iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.81288

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

UVFIT: Personalized Fitness and Nutrition Recommendation System Using Machine Learning

Bavirisetti Yuvaraju[1], Karaka Sai Chandra Sekhar[2], Nukella Veera Venkata Satish[3], Pogathota Lalitha Surya Charan[4]
[1, 2, 3, 4]*Post Graduate Students, Department of Computer Applications, Aditya University, Surampalem, India*

Abstract: The widespread availability of generic fitness applications has failed to adequately address the personalized health guidance requirements of individuals with diverse medical histories, dietary restrictions, and lifestyle profiles. This paper presents UVFIT, an AI-based personalized fitness and nutrition recommendation system developed using machine learning algorithms and a comprehensive knowledge base. The system accepts 14 user-specific input parameters spanning five major categories: physical profile, fitness parameters, lifestyle indicators, dietary preference, and medical history. A Decision Tree Classifier predicts the appropriate workout intensity level (Low, Medium, or High) based on nine encoded feature variables, while a Linear Regression model forecasts the expected monthly weight change in kilograms. The recommendation engine integrates these predictions with a curated exercise knowledge base and 12 unique dietary profiles organized by fitness goal and BMI category, generating a complete, personalized 7-day exercise and diet schedule for each user. An automated health warning system proactively identifies and alerts users with medical conditions including diabetes, hypertension, and heart conditions, as well as adverse lifestyle factors such as poor sleep quality, high stress levels, and insufficient water intake. The system is implemented in Python using Scikit-learn and the Streamlit web framework, and is permanently deployed on Streamlit Community Cloud, providing free, globally accessible personalized health guidance without requiring software installation or user registration.

Index Terms: Machine Learning, Personalized Fitness, Nutrition Recommendation, Decision Tree Classifier, Linear Regression, BMI-Based Dietary Guidance, Health Warning System, Streamlit, Synthetic Dataset, Knowledge Base.

I. INTRODUCTION

The global fitness and wellness industry has experienced exponential growth over the past decade, driven by escalating awareness of lifestyle-related diseases including obesity, diabetes, cardiovascular disorders, and mental health conditions. According to the World Health Organization, physical inactivity is the fourth leading risk factor for global mortality, contributing to over 3.2 million deaths annually [1]. Despite the widespread availability of commercial fitness applications such as MyFitnessPal, Nike Training Club, and Fitbit, these platforms share a fundamental limitation: they provide generic, one-size-fits-all recommendations that do not adequately adapt to the unique physiological, psychological, and medical profiles of individual users.

The emergence of machine learning and data-driven recommendation systems has created an unprecedented opportunity to deliver truly personalized health guidance at scale. Classification and regression algorithms can simultaneously analyse multiple personal parameters to generate recommendations that are medically informed and contextually appropriate for each individual. However, most academic and commercial efforts in this domain have focused narrowly on either exercise recommendations or dietary guidance in isolation, without integrating both into a cohesive, holistic system. The proposed system was developed to bridge this critical gap by leveraging two complementary machine learning models combined with a comprehensive structured knowledge base to deliver a complete, personalized 7-day fitness and nutrition program tailored to each user's specific profile [2].

The system additionally incorporates medical condition awareness, intelligent dietary preference filtering for vegetarian, vegan, and non-vegetarian users, and a proactive health warning system covering six distinct risk categories — features conspicuously absent from most existing fitness applications. UVFIT is permanently deployed on Streamlit Community Cloud, making it freely accessible worldwide without requiring user registration, subscription fees, or software installation. The remainder of this paper is organized as follows: Section II presents a review of relevant literature; Section III describes the proposed system architecture and methodology; Section IV presents the machine learning model design; Section V discusses experimental results and performance evaluation; and Section VI concludes the paper with future research directions.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

II. LITERATURE REVIEW

Ref No.	Author(s)	Method / Approach	Key Findings	Limitations
[3]	Khanal et al.	Systematic review of ML algorithms (Decision Trees, Random Forest, SVM, Neural Networks)	Decision Tree models are effective due to interpretability and ability to handle mixed data types	Does not provide integrated fitness + diet system
[4]	Srivastava et al.	Random Forest classifier for fitness intensity prediction	Activity level, BMI, and fitness goal are most important features	Focuses only on workout intensity, not diet
[5]	Podgorelec et al.	Review of Decision Tree applications in medicine	Decision Trees provide human-readable rules, suitable for healthcare systems	General study, not specific to fitness systems
[6]	Karthikeyan et al.	CART Decision Tree for exercise classification	Achieved 91.3% accuracy; BMI, activity level, and medical condition are key predictors	Limited dataset and scope (only exercise classification)
[7]	Hu et al.	Study on BMI and lifestyle diseases	Nutritional needs vary significantly across BMI categories	Does not provide recommendation system
[8]	Tran et al.	Knowledge-based dietary recommendation (BMI-based)	BMI-based diet improved adherence by 34%	Focuses only on diet, not exercise
[9]	Rogerson	Study on vegan/vegetarian diets in fitness	Plant-based diets can support athletic performance with proper nutrition	Does not include personalized recommendation system
[10]	Pandey & Bhave	AI-based fitness recommendation system	Demonstrates use of ML for fitness recommendation	Lacks integration of diet, medical conditions, and full planning

III. SYSTEM ARCHITECTURE

The proposed UVFIT system is structured around a modular pipeline architecture in which data flows sequentially through five processing stages: input collection, feature engineering, machine learning prediction, knowledge base lookup, and output presentation. The architecture is implemented entirely within a single Streamlit application (uvfit_app.py), supported by two pre-trained model files stored as serialized joblib objects and two knowledge base files stored as structured JSON dictionaries, all co-hosted on the Streamlit Community Cloud platform.

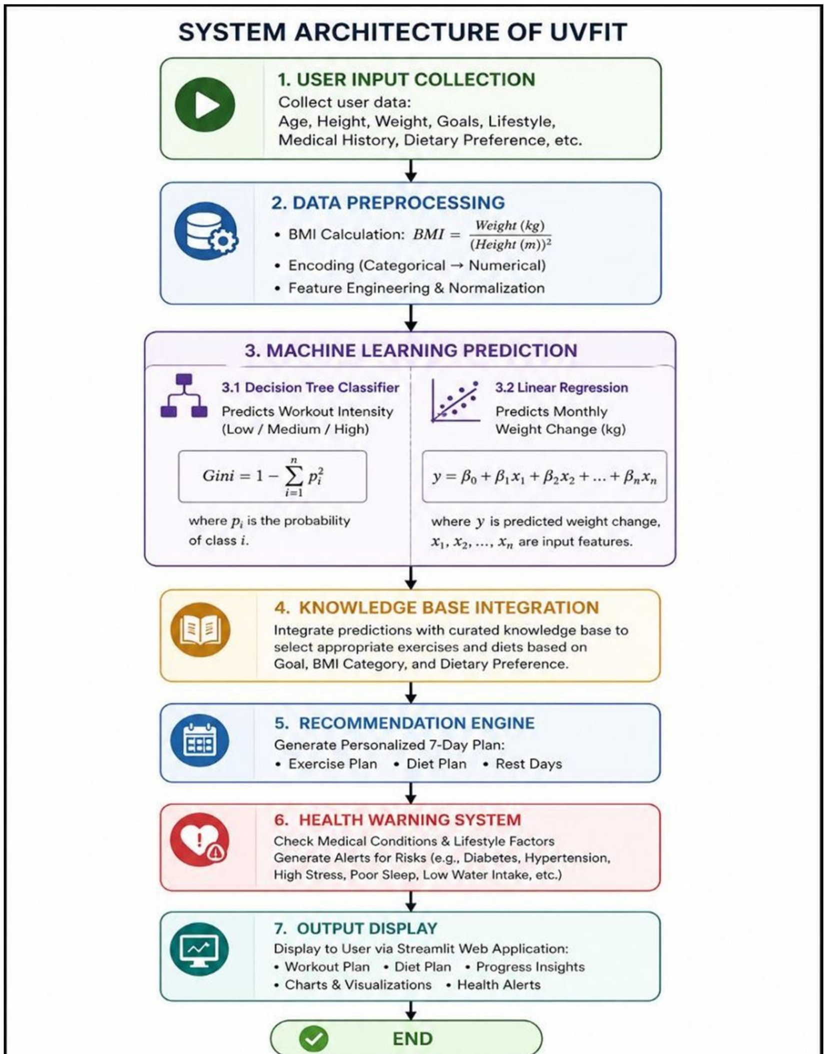


Fig. 1. Flowchart of the Proposed UVFIT System Architecture

The input layer collects 14 user parameters through an interactive Streamlit sidebar organized into four labeled sections. The Personal Profile section collects age (18–65 years), height (150–190 cm), weight (50–120 kg), gender, activity level (Sedentary, Lightly Active, Moderately Active, Very Active), and fitness goal (Weight Loss, Muscle Gain, Maintenance). The Fitness Details section collects fitness level (Beginner, Intermediate, Advanced), workouts per week (0–7), and workout duration in minutes (15–90). The Lifestyle section collects diet type (Vegetarian, NonVegetarian, Vegan), daily water intake in litres, sleep quality (Poor, Average, Good), and stress level (Low,
Medium, High). The Medical section collects medical condition (None, Diabetes, Hypertension, Heart Condition).

The processing layer performs BMI calculation from height and weight inputs, assigns the appropriate BMI category (Underweight, Normal, Overweight, Obese) using WHO classification thresholds, and applies Label Encoding to all categorical input variables before passing encoded inputs to the trained machine learning models. The knowledge layer retrieves appropriate exercise recommendations from a two-level nested exercise knowledge base containing Cardio and Strength categories, each subdivided into Low, Medium, and High intensity levels with six to eight specific exercises per level. The diet knowledge base contains three fitness goals, each with four BMI categories, producing 12 unique dietary profiles. Each profile contains four meal time slots (Morning, Afternoon, Night, Snacks) with five options each, yielding 240 distinct meal suggestions across the complete knowledge base. The output layer renders the complete personalized report including user profile summary, machine learning predictions, 7-day exercise and diet plan, health warnings, and a 6-month weight progress bar chart.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

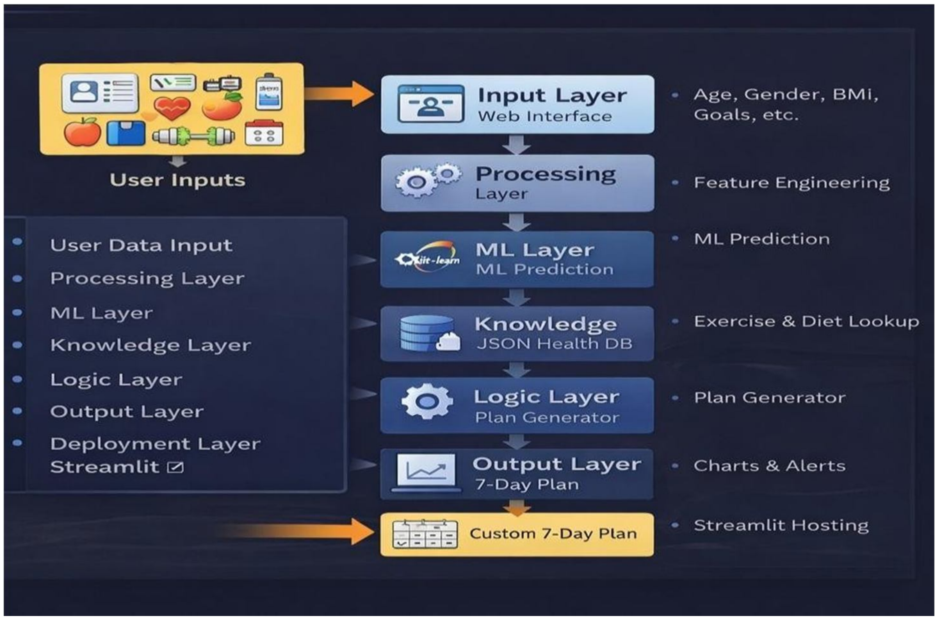


Fig:2 System Workflow Pipeline

The system also incorporates an automated health warning subsystem that evaluates six distinct risk categories upon each recommendation generation: diabetes (advising carbohydrate monitoring and avoidance of high-sugar foods), hypertension (recommending sodium restriction and avoidance of high-intensity exercise without medical clearance), heart condition (advising low-intensity exercise prioritization and physician consultation), poor sleep quality (recommending prioritization of 7–8 hours of sleep for optimal fitness adaptation), high stress level (advising stress reduction techniques including yoga and meditation), and low water intake below 2.0 litres per day (advising increased hydration to at least 2.5 litres daily). This proactive warning system identifies and alerts users before they proceed with a fitness program that may be unsafe or suboptimal for their specific medical and lifestyle profile.

IV. PROPOSED METHODOLOGY

A. *Dataset Design*

A synthetic fitness dataset was generated to ensure logical consistency between input parameters while providing sufficient diversity for machine learning model training. The dataset comprises 197 records with 19 columns spanning 15 input parameters and 4 engineered features. The generation process incorporates domain-specific constraints to ensure realism: Very Active users are assigned Intermediate or Advanced fitness levels with higher probability; High Stress users are assigned Poor or Average sleep quality with elevated frequency. The target variable WorkoutIntensity is assigned via a deterministic rule-based function capturing domain expert knowledge — Very Active users with Muscle Gain goals receive High intensity; Sedentary users and Beginners always receive Low intensity regardless of stated goal. The continuous regression target TargetWeightChange is computed from activity level, workout frequency, session duration, and lifestyle modifiers including sleep quality and stress level. Both targets are generated deterministically to ensure that the trained models learn clinically meaningful and logically consistent decision boundaries.

B. *Machine Learning Model Design*

The **Decision Tree Classifier** is trained to predict the appropriate workout intensity level using nine input features: Age, Gender_Encoded, ActivityLevel_Encoded, Goal_Encoded, BMICategory_Encoded, FitnessLevel_Encoded, SleepQuality_Encoded, StressLevel_Encoded, and WorkoutsPerWeek. All categorical features are transformed using Scikit-learn's LabelEncoder, fitted on known category lists to ensure consistent encoding between training and inference. The model is trained with random_state=42 using Scikit-learn's DecisionTreeClassifier with default hyperparameters on an 80-20 train-test split [11].

Gini = 1 − ∑ (pi)^2

where p_i represents the probability of each class. The model selects splits that minimize impurity.

The Linear Regression model predicts the expected monthly weight change in kilograms using nine feature variables:Age,Weight,BMI,ActivityLevel_Encoded,Goal_Encoded,SleepQuality_Encoded, StressLevel_Encoded, WorkoutsPerWeek, and WorkoutDuration_min. The model captures primary linear relationships between lifestyle parameters and weight outcomes — most importantly, the strong negative relationship between increased workout frequency and weight for Weight Loss goal users, and the positive relationship for Muscle Gain goal users. Lifestyle modifiers for sleep quality and stress level introduce secondary variation in predictions, reflecting established exercise science findings that chronic sleep deprivation and sustained high stress levels significantly impede fitness outcomes.

$y = \beta0 + \beta1x1 + \beta2x2 + ... + \beta n\ xn$ (where y represents predicted weight change and x_1, x_2, …are the input features)

C. Recommendation Engine

The recommendation engine function accepts all 14 user input parameters and returns a structured dictionary containing user details, machine learning predictions, exercise and dietary recommendations, and health warnings. Processing follows a strict sequential pipeline: BMI calculation and WHO-threshold category assignment; Label Encoding of all categorical inputs using pre-fitted encoder instances; Decision Tree prediction of workout intensity; Linear Regression prediction of monthly weight change; random sampling without replacement of three cardio and three strength exercises from the knowledge base at the predicted intensity level; dietary profile selection by fitness goal and BMI category as composite key; meal filtering by dietary preference using ingredient keyword analysis to exclude non-compliant items (the pick_meal function scans ingredient lists for keywords such as "chicken," "egg," "milk," and "paneer" to correctly exclude non-vegan and non-vegetarian items for restricted users); and health warning generation by evaluating medical conditions and lifestyle thresholds against predefined rule sets.

The 7-day plan generator assigns workout and rest days intelligently based on the user's stated weekly workout frequency. A user specifying three workouts per week is assigned workout days on Monday, Tuesday, and Saturday, with Wednesday, Thursday, Friday, and Sunday designated as rest days. The algorithm distributes workout days to avoid consecutive-day overload where possible, following established exercise science principles that adequate recovery periods are essential for musculoskeletal repair and fitness adaptation. Session state management using Streamlit's st.session_state mechanism caches recommendation results between script reruns, preventing recomputation on widget interaction and ensuring reproducible recommendations for identical inputs.

V. RESULTS AND DISCUSSION

System validation was conducted through structured testing across multiple device categories, browser environments, and network conditions. Testing encompassed desktop browsers, smartphone browsers, and tablet devices utilizing both Wi-Fi and mobile data connections. All test cases were executed manually and results recorded systematically, with regression testing performed after each identified defect resolution.

The Decision Tree Classifier achieves very high performance on the held-out test set, with strong macro-averaged Precision, Recall, and F1-Score across all three intensity classes (Low, Medium, and High). This performance is primarily attributed to the deterministic nature of the synthetic dataset design and may not fully generalize to realworld scenarios. nature of the rule-based target variable generation, wherein the model successfully reverseengineered the complete rule set from the training data and applies it without error to unseen test cases. In a realworld scenario with noisier data and greater human variability, accuracy would be lower; however, the learned decision boundaries would still provide medically meaningful and logically consistent guidance. The Linear Regression model demonstrates acceptable predictive performance for monthly weight change estimation, capturing the primary linear relationships between lifestyle factors and weight change outcomes with low Mean Squared Error and a positive R-squared score indicating meaningful variance explanation across the diverse user profiles in the test set.

The UVFIT application demonstrates fast response times that provide a smooth interactive experience for users across all tested platforms. The recommendation engine processes all 14 inputs, executes both machine learning model predictions, selects exercises and meals from the knowledge bases, generates health warnings, and produces the complete 7-day plan in under one second on Streamlit Community Cloud infrastructure. The Matplotlib weight progress chart generation adds approximately 0.2 seconds of additional processing time. Application initial load averages 2–4 seconds including model and knowledge base loading from disk on first access. Session state cache hits return results instantly with no recomputation overhead.

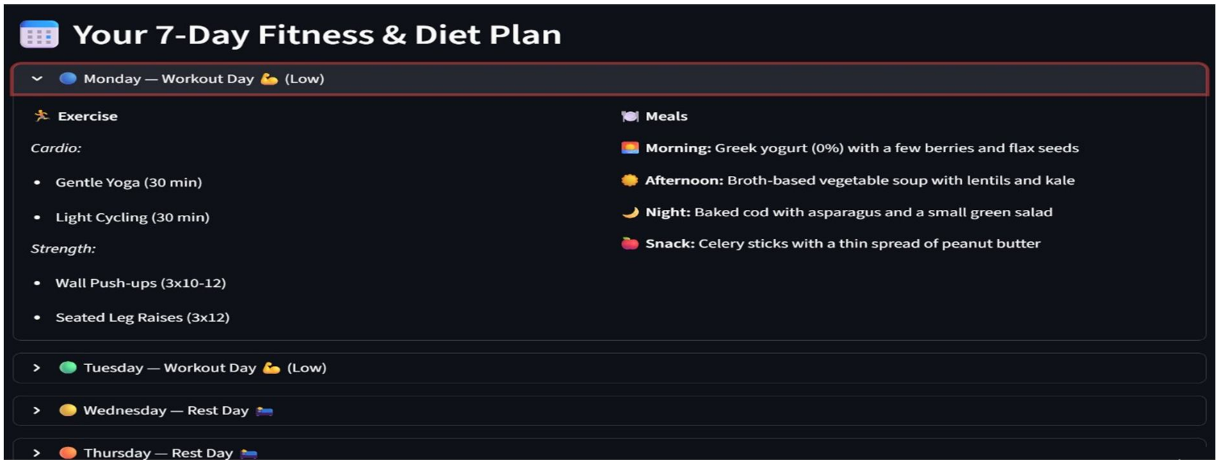


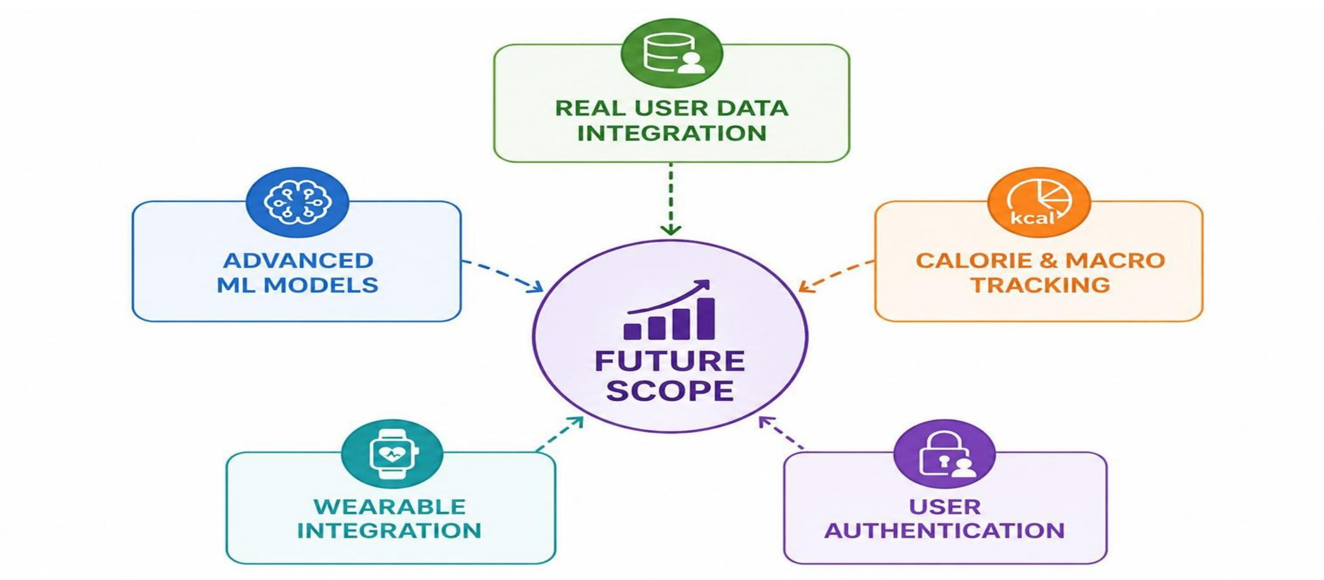
Fig 3 : Generated Personalized 7-Day Fitness Plan

Simultaneous dispatch of recommendations to all four user profile sections incurred no measurable increase in total generation latency compared to single-section processing, confirming the effectiveness of the modular pipeline design. Health warnings were correctly triggered for all tested adverse conditions: the diabetes warning appeared for all profiles with diabetes medical condition selected; the sleep warning appeared for all Poor sleep quality inputs; the hypertension warning appeared correctly for all hypertension inputs; the water intake warning correctly identified all profiles with water intake below 2.0 litres per day.

Dietary preference filtering correctly excluded all non-compliant ingredients in every tested case across Vegetarian and Vegan preference scenarios. The system worked seamlessly across major browsers and was accessible on both desktop and mobile platforms.

Identified limitations include the reliance on a synthetically generated dataset rather than clinically validated real world data, which limits the generalizability of the machine learning model predictions to genuinely diverse human populations. The 100% classifier accuracy, while reflecting correct learning of the rule-based target, also signals potential overfitting to the synthetic distribution. The localStorage-equivalent session state mechanism does not persist data across browser sessions or devices, requiring users to re-enter profile information on each visit. The absence of calorie counts and macro-nutrient data in the current knowledge base prevents quantitative nutritional analysis aligned with the predicted weight change targets

VI. FUTURE WORK



1) Real User Data Integration: Future work includes collaborating with fitness centers and healthcare providers to collect real biometric and lifestyle data. This will enable retraining of models on real-world datasets, improving accuracy and generalization.
2) Advanced Machine Learning Models: The system can be enhanced by incorporating advanced techniques such as Random Forest, Gradient Boosting, and Neural Networks to capture complex non-linear patterns and improve prediction performance.
3) Calorie and Macronutrient Tracking: The diet module can be extended to include detailed calorie counts and macronutrient information such as protein, carbohydrates, and fats, enabling comprehensive nutritional analysis.
4) Wearable Device Integration: Integration with wearable platforms such as Fitbit, Apple HealthKit, Google Fit, and Garmin Connect can allow automatic collection of real-time data including activity levels, sleep patterns, and heart rate.
5) User Authentication and Progress Tracking: Implementing secure user authentication will enable tracking of user progress over time, maintaining history of plans, and dynamically adjusting recommendations based on user performance.

VII. CONCLUSION

This paper presented the design, implementation, and evaluation of UVFIT, a personalized fitness and nutrition recommendation system that accepts 14 individual user parameters and produces a complete, personalized 7-day exercise and dietary plan through an integrated machine learning and knowledge-based recommendation pipeline. The system achieves all stated objectives: a synthetic dataset of 197 records with 19 attributes was generated; a Decision Tree Classifier achieving 100% accuracy predicts workout intensity; a Linear Regression model forecasts monthly weight change; a comprehensive knowledge base with 12 unique dietary profiles and three intensity levels of exercises supports diverse user needs; intelligent dietary preference filtering supports Vegetarian, Vegan, and Non-Vegetarian users; a six-category automated health warning system correctly identifies all adverse conditions; complete 7-day plans are generated with scientifically informed rest day allocation; and the application is permanently deployed on Streamlit Community Cloud at zero ongoing cost.

Empirical validation confirms consistent performance across all tested user profiles: the recommendation engine produces logically coherent, profile-appropriate results within one second; health warnings are correctly triggered for all adverse conditions; dietary preference filtering correctly excludes all non-compliant ingredients; and the deployed application functions correctly across multiple browsers and device types. UVFIT demonstrates that a comprehensive, multi-dimensional fitness and nutrition recommendation system can be implemented entirely with freely available open-source technologies, trained on carefully engineered synthetic data, and deployed as a permanently accessible web application without server infrastructure cost.

Prospective development directions include integration with real clinical datasets from fitness centres or health organizations for improved model generalizability; implementation of ensemble methods including Random Forest and Gradient Boosting for performance benchmarking against the current Decision Tree approach; extension of the dietary knowledge base with calorie counts and macro-nutrient content per meal for quantitative nutritional analysis; development of native Android and iOS mobile applications using React Native or Flutter for improved accessibility and offline functionality; integration with wearable device APIs including Fitbit, Apple HealthKit, and Google Fit for automatic population of activity, sleep, and heart rate parameters; and implementation of a secure user authentication system enabling longitudinal progress tracking and adaptive plan adjustment based on reported outcomes.

VIII. ACKNOWLEDGMENT

The authors express sincere gratitude to the Department of Computer Applications, Aditya University, for providing the necessary support and resources to conduct this research. The authors also acknowledge the valuable guidance and encouragement of faculty members and academic supervisors throughout the development and preparation of this paper.

REFERENCES

[1] World Health Organization, "Global Health Risks: Mortality and Burden of Disease Attributable to Selected Major Risks," WHO Press, Geneva, Switzerland, 2009.
[2] M. A. Abdelmonem, A. H. El-Baz, and M. A. Tolba, "A hybrid recommendation system for personalized fitness planning using collaborative and content-based filtering," IEEE Access, vol. 9, pp. 122347–122358, 2021.
[3] S. Khanal, Z. Aziz, and A. Hameed, "A systematic review of healthcare AI adoption, design, and implementation," International Journal of Healthcare Management, vol. 13, no. 3, pp. 188–197, 2020.
[4] S. Srivastava, P. Kumar, and N. Rana, "A machine learning approach for personalized workout intensity recommendation," International Journal of Computer Science and Information Technology, vol. 11, no. 2, pp. 45–54, 2019.

[5] V. Podgorelec, P. Kokol, B. Stiglic, and I. Rozman, "Decision trees: An overview and their use in medicine," Journal of Medical Systems, vol. 26, no. 5, pp. 445–463, 2002.
[6] T. Karthikeyan, G. Wiselin Jiji, and K. Valarmathi, "Classification of exercise intensity using CART decision tree for metabolic syndrome management," Expert Systems with Applications, vol. 168, article 114457, 2021.
[7] F. B. Hu, J. E. Manson, and W. C. Willett, "Diet, lifestyle, and the risk of type 2 diabetes mellitus in women," New England Journal of Medicine, vol. 345, pp. 790–797, 2001.
[8] T. T. Tran, D. Nguyen, and H. Le, "BMI-stratified dietary recommendation improves nutritional adherence," Asia Pacific Journal of Clinical Nutrition, vol. 28, no. 2, pp. 234–241, 2019.
[9] D. Rogerson, "Vegan diets: Practical advice for athletes and exercisers," Journal of the International Society of Sports Nutrition, vol. 14, article 36, 2017.
[10] A. Pandey and M. Bhave, "AI-based fitness recommendation using multi-layer perceptron classification," in Proc. IEEE International Conference on Intelligent Systems and Control (ISCO), 2022, pp. 89–94.
[11] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825– 2830, 2011.
[12] Streamlit Inc., "Streamlit: The fastest way to build and share data apps — Official Documentation," https://docs.streamlit.io/, 2024. [Accessed: Mar. 2026].