods of estimating calories burned during physical activities often rely on general approximations that fail to consider individual differences. As a result, they may provide inaccurate information, leading to ineffective fitness planning. This growing need for precision in fitness guidance has driven the development of more advanced methods to estimate calorie expenditure, bringing data science and machine learning to the forefront of health and wellness applications.

Machine learning, particularly regression techniques, offers a powerful approach to modeling and predicting calorie burn. By leveraging large datasets, machine learning models can identify patterns and relationships between different factors that influence calorie expenditure. These factors include the type of physical activity, its duration and intensity, as well as individual characteristics such as age, weight, gender, and fitness level. Through these variables, machine learning algorithms can create personalized predictions that are far more accurate than traditional methods. This innovation empowers users to make informed decisions about their fitness routines, leading to more effective and efficient outcomes.

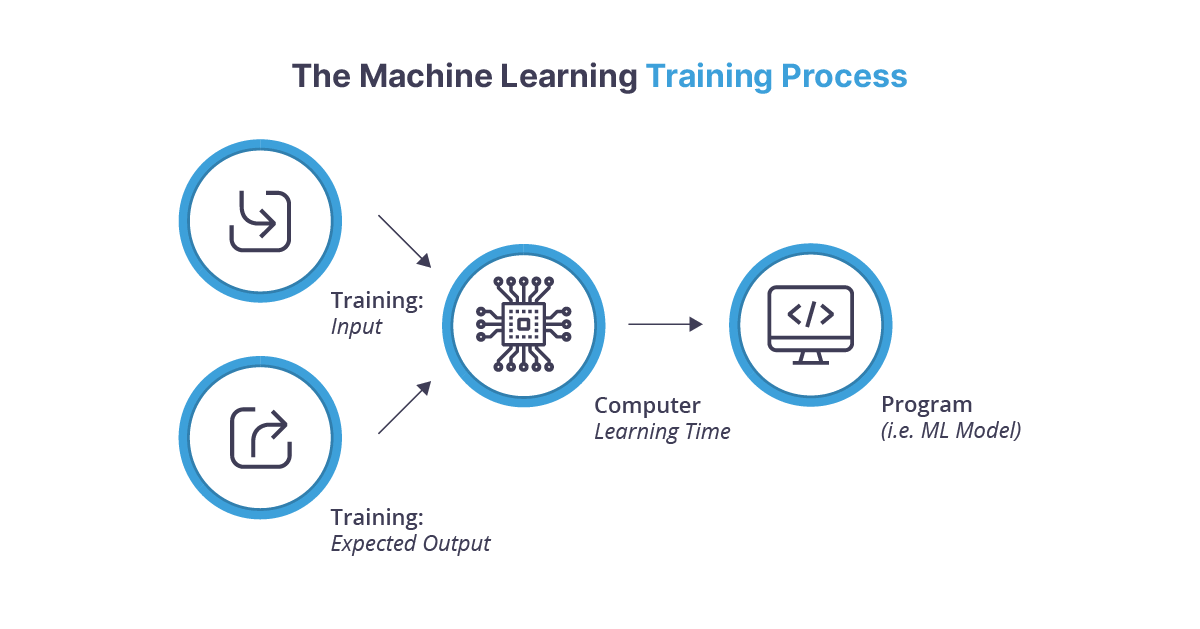


Fig1:Machine Learning

The development of a predictive model for calorie burn is a complex task that requires careful selection and preprocessing of data. Datasets that capture a variety of physical activities, along with detailed personal attributes, are necessary for training robust models. Once the data is collected, it undergoes preprocessing steps such as normalization, handling missing values, and feature selection to ensure the model performs optimally. The choice of regression algorithms—such as linear regression, decision trees, or more advanced methods like gradient boosting and neural networks—plays a crucial role in the accuracy of the predictions. These models are evaluated using metrics like Mean Squared Error (MSE) or R-squared to ensure their reliability in predicting calorie burn across different scenarios.

A predictive model of this nature can have significant implications for fitness enthusiasts, healthcare professionals, and the broader wellness industry. For individuals, it offers a personalized approach to managing their calorie intake and expenditure, allowing for more effective workout routines and dietary planning. Fitness professionals can use these insights to tailor training programs that align with their clients' goals, while healthcare providers may utilize the model to assist in rehabilitation and recovery plans that require specific activity levels. Moreover, this model can be integrated into fitness apps and wearable devices, providing real-time calorie burn predictions that enhance user experience and engagement.

**1.1Motivatio*n***

The increasing need for precise and customised exercise advice is what inspired this idea. The accuracy of traditional calorie burn estimation methods is frequently lacking, resulting in less than ideal health effects. The increasing popularity of data-driven health solutions and wearable fitness technologies calls for the development of more dependable models that take individual parameters like age, weight, and activity level into account. This research intends to close that gap by utilising machine learning to provide users personalised calorie burn projections to maximise exercise regimens, improve health management, and ultimately facilitate better and more informed lifestyle choices.

**1.2 Objectives:**

* Build a machine learning regression model that accurately predicts calories burned during physical activities by analyzing input features such as activity type, duration, intensity, age, weight, and gender.
* Provide personalized insights into calorie expenditure that can be tailored to individual needs, helping users create optimized workout routines and dietary plans based on their specific characteristics and fitness goals.
* Leverage data science and machine learning techniques to enhance the precision of calorie management systems, making them more reliable than traditional estimation methods and contributing to better health outcomes.
* Design the model to be adaptable for integration into fitness apps and wearable devices, offering real-time calorie burn predictions and insights to users, thereby improving user engagement and experience.
* Continuously refine the model by experimenting with different regression algorithms, feature engineering techniques, and evaluation metrics to achieve the highest possible accuracy and reliability in various scenarios.

**1.3Contributions of Proposed System:**

The proposed system contributes to the field of fitness and health management by introducing a machine learning-based predictive model for accurately estimating calories burned during physical activities. Unlike traditional estimation methods, which often rely on generalized formulas, this system leverages regression techniques to account for individual-specific factors such as age, weight, gender, and the type, duration, and intensity of physical activities. By providing personalized calorie burn predictions, the system empowers users to make more informed decisions regarding their workout routines and dietary plans, ultimately leading to better health and fitness outcomes. This approach bridges the gap between generic estimations and personalized fitness guidance, offering a more precise tool for calorie management.

Additionally, the system’s flexibility allows for integration with modern wearable devices and fitness applications, enabling real-time predictions and recommendations for users. This enhances user engagement and helps individuals track their progress more effectively, fostering a data-driven approach to achieving fitness goals. The model's adaptability ensures that it can be continuously refined and optimized as new data becomes available, keeping pace with advancements in the health and fitness industry. By contributing to a more personalized, data-driven approach to fitness, the proposed system has the potential to improve the overall effectiveness of physical activity planning and health management.

**II.RELATED WORK**

More complex techniques for estimating calorie burn have recently been developed thanks to developments in machine learning, with an emphasis on enhancing accuracy through data-driven techniques[5]. In their comparison study of many machine learning methods for calorie burn prediction, Vijayalakshmi and Sridurga emphasised the variations in model performance between decision trees, random forests, and support vector machines. In general, ensemble approaches fared better in terms of accuracy and resilience than individual models, according to their research, which emphasises the significance of algorithm selection in the construction of trustworthy predictive models. This work shows how machine learning can be used to improve calorie prediction systems and offers insightful information about which algorithms work best with various kinds of datasets.

Vinoy and Joseph examined the use of machine learning techniques in further detail by examining the XGBoost and linear regression algorithms for calorie burn prediction[6]. Their research showed that while linear regression offers simplicity and interpretability, more complex algorithms like XGBoost can provide higher accuracy due to their ability to capture non-linear relationships between input features and calorie expenditure. The findings suggest that advanced algorithms like XGBoost are particularly effective in scenarios where the data exhibits complex interactions, making them a promising choice for real-world applications in calorie prediction.

Ratnakar and V. S. expanded on this by examining the broader application of machine learning models for calorie burn prediction across different datasets. Their study demonstrated the versatility of these models in handling diverse data inputs, such as physical activity type, duration, intensity, and personal attributes like age, weight, and gender[7]. By leveraging feature engineering and optimizing hyperparameters, their work showed significant improvements in prediction accuracy. Collectively, these studies underscore the potential of machine learning to transform traditional calorie estimation methods, providing more accurate, personalized predictions that can be integrated into health and fitness technologies.

Calorie burn estimation has long been a key focus in health and fitness research, with various approaches developed to provide insights into energy expenditure during physical activities[1]. Bubnis discusses traditional methods for calculating daily calorie burn, which typically involve estimating basal metabolic rate (BMR) and accounting for physical activities throughout the day. These methods, while helpful, often rely on generic formulas that fail to consider individual differences in metabolism, body composition, and activity intensity. As a result, there is a growing need for more accurate, personalized methods that can account for the unique characteristics of each individual.

Furthering this understanding, a study conducted by Kansas State University highlights the impact of social factors on calorie expenditure, suggesting that individuals are more likely to burn more calories when working out with someone they perceive as being better or more skilled[2]. This finding introduces an important psychological dimension to physical activity, demonstrating that factors beyond physiological metrics can influence calorie burn. While traditional models often overlook such variables, they are critical to understanding the full picture of energy expenditure and developing more comprehensive predictive models.

Tingley adds a contemporary perspective by exploring how advancements in science have deepened our understanding of calorie burn and its variability among individuals[3]. The research emphasizes the role of age, weight, and other biological factors in determining how the body processes energy. Tingley’s work underscores the limitations of one-size-fits-all calorie estimation methods, advocating for personalized approaches driven by data. This growing body of literature reflects the ongoing evolution in the field, where machine learning and data-driven techniques are poised to provide more accurate and customized predictions for calorie expenditure.

**2.1Machine Learning Algorithm:**

Here are the Machine Learning Algorithm that are used.

**2.2.1SVR Model:**

A potent class of supervised machine learning techniques, support vector machines are frequently used for regression and classification problems. The primary goal of support vector machines is to identify the best hyperplane for efficiently dividing a dataset's classes. The margin, or the distance between the hyperplane and the closest data points from each class, is maximised by positioning this hyperplane in that manner.

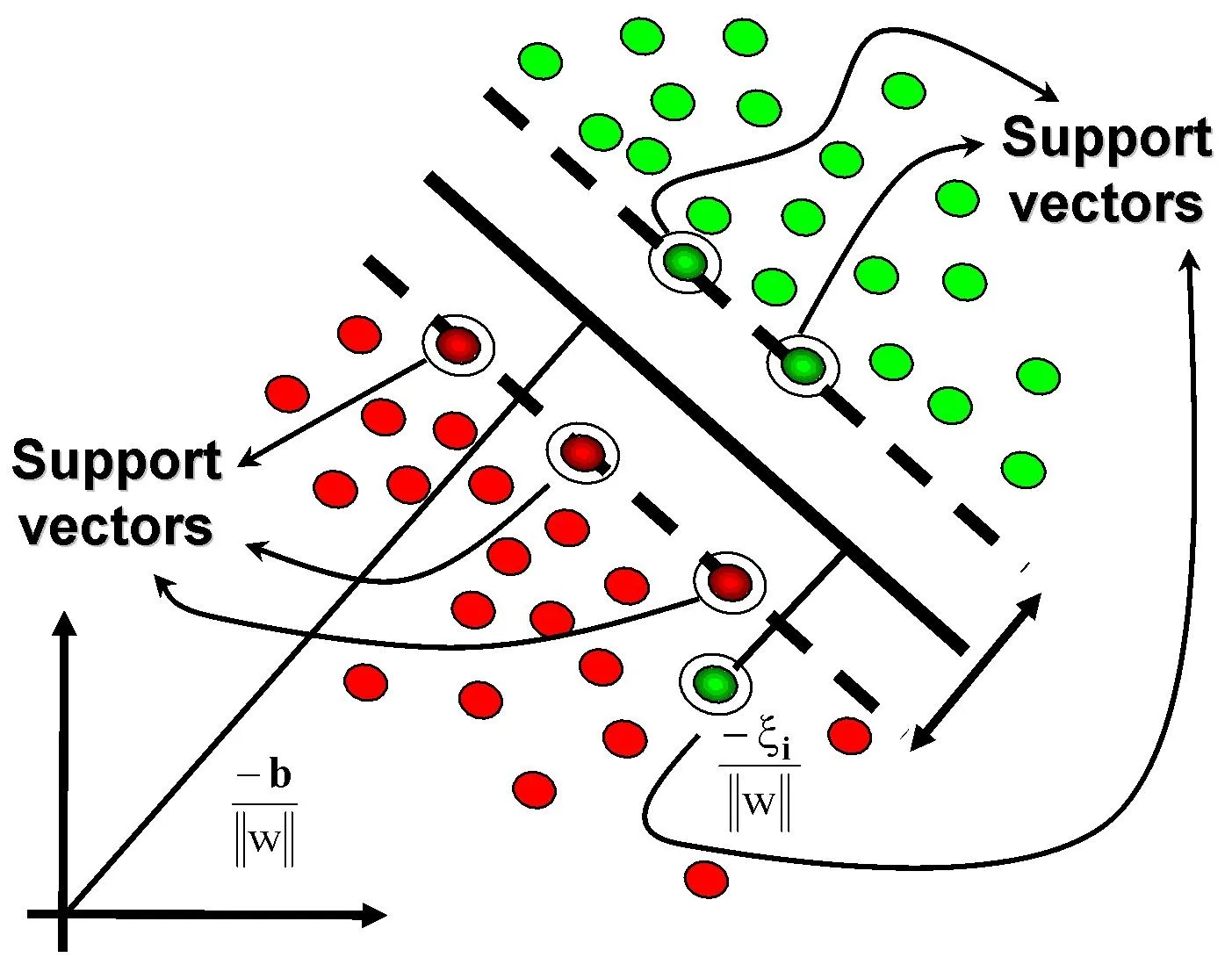
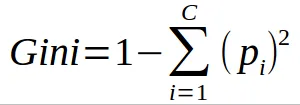


Fig2:SVM

The process of creating a hyperplane in Support Vector Machines requires taking into account four important components: the data points (x), class labels (y), weight vector (w), and bias (b). The features of the dataset, such as pixel values in image classification tasks, are represented by the data points (x). Class labels (y) show which categories or classes have been applied to each data item; for instance, classifying photographs featuring dogs or cats differently. Representing the coefficients allocated to each feature, the weight vector (w) is essential in setting the orientation of the hyperplane. Finding and estimating the ideal weight vector that maximises the margin—the space between the hyperplane and the closest data points from each class is the main goal of support vector machines.

**2.2.2 Random Forest Regressor:**

An ensemble learning method that is frequently used for both classification and regression applications is the Random Forest algorithm. The Random Forest, which consists of several decision trees, combines their forecasts to improve overall precision and resilience[11]. Each tree is built during training using a random subset of the dataset, and a random subset of characteristics is taken into consideration for splitting at each node. Because of its randomness, the model is less prone to overfitting and is better able to generalise to new, untested data.



The Gini index is an important decision tree node splitting metric in Random Forests classification. The Gini index takes into account the distribution of classes among the data points to quantify the impurity or disorder inside a node[14]. A purer node with a more homogeneous class distribution is indicated by lower Gini values, which direct the decision tree to make splits that are effective. By merging the advantages of several decision trees, this iterative procedure is used to build trees in a Random Forest, improving the ensemble's capacity for reliable and accurate prediction-making.

**2.2.3:Decision Tree Regressor:**

A supervised machine learning approach for classification and regression problems is called a decision tree. Recursively dividing the dataset into subgroups according to the most important attribute at each decision node is how it operates[8]. The objective is to build a structure like a tree, with each leaf node standing in for the anticipated result. Recursive binary splitting is the method used to create decision trees, in which the dataset is divided into two subsets according to a selected feature and a threshold value. Criteria like mean squared error for regression and Gini impurity for classification serve as a roadmap for decision-making at each node. Decision trees can capture complicated relationships in the data and are easily interpretable and understood.

**2.2.4 Lasso And Ridge Regression:**

Lasso and Ridge regression are two fundamental techniques in linear regression that address the challenges of multicollinearity and overfitting, each with distinct approaches to regularization. **Ridge regression**, also known as Tikhonov regularization, applies a penalty proportional to the square of the magnitude of the coefficients (L2 norm), which helps to shrink the coefficients of less important features and stabilize the model. This approach is particularly effective when dealing with multicollinearity, as it reduces the variance of the model and improves its generalization by incorporating all features while controlling their influence.

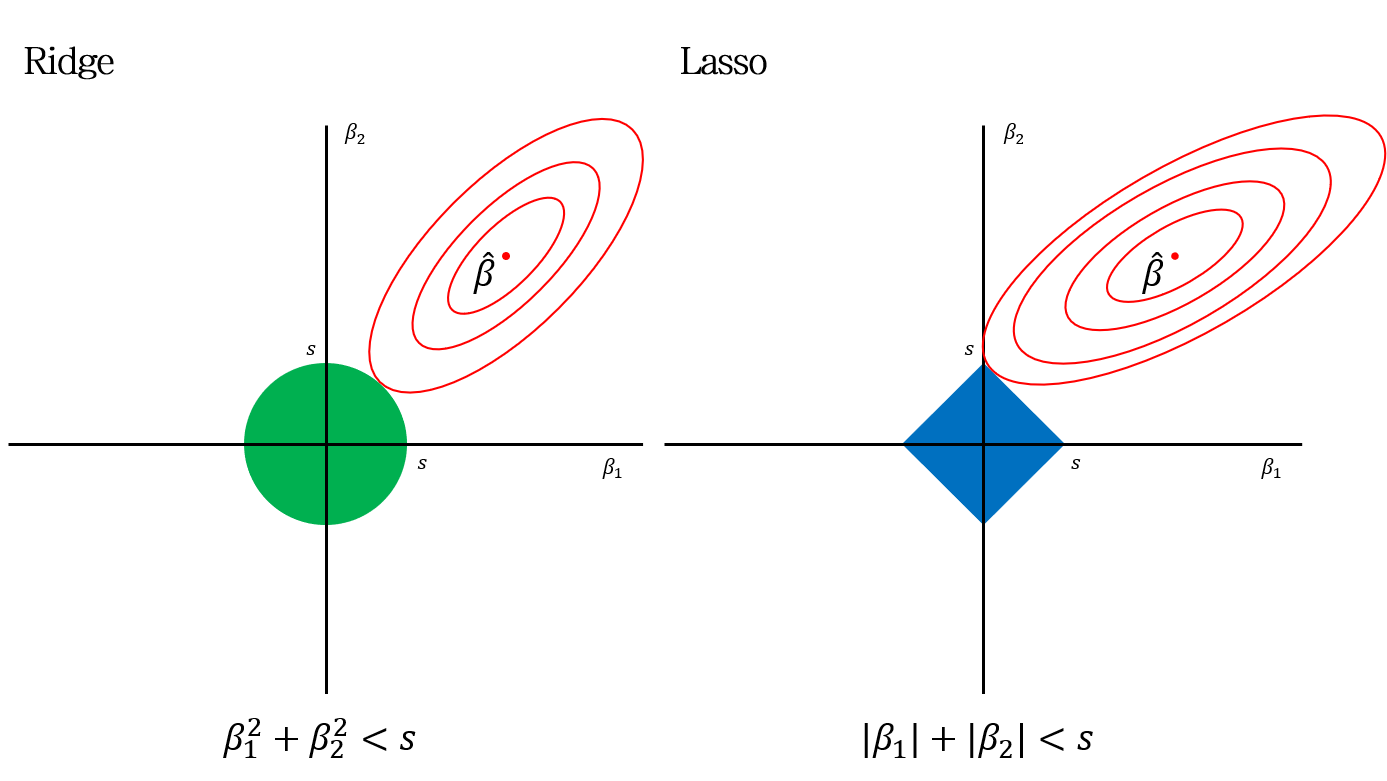


Fig3:Ridge nad Lasso

**Lasso regression**, on the other hand, applies a penalty proportional to the absolute value of the coefficients (L1 norm), leading to some coefficients being exactly zero. This characteristic not only reduces overfitting by penalizing large coefficients but also performs feature selection, making it useful for models where feature reduction is desired.

**III.PROPOSED METHODOLOGY:**

The proposed method for predicting calorie burn employs a combination of regression machine learning techniques to create a robust predictive model that accounts for various factors influencing energy expenditure. The methodology begins with the collection and preprocessing of data, which includes physical activity type, duration, intensity, and individual characteristics such as age, weight, and gender. This data is sourced from comprehensive datasets that capture a wide range of activities and demographic information, ensuring that the model is trained on diverse and representative samples. Preprocessing involves cleaning the data to handle missing values, normalizing numerical features, and encoding categorical variables to prepare the dataset for effective model training.

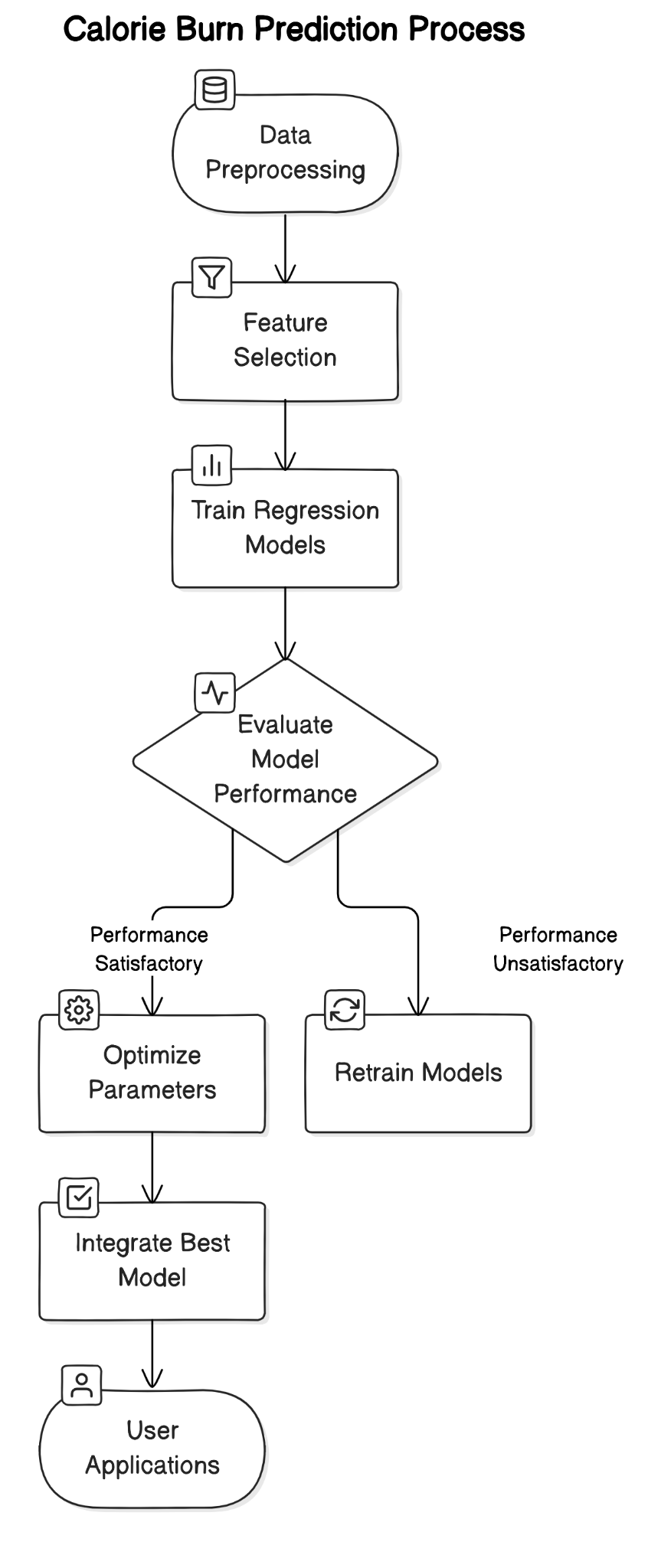


Fig4:Proposal Diagram

Next, feature selection is performed to identify the most relevant variables that contribute to calorie burn prediction. This step is crucial for reducing the dimensionality of the dataset and improving the model's performance. Techniques such as correlation analysis, Recursive Feature Elimination (RFE), and feature importance scores from preliminary models are used to select a subset of features that have the highest impact on calorie expenditure. By focusing on the most influential features, the model is better equipped to make accurate predictions and avoid overfitting to irrelevant or redundant data.

The core of the proposed method involves training and evaluating various regression models, including linear regression, Ridge regression, Lasso regression, and more advanced techniques like Gradient Boosting and XGBoost. Each model is trained on the prepared dataset, and its performance is evaluated using metrics such as Mean Squared Error (MSE), R-squared, and Cross-Validation scores. This comparative analysis helps in selecting the best-performing model based on accuracy and generalization capability. Additionally, hyperparameter tuning is conducted using techniques like Grid Search or Random Search to optimize the model's parameters and enhance its predictive power.

Once the optimal model is identified, it is validated using a separate test set to ensure its reliability and robustness in real-world scenarios. The validation process involves assessing the model's performance on unseen data, checking for potential issues like overfitting or underfitting, and ensuring that the predictions are consistent with actual calorie expenditure. Post-validation, the model is fine-tuned and adjusted based on feedback and performance metrics to achieve the highest possible accuracy.

Finally, the predictive model is integrated into a user-friendly application or platform, such as a fitness app or wearable device, to provide real-time calorie burn predictions and personalized recommendations. This integration includes designing an intuitive user interface, ensuring seamless data input and output, and implementing features that allow users to track their progress and adjust their fitness plans accordingly. The goal is to offer a practical tool that empowers individuals to manage their calorie intake and physical activities effectively, ultimately supporting healthier lifestyle choices and more informed fitness decisions.

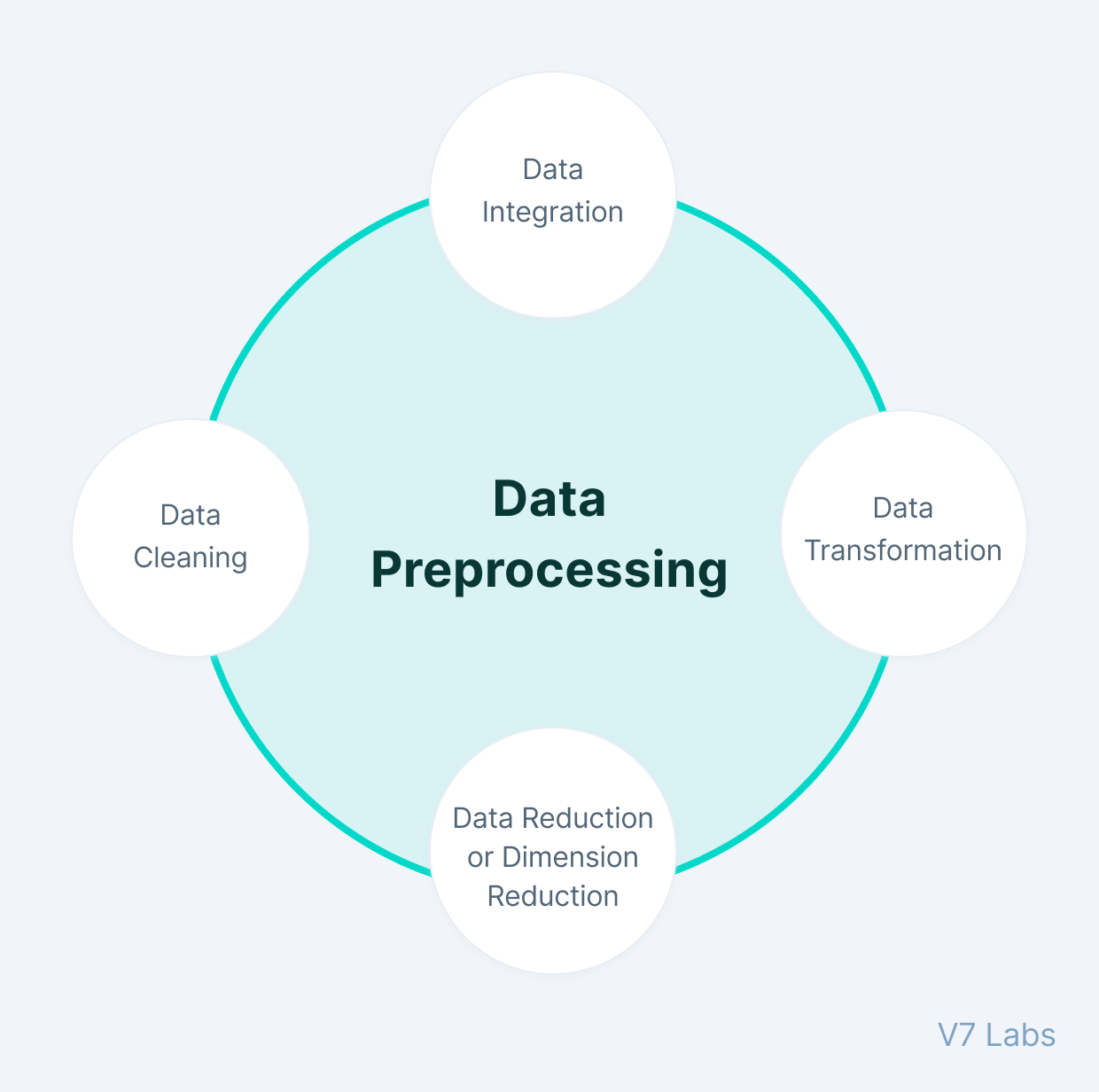
**3.1Dataset Collection:**

Data collection for the calorie burn prediction model involves gathering comprehensive records that include individual identifiers (User\_ID), personal attributes (Gender, Age, Height, Weight), and exercise-related metrics (Duration, Heart\_Rate, Body\_Temp). The dataset captures the total calories burned by each user (Calories) during various physical activities. It integrates data from diverse sources, ensuring a wide range of activity types and demographic profiles. By combining personal and exercise-specific data, this approach aims to create a robust model that accurately predicts calorie expenditure, tailored to individual characteristics and activity levels, ultimately enhancing fitness and health management..

**3.2Data Preparation:**

Data preparation involves several key steps to ensure the dataset is ready for model training. Initially, missing or inconsistent values are addressed through imputation or removal to maintain data quality. Next, categorical variables, such as Gender, are encoded into numerical formats to facilitate model processing. Numerical features like Age, Height, Weight, Duration, Heart\_Rate, and Body\_Temp are normalized or standardized to ensure uniform scaling. Feature selection techniques are then applied to identify and retain the most relevant variables for predicting calorie burn. Finally, the dataset is split into training and test subsets to evaluate model performance effectively.

**3.2.1:Data Preprocessing:**

Data preprocessing involves cleaning and transforming raw data to prepare it for analysis. This includes handling missing values through imputation or exclusion, encoding categorical variables like Gender into numerical formats, and normalizing or standardizing numerical features such as Age, Height, Weight, Duration, Heart\_Rate, and Body\_Temp. Outliers are identified and managed to prevent skewed results. Feature scaling ensures that all variables contribute equally to the model. Additionally, feature engineering may be applied to create new variables or interactions that improve model performance. Finally, data is split into training and test sets to validate model accuracy and generalization. F Fig 5:Data preprocessing

Data visualization of gender distribution involves creating visual representations, such as bar charts or pie charts, to illustrate the proportion of different genders within the dataset. This helps in understanding the demographic makeup of the participants and identifying any imbalances or biases. For instance, a bar chart may display the count or percentage of male versus female participants, offering a clear view of the gender distribution. This visualization is crucial for ensuring diverse representation in the data, which can impact the model’s performance and generalizability. By examining gender distribution, insights into potential data biases and their implications on predictive accuracy can be gained.

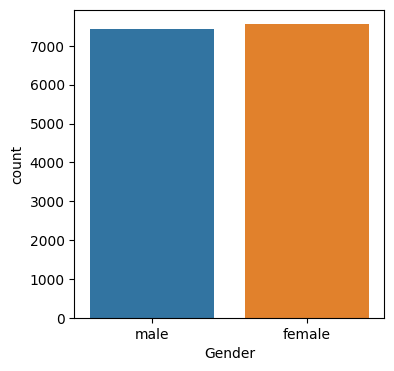


Fig7:Gender Distribution

**3.3 Model Training**:

To forecast calorie burn, several regression models are built and trained. This process is known as model construction. Assuming a linear connection between characteristics and target, Linear Regression creates a baseline by fitting a straight line to the data. While Ridge Regression use L2 regularisation to reduce coefficients and handle multicollinearity, Lasso Regression boosts sparsity in feature selection by adding L1 regularisation. Both techniques assist in managing overfitting and enhancing model generalisation.

Moreover, Support Vector Regression (SVR) maximises the margin of error through the use of support vector machines to handle non-linear connections and outliers. Decision Tree Regressor divides the data according to feature values to construct a model that has a structure like a tree for predictions. In order to improve accuracy and resilience, decrease overfitting, and boost predictive performance, the Random Forest Regressor combines many decision trees. All of them model is evaluated for accuracy and adjusted as needed to optimize performance.

**IV.RESULTS**

**6.1 Evaluation and Metrics**

A classifier's standard evaluation is based on a number of predefined performance indicators. Our models are assessed using the following metrics: Accuracy, Specificity, F-score, Precision, and Recall.

First, we used a confusion matrix and made inferences from it for the classification algorithms. A confusion matrix is a table that, given a set of test data for which the real values are known, provides details on the quality or performance of a model for two or more types of classes. The simplest confusion matrix for the classifier 2 class, as seen in Figure 5, is a two-dimensional confusion matrix.

The plot generated using `scipy.stats.probplot` provides a graphical assessment of the normality of the residuals from the model. By comparing the quantiles of the residuals (`test\_res`) against the theoretical quantiles of a normal distribution, the plot visually indicates whether the residuals follow a normal distribution. In this plot, if the data points closely align with the reference line, it suggests that the residuals are normally distributed, which is a key assumption for many statistical models. Deviations from the line may indicate deviations from normality, suggesting potential issues with model fit or residual behavior.

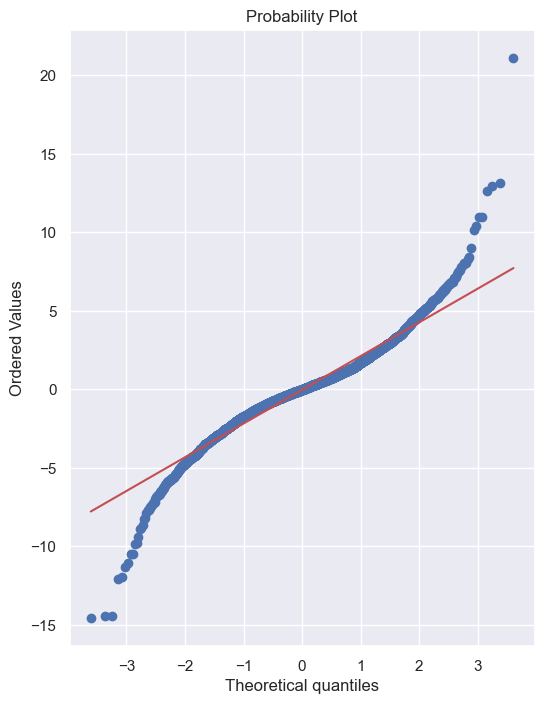


Fig6:Probability Plot

The heatmap provides insights into the relationships between various factors by visualising the correlation between distinct elements in the calorie burn prediction dataset. The heatmap finds patterns and associations between data like Duration, Heart\_Rate, and Body\_Temp by utilising colour gradients to show the intensity and direction of correlations. The characteristics that exhibit strong associations with both calorie expenditure and each other are highlighted. This data is essential for feature engineering and selection since it directs the selection of variables to incorporate into the model, hence increasing the predictability and interpretability of the calorie burn estimates.

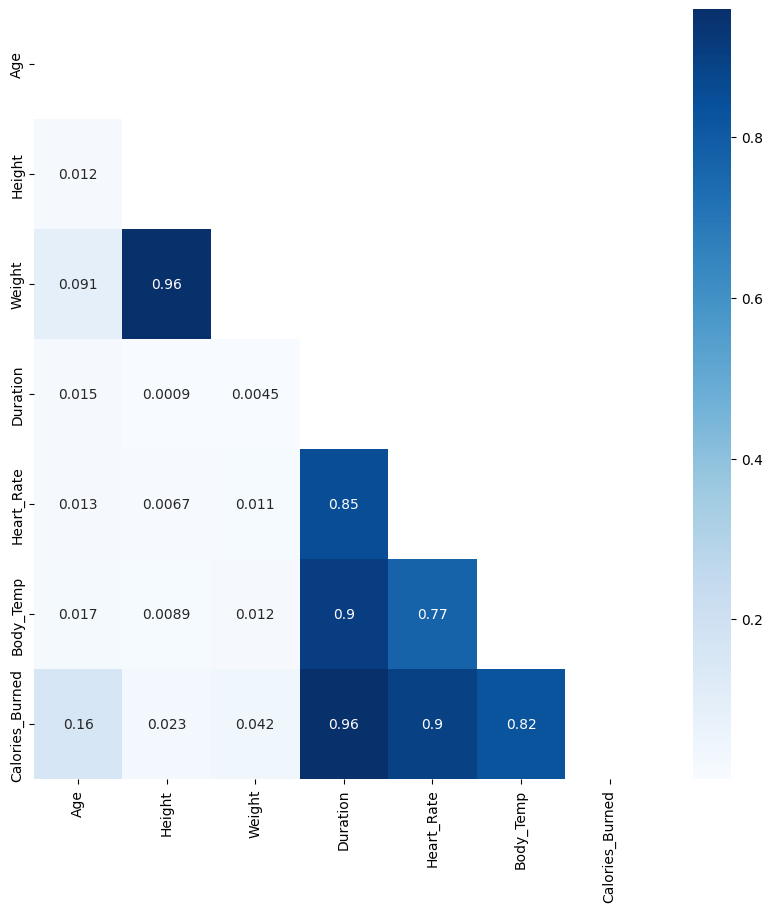


Fig7:Correlation HeatMap

The performance of various regression models in predicting calorie burn reveals significant differences in accuracy and effectiveness. Linear Regression and Ridge Regression exhibit similar performance, with Mean Absolute Errors (MAE) of approximately 8.27 and 8.27, respectively, and Mean Squared Errors (MSE) around 125.39. Their Root Mean Squared Errors (RMSE) are also close, about 11.20, indicating a reasonably good fit for the data. Both models achieve high R-squared values, around 0.97, reflecting their strong ability to explain the variance in calorie burn. Despite their good performance, their predictive accuracy is surpassed by more advanced models.

Lasso Regression performs slightly worse compared to Linear and Ridge Regression, with a higher MAE of 9.09 and MSE of 150.83. Its RMSE is also higher at 12.28, which indicates a greater degree of prediction error. However, it still maintains a high R-squared value of 0.96, suggesting that while it may be less accurate than Linear and Ridge Regression, it is still effective in capturing the relationship between features and calorie expenditure. The higher error metrics in Lasso Regression can be attributed to its feature selection properties, which might lead to a loss of some predictive accuracy in exchange for model simplicity.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Mean Absolute Error** | **Mean Squared Error** | **Root Mean Squared Error** | **R-squared (R²)** |
| Linear Regression | 8.272389 | 125.390475 | 11.197789 | 0.967390 |
| Lasso | 9.094562 | 150.833545 | 12.281431 | 0.960773 |
| Ridge | 8.272578 | 125.401244 | 11.198270 | 0.967387 |
| SVR | 2.248260 | 29.640191 | 5.444281 | 0.992291 |
| Decision Tree Regressor | 3.366667 | 26.203333 | 5.118919 | 0.993185 |
| Random Forest Regressor | 1.696080 | 6.991670 | 2.644177 | 0.998182 |

Table 1:Comparison

In contrast, SVR, Decision Tree Regressor, and Random Forest Regressor show superior performance with lower error metrics and higher R-squared values. SVR has an MAE of 2.25 and MSE of 29.64, with an RMSE of 5.44 and an R-squared of 0.99. The Decision Tree Regressor follows closely with an MAE of 3.37, MSE of 26.20, and RMSE of 5.12, achieving an R-squared value of 0.99. The Random Forest Regressor outperforms all other models with the lowest MAE of 1.70, MSE of 6.99, RMSE of 2.64, and the highest R-squared value of 0.998. These metrics indicate that ensemble methods like Random Forest provide the most accurate and reliable predictions for calorie burn, effectively handling complex patterns and interactions in the data.

**V. Conclusion:**

The study shows that, in comparison to conventional techniques, sophisticated regression models greatly increase the accuracy of calorie burn prediction. The model that performed the best was the Random Forest Regressor, which also had the greatest R-squared value and the lowest Mean Absolute Error, Mean Squared Error, and Root Mean Squared Error. The robustness of SVR and Decision Tree Regressor in capturing intricate linkages is demonstrated by their strong performance. Although Lasso, Ridge, and Linear regression models work well, their prediction accuracy is somewhat lower. All things considered, our results emphasise how crucial it is to use the right models and methods for accurate calorie burn calculations, improving fitness and health care.

**VI. Future Scope:**

Future research might concentrate on improving the calorie burn model's predicted accuracy by adding other data sources, such metabolic rates and physiological measures taken in real time by wearables. Examining more complex machine learning strategies, including ensemble approaches or deep learning, may help the model perform better still. Furthermore, by including user input and personalised fitness data, forecasts may be more individually tailored, improving the model's applicability and usefulness. Generalisability would also be enhanced by expanding the dataset to cover a variety of exercise kinds and varied demographics. Lastly, creating a user-friendly application with the ability to make predictions in real time might offer healthcare professionals and fitness enthusiasts useful information.

**VII. References:**

[1]D. Bubnis, "Calculating how many calories are burned in a day," *Medical News Today*, 01 January 2020.

[2]K. S. University, "Burning more calories is easier when working out with someone you perceive as better," 26 November 2012.

[3]B. K. Tingley, "The New Science on How We Burn Calories," *The New York Times Magazine*, 14 September 2021.

[4]S. T and V. K, "PREDICTION OF USER’S CALORIE ROUTINE USING CONVOLUTIONAL NEURAL NETWORK," *International Journal of Engineering Applied Sciences and Technology*, vol. 5, no. 3, pp. 189-195, 2020.

[5]G. Vijayalakshmi and T. Sridurga, "COMPARING MACHINE LEARNING ALGORITHMS FOR PREDICTING CALORIES BURNED," *Journal of Emerging Technologies and Innovative Research (JETIR)*, vol. 10, no. 3, pp. 519-527, March 2023.

[6]S. P. Vinoy and B. Joseph, "Calorie Burn Prediction Analysis Using XGBoost Regressor and Linear Regression Algorithms," in *Proceedings of the National Conference on Emerging Computer Applications (NCECA)*, Kottayam, 2022.

[7]S. S. Ratnakar and V. S, "Calorie Burn Prediction using Machine Learning," *International Advanced Research Journal in Science, Engineering and Technology*, vol. 9, no. 6, pp. 781-787, June 2022.

[8]R. K. Singh and V. Gupta, "Calories Burnt Prediction Using Machine Learning," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 11, no. 5, May 2022.

[9]M. Nipas, A. G. Acoba, J. N. Mindoro, M. A. F. Malbog, J. A. B. Susa and J. S. Gulmatico, "Burned Calories Prediction using Supervised Machine Learning: Regression Algorithm," in *2022 Second International Conference on Power, Control and Computing Technologies (ICPC2T)*, Raipur, India, March 2022.

[10]R. S. Biyani and M. S. Nandini, "CALORIES PREDICTION BASED ON FOOD IMAGES," *International Research Journal of Engineering and Technology (IRJET)*, vol. 7, no. 8, pp. 2122-2125, August 2020.

[11]K. Westerterp, "Control of energy expenditure in humans," *European Journal of Clinical Nutrition*, vol. 71, pp. 340-344, 30 November 2016.

"calories\_burnt\_data," Kaggle, 2022.

[12]K. K. Al-jabery, T. Obafemi-Ajay, G. R, Olbricht and D. C. W. II, "Computational Learning Approaches to Data Analytics in Biomedical Applications," *ACADEMIC PRESS*, 2020, pp. 7-27.

[13]K. Nighania, "Various ways to evaluate a machine learning model’s performance," *Towards Data Science*, 30 December 2018.

[14]T. Z. Phyu and N. N. Oo, "Performance Comparison of Feature Selection Methods," in *MATEC Web of Conferences*, EDP Sciences, 2016.