Model Optimization for Personalized Health Metrics Analysis

Madhuranga Perera
Dept of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21159930@my.sliit.lk

Amasha Wijesiriwardena
Dept of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21182082@my.sliit.lk

Amandi Pathirana
Dept of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21181306@my.sliit.lk

Lakshan Gamaathige
Dept of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it21182532@my.sliit.lk

Pipuni Wijesiri
Dept of Computer Systems Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
pipuniwijesiri@gmail.com

Anuradha Jayakody
Dept of Computer Systems Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
anuradha.j@sliit.lk

2025 5th International Conference on Advanced Research in Computing (ICARC) | 979-8-3315-3098-3/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICARC64760.2025.10963304

Abstract— **This paper investigates the development and application of four machine learning models designed to enhance personalized health management, specifically targeting young adults aged 15–30. The research addresses common health challenges, such as obesity and lifestyle-induced diseases, through data-driven methodologies that provide personalized meal plans, workout recommendations, and progress monitoring. The first model generates optimized personalized recommendations according to the user's health condition using Random Forest and Decision Tree algorithms. The second model utilizes an ensemble of Random Forest, LightGBM, and XGBoost, combined through a stacking technique with Linear Regression as the meta-model, to generate optimized personalized meal plans according to health condition. The third model generates optimized workout plans using Gradient Boosting and XGBoost classifiers, accounting for individual fitness objectives, body compositions, and medical conditions. A fourth model predicts goal achievement timelines by analyzing features such as caloric balance and hydration efficiency, providing users with actionable feedback using XGBoost. The integration of these AI-driven components into a scalable digital platform demonstrates the potential of machine learning in transforming health management. Future enhancements include improving model accuracy, enabling real-time feedback, and deploying the system as an accessible mobile application.**

Keywords— Personalized, ensemble, prediction, machine learning, recommendation.

I. Introduction

This research focuses on the development and evaluation of four machine learning models designed to optimize personalized meal planning, workout recommendations, progress tracking, and health forecasting. Each model incorporates unique methodologies to deliver precise and actionable outputs, leveraging data such as user demographics, activity metrics, and health conditions. By employing ensemble techniques, hyperparameter optimization, and robust validation strategies, these models achieve high accuracy and reliability in their predictions. The objective of this paper is to present the technical details and performance evaluation of the machine learning models developed for each component. The discussion is limited to the algorithms, preprocessing methods, and optimization techniques utilized, highlighting their integration into a cohesive system for personalized health management [5]. This work underscores the potential of AI-driven approaches to enhance health outcomes and user engagement in wellness planning [3]. The problem statement is young adults (15–30) face increasing health challenges like obesity, diabetes, and heart disease due to poor diets and sedentary lifestyles. Current health management systems lack personalization, offering generic solutions that fail to meet individual needs [6]. Additionally, these systems often lack integration of advanced machine learning, limiting their ability to deliver precise, adaptive recommendations for meal planning, exercise, and progress tracking.

The objectives include a Disease Identification Model that predicts health conditions from user symptoms and provides actionable advice using robust datasets and optimized algorithms, alongside Personalized Health Recommendations offering tailored lifestyle adjustments based on metrics like BMI, blood pressure, and activity levels. A Meal Planning Model suggests daily calorie targets and recipes aligned with user preferences, restrictions, and health goals, while a Workout Routine Model generates adaptable exercise plans considering fitness levels, medical conditions, and equipment availability. Additionally, a Muscle Targeting Model identifies muscle groups for safety, goal-specific workouts, and a BMI Prediction Model estimates timelines for achieving target BMI using factors such as calorie balance and hydration.

II. LITERATURE REVIEW

Personalized health management has emerged as a critical area in addressing the growing prevalence of lifestyle-related diseases such as obesity, diabetes, and heart disease among young adults. Research underscores the benefits of tailored interventions, highlighting their role in improving adherence to health regimens and achieving better health outcomes.

Personalized Meal Planning- Studies reveal that personalized meal plans based on user-specific factors like dietary preferences, health conditions, and caloric needs enhance compliance and long-term success. Algorithms that incorporate user feedback and integrate external recipe databases enable the creation of adaptive meal plans that align with nutritional goals while accommodating individual restrictions and preferences [1] [2].

Customized Workout Recommendations- The effectiveness of personalized exercise plans has been widely documented, with research showing that individualized routines improve user engagement and reduce injury risks [3]. Machine learning models, such as Gradient Boosting and XGBoost, have been successfully applied to generate workout plans tailored to fitness levels, health constraints, and goals. These models consider factors like user demographics, activity levels, and medical history to deliver safe and effective recommendations [4].

Progress Tracking and Health Forecasting- Progress tracking is essential for maintaining user motivation and ensuring adherence to health goals. Predictive models, such as those utilizing XGBoost and other ensemble techniques, have demonstrated accuracy in estimating timelines for achieving targets like BMI. Incorporating features such as caloric balance, hydration, and sleep patterns further enhances prediction reliability [4]. Health forecasting models also aid in identifying potential risks and providing preventive recommendations, empowering users to make informed decisions. Advancements in Machine Learning for Health Management- Recent advancements in machine learning have enabled the development of systems that provide precise and adaptive health solutions. Ensemble models, such as those combining Random Forest, LightGBM, and XGBoost, have proven effective in improving the accuracy and robustness of health predictions. Techniques like stacking, hyperparameter optimization, and cross-validation ensure that these models achieve high reliability in diverse scenarios.The integration of AI-driven models into scalable platforms has transformed health management. These systems combine user data with advanced algorithms to deliver comprehensive solutions for meal planning, exercise, and health monitoring. The adoption of APIs for real-time feedback and personalized recommendations has further enhanced user engagement and system accessibility. The literature highlights the transformative potential of personalized health management systems powered by machine learning. By leveraging user data and integrating advanced algorithms, these systems address the limitations of generic health recommendations, offering precise and adaptive solutions. This research builds upon existing studies by focusing on the integration of multiple personalized components into a unified platform, targeting young adults' unique health needs.

III. METHODOLOGY

This research developed four models to provide personalized health solutions, focusing on meal planning, workout recommendations, progress tracking, and health forecasting. Each model leverages user data, including demographics, health metrics, and activity levels, to deliver tailored insights.

The **Disease Identification Model** analyzes user symptoms to predict potential health conditions utilizing decision tree and Random Forest. Preprocessing involved handling missing values and mapping symptoms to diseases. Classifiers were trained on a labeled dataset, with hyperparameter tuning and cross-validation ensuring robustness. The model outputs a predicted disease, description, and actionable home care advice, making it user-friendly [1]. The **Personalized Recommendation Model** utilizes detailed health metrics such as BMI, blood pressure, and activity levels to generate individualized recommendations utilizing Random Forest. Preprocessed inputs were scaled and standardized, and regression techniques provided tailored advice for lifestyle adjustments and preventive measures.

The **Personalized Meal Planning Model** combines user factors like BMI, age, and activity levels to predict daily calorie needs utilizing [1] LightGBM, XGBoost and Random Forest. Ensemble algorithms with stacking regressors and hyperparameter optimization with GridSearchCV [2], the model suggests personalized meal plans by integrating with an external recipe API.

The **Workout Recommendation Model** predicts exercise plans tailored to user fitness goals and health constraints using XGBoost and Gradiant Boost Clssifier. Input data was standardized, and classifiers were optimized using hyperparameter tuning with GridSearchCV and crossvalidation to ensure performance and accuracy [2].

Finally, the **BMI Prediction Model** estimates the time required to achieve target BMI. Features such as calorie intake, hydration efficiency, and sleep data were engineered and standardized. The XGBoost regressor, optimized with RandomizedSearchCV, provides accurate predictions, supporting users in achieving their health goals [3]

FUNCTION

PersonalizedHealthAssistant(user_data, weekly_updates):
INPUT: User data (age, weight, height, gender, activity level, health metrics, fitness goals), Weekly updates (calories, weight changes)

OUTPUT: Personalized meal plan, workout plan, and progress report

1: Preprocess Data

HANDLE missing values using mean/mode
SCALE numerical data (e.g., BMI, calories burned) LABEL-ENCODE categorical data (e.g., gender, preferences)
STORE processed data

2: Generate Meal Plan

LOAD pre-trained ensemble model
CALCULATE recommended calorie intake
RETRIEVE recipes from Spoonacular API using calories and preferences
STORE meal plan (breakfast, lunch, dinner)

3: Generate Workout Plan

LOAD pre-trained fitness model PREDICT workouts and muscle groups
OPTIMIZE the workout schedule dynamically by analyzing user feedback, performance metrics, and progress over time to ensure effectiveness and adaptability

STORE the personalized weekly workout plan securely, allowing users to access and update it as needed for continuous improvement and goal alignment. Regular updates provide variety to maintain user motivation, while detailed analytics ensure plans remain tailored to individual needs.

4: Track Progress

UPDATE user metrics (BMI, hydration, sleep)
PREDICT goal achievement timeline with XGBoost
GENERATE progress report with insights and recommendations.

END FUNCTION

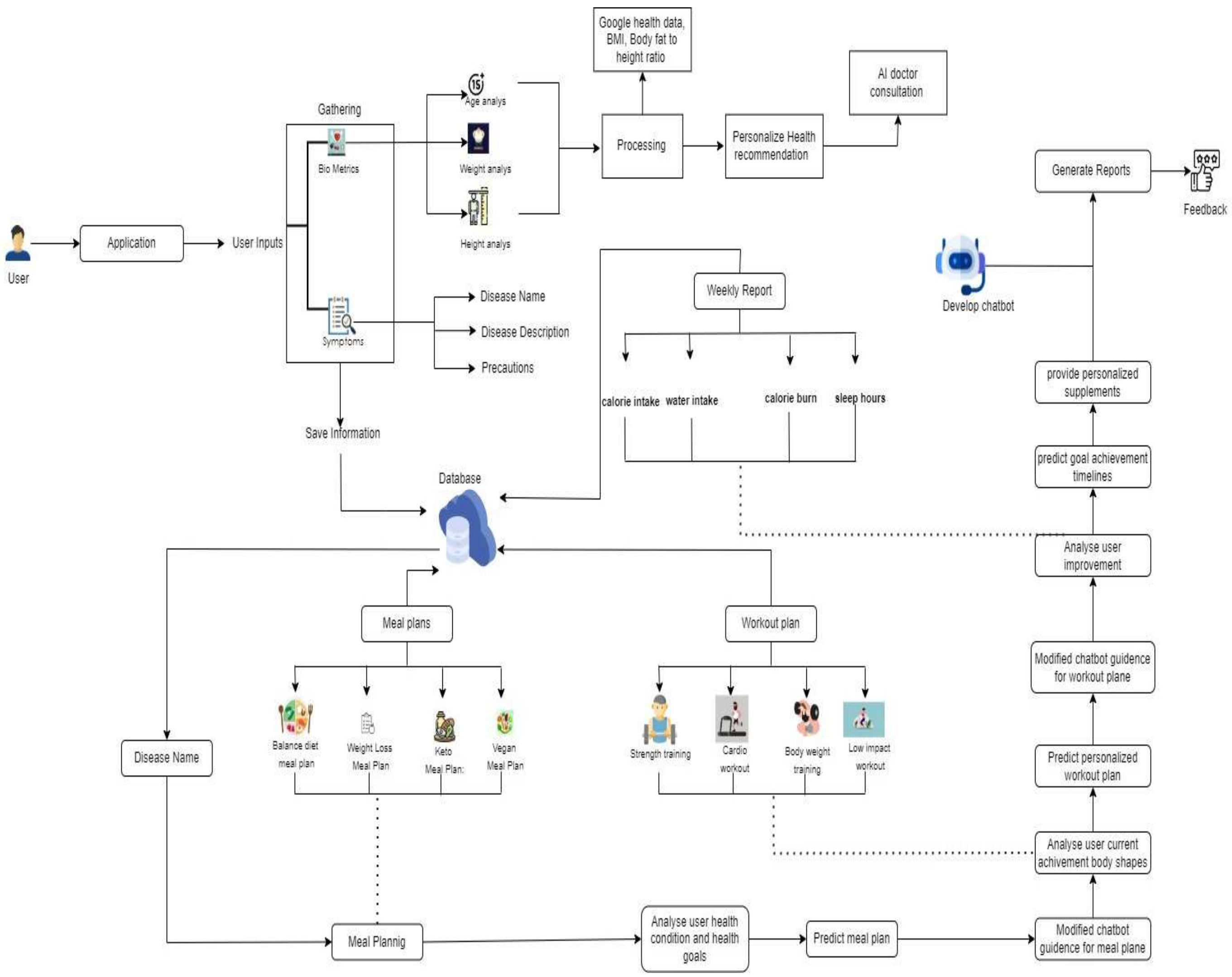

Fig. 1. System overview diagram

EQUA TIONS

The training process involves optimizing the hyperparameters using GridSearchCV, which tunes the model to minimize the mean absolute error (MAE). Hyperparameter Tuning:

- **Grid Search** optimizes the parameters such as the number of trees in RandomForest (`n_estimators`) and the depth of the trees (`max_depth`).
- The model is trained to minimize,

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} \mid y_i - \text{y\^{}i} \mid$$

Where y_i is the true value and yi is the predicted value for the i-th sample [9].

The **Personalized Recommendation Phase** includes the calculation of critical health metrics, such as Body Mass Index (BMI) and Body Fat-to-Weight Ratio (BFWR), which are pivotal for tailoring user-specific health recommendations.

Body Mass Index (BMI):

BMI is calculated using the following equation:

BMI=Weight (kg) / Height (m)2

BMI is employed to classify users into health categories, such as underweight, normal weight, overweight, or obese, aiding in accurate health profiling.

Where *Weight (kg)* represents the weight of the individual in kilograms, and *Height (m)2* denotes the square of the individual's height in meters.

Body Fat-to-Weight Ratio (BFWR):

The BFWR percentage is computed using:
BFWR (%)=1.2×BMI+0.23×Age (years)−10.8×Gender−5.4
Where *Body Fat (kg)* represents the total body fat in kilograms, and *Weight (kg)* is the individual's total weight in kilograms. This ratio is used to assess the proportion of body fat relative to total body weight, providing insights into an individual’s overall health and fitness status.

IV. RESULTS AND DISCUSSION

The error distribution as depicted in Fig. 2 illustrates the differences between actual and predicted calorie intake values. The histogram, combined with a kernel density estimation (KDE) curve, reveals a near-normal distribution centered around zero.

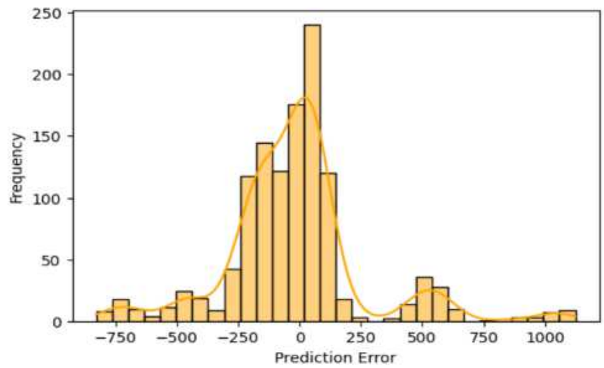

Fig. 2. Error distribution

The scatter plot in Fig. 3 shows the alignment between actual and predicted calorie intake values. The red dashed line represents perfect predictions (Most data points cluster near this line, indicating the model's effectiveness.

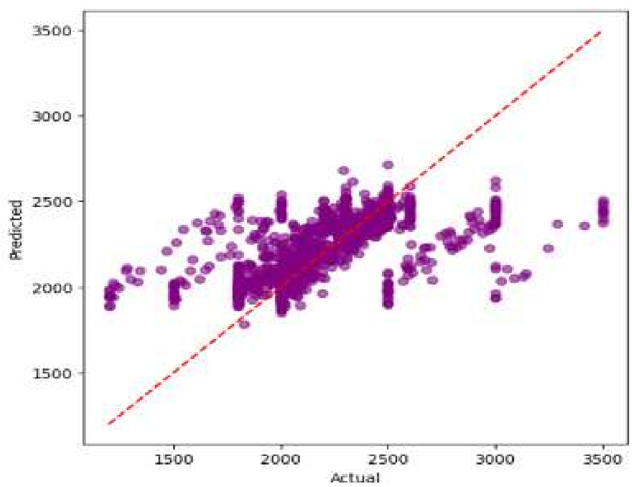

Fig. 3. Actual and predicted calorie intake values.

The chart presented in Fig. 4 visualizes the dataset after removing null values. The x-axis represents the column names, including *Disease* and the 17 symptom features, while the y-axis reflects the frequency of null values across the dataset. The flat line at **0.00** confirms that all null values have been successfully eliminated from the dataset. This preprocessing step ensures data integrity and improves the performance of the machine learning model by providing clean, complete data for training and testing.

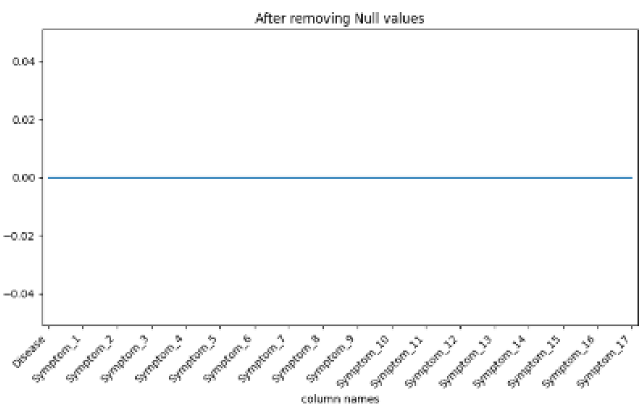

Fig. 4. Analysis of null values in the dataset

Fig. 5 shows the feature importance of various input factors in the workout routine model. Daily calorie intake is the most significant feature, followed by disease status and BMI. Water intake has the lowest contribution to the model's predictions.

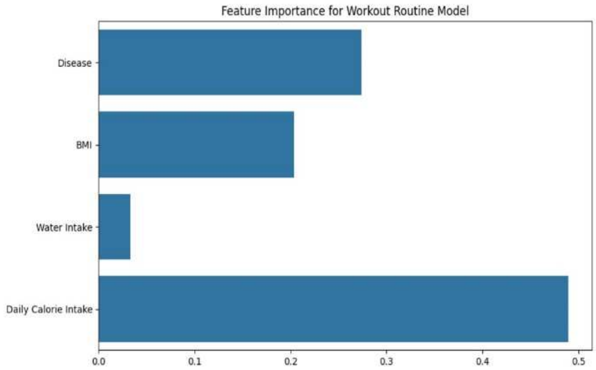

Fig. 5. Feature importance for workout routine

The heatmap presented in Fig. 6 demonstrates the accuracy of the Random Forest model in disease identification by comparing actual versus predicted values for each condition. The diagonal line of higher intensity (warmer colors) represents cases where the predicted diseases align perfectly with the actual diseases. This highlights the model's overall effectiveness in correctly classifying diseases such as **Hypertension** and **Diabetes**.

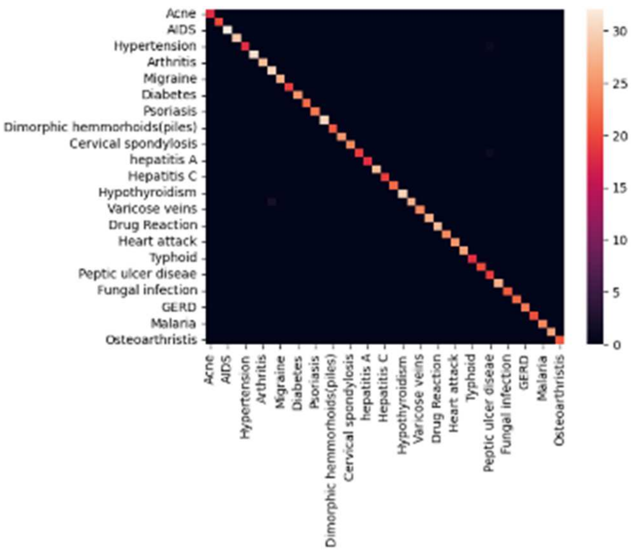

Fig. 6. Heatmap of actual vs. predicted disease classifications using random Forest Model.

Fig.7 scatter plot compares the actual time required to achieve a goal BMI (on the x-axis) with the predicted time (on the y-axis). The red dashed line represents the ideal scenario where predicted and actual values perfectly align.
The clustering of points around the line indicates that the prediction model is highly accurate, with minimal deviation between actual and predicted values.

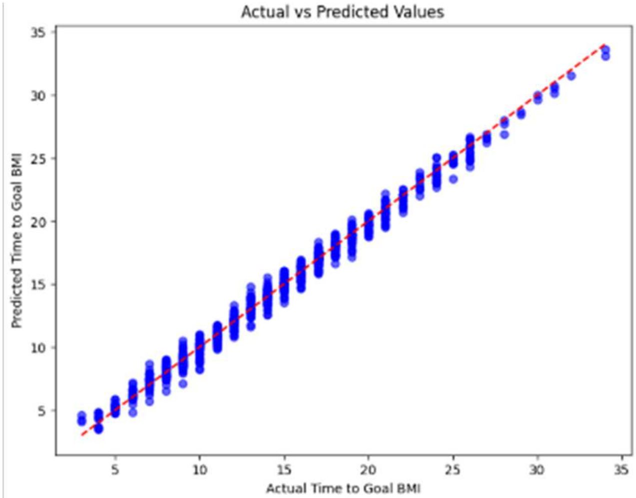

Fig. 7. Scatter plot compares the actual time required to achieve a goal

Fig. 8 graph demonstrates the relationship between the training set size and the model's performance, represented by the mean squared error (MSE). The blue curve shows the training error, which remains low, indicating good model fit during training. The orange curve represents validation error, which decreases as the training set size increases, eventually stabilizing. This suggests that the model generalizes well and benefits from a larger dataset, improving its performance.

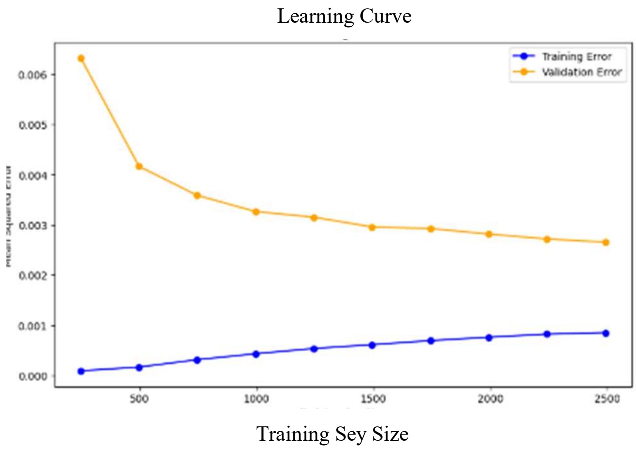

Fig. 8. Learning curves demonstrate the relationship between the training set size and the model's performance

The histogram in Fig. 9 compares the actual and predicted workout routines. The overlapping bars illustrate the distribution of predictions against actual data points. While some deviations exist, the considerable alignment between actual (blue) and predicted (red) frequencies suggests the model performs reliably in estimating workout routines for personalized fitness planning.

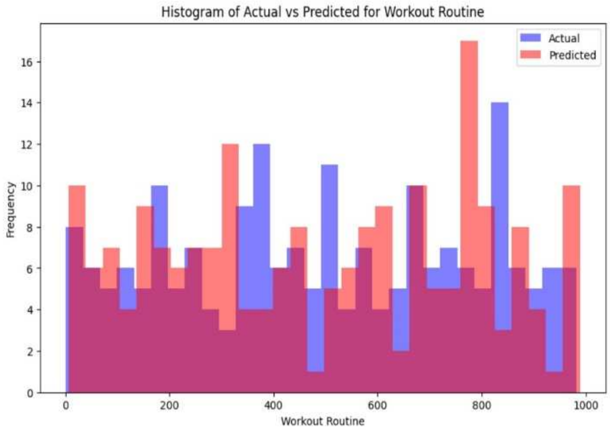

Fig. 9. Feature Importance for Workout Routine Model

V. CONCLUSION

This research demonstrates the potential of machine learning in personalized health management for young adults. By integrating six AI-driven models, we addressed key challenges in meal planning, workout customization, progress tracking, and health forecasting. Techniques like ensemble learning, hyperparameter tuning, and robust preprocessing ensure accurate, reliable predictions. The models provided tailored recommendations and timelines for achieving health goals, showcasing the effectiveness of algorithms such as Random Forest, XGBoost, and LightGBM. Future efforts will focus on enhancing accuracy, expanding datasets, and deploying a mobile application for real-time feedback, making health management more accessible and impactful.

VI. REFERENCES

[1] Y. H. K. H. Min Chen, " Disease Prediction by Machine Learning over Big Data from Healthcare Communities," *Disease Prediction by Machine Learning over Big Data from Healthcare Communities,* p. 10, 2017.

[2] T. S. J. P. S. T. N. P. F. M. J. J. E. S. B. S. J. V. Mika Kivimäki, " Body-mass index and risk of obesity-related complex multimorbidity: an observational multicohort study," *Body-mass index and risk of obesity-related complex multimorbidity: an observational multicohort study,* p. 11, 2022.

[3] D. Gondocs, "AI in medical diagnosis: AI prediction & human judgment," *AI in medical diagnosis: AI prediction & human judgment ,* p. 12, 2024.

[4] A. B. Taieb, "Understanding the risk of developing weight-related complications associated with different body mass index categories: a systematic review," *Understanding the risk of developing weight-related complications associated with different body mass index categories: a systematic review,* p. 21, 2022.

[5] A. Chen, "Simulation of a machine learning enabled learning health system for risk prediction using synthetic patient data," *Simulation of a machine learning enabled learning health system for risk prediction using synthetic patient data,* p. 11, 2022.

[6] H. Habehh, "Machine Learning in Healthcare," *Machine Learning in Healthcare,* p. 10, 2021.

[7] RACIEL YERA TOLEDO1, "A Food Recommender System Considering Nutritional Information and User Preferences," *A Food Recommender System Considering Nutritional Information and User Preferences,* 2021.

[8] M. Ankita Yadav, "Workout Prediction using Machine Learning," *Workout Prediction using Machine Learning,* 2023.

[9] A. R. Joyita, "Human-Behavior-Based Personalized Meal Recommendation and Menu Planning Social System," *Human-Behavior-Based Personalized Meal Recommendation and Menu Planning Social System,* 2019.

[10] M. Anusari, "SriHealth: A Single Platform for Meal Plans,Workouts, Yoga Schedules Based on SriLankan Lifestyle," *SriHealth: A Single Platform for Meal Plans,Workouts, Yoga Schedules Based on SriLankan Lifestyle,* p. 6, 2024.

[11] A. Yadav, " Workout Prediction using Machine," *Workout Prediction using Machine ,* 2023.

[12] A. Jadhav, "Workout Whiz – Your Personalized AI PAL," *Workout Whiz – Your Personalized AI PAL,* p. 6, 2023.

[13] M. S. A. Tehrani, " Personalized Meal Planning for Diabetic Patients Using a Multi-Criteria Decision Making Approach," *Personalized Meal Planning for Diabetic Patients Using a Multi-Criteria Decision Making Approach,* p. 6, 2019.