© DEC 2025 | IRE Journals | Volume 9 Issue 6 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I6-1713155

Body Fat Percentage Detection Using Machine Learning

HARISH H K[1], LANCHANA A K[2], UNNATHI S R[3], USHA K S[4], VINUTHA H S[5]
[1]*Asst Professor, Department of Computer Science & Engineering, Maharaja Institute of Technology, Mysore, Karnataka, India.*
[2,3,4,5] *Student, Department of Computer Science & Engineering, Maharaja Institute of Technology, Mysore, Karnataka, India.*

Abstract- This study presents a machine leaming-based approach for predicting body fat percentage and identifying associated health risks using dietary intake patterns and physical attributes. The primary objective of the work is to develop an accessible and non-invasive method that eliminates the need for specialized medical. equipment. The system collects anthropometric measurements such as height, weight, waist, hip, and forearm length, along with daily meal-level food entries including breakfast, lunch, and dinner items. A Linear Regression model is employed to estimate the body fat percentage based on engineered features derived from both physical and dietary data. Subsequently, an XGBoost classifier analyzes the predicted body fat percentage to detect potential diseases and risk levels related to obesity, such as diabetes and cardiovascular conditions. Experimental evaluation demonstrates that the model achieves consistent prediction accuracy and effectively classifies health risks, offering a practical decision-support tool for early health assessment. The findings Indicate that integrating real-world diet patterns with physical measurements provides meaningful insights for personalized health monitoring. Overall, the system contributes to improved awareness of body composition, highlights risk factors linked to unhealthy body fat levels, and supports preventive healthcare through data-driven analysis.

Index Terms— Body Fat Percentage, Machine Learning, Linear Regression, Xgboost, Health Risk Detection

I. INTRODUCTION

Body fat percentage is an essential indicator for evaluating an individual's overall health, as it provides a more accurate representation of body composition than conventional metrics such as weight or Body Mass Index (BMI). High or abnormal body fat levels are strongly linked to various chronic illnesses, including obesity, cardiovascular diseases, and diabetes. As lifestyle-related disorders continue to rise globally, there is a growing need for reliable and accessible methods to assess body fat and associated health risks. This growing importance has motivated research toward developing non-invasive, user-friendly, and data-driven systems. Recent advancements in machine learning have enabled more precise analysis of complex health-related data, particularly concerning diet and anthropometric measurements. Current research efforts focus on the integration of physical attributes with dietary patterns to improve body composition estimation. However, many existing approaches rely on expensive diagnostic tools or require users to provide detailed nutrient-level dietary information, which may not be practical for daily monitoring. This creates a gap for systems that can analyze simpler, meal-based food inputs while still providing accurate predictions.

In response to these limitations, the present study introduces a machine-learning-based system that predicts body fat percentage using physical measurements and meal-level dietary entries. The model employs Linear Regression to estimate body fat and uses XGBoost to detect diseases associated with abnormal fat levels. By combining practical input methods with robust predictive algorithms, this work aims to offer an affordable, non invasive, and effective solution for early health assessment and personalized wellness tracking.

II. METHODOLOGY

The present research on Body Fat Percentage Detection using Machine Leaming focuses on integrating physical measurements and meal-level dietary inputs to build a predictive health assessment system. The methodology begins with the collection of anthropometric data such as height, weight, waist length, hip length, forearm length. In parallel, the system records daily food consumption by capturing the actual food items consumed during break fast,lunch and dinner.These raw dietary inputs are preprocessed by mapping each food item to a corresponding nutrient profile using a predefined food database .The extracted nutritional attributes are then aggregated to create meaningful features for machine learning analysis. After preprocessing, a Linear Regression model is trained to estimate body fat percentage from the combined physical and dietary features. The predicted body fat percentage is then used as an input to the XGBoost classifier, which identifies possible diseases or health risks associated with abnormal fat levels. This multi-stage approach ensures a systematic and accurate prediction pipeline.

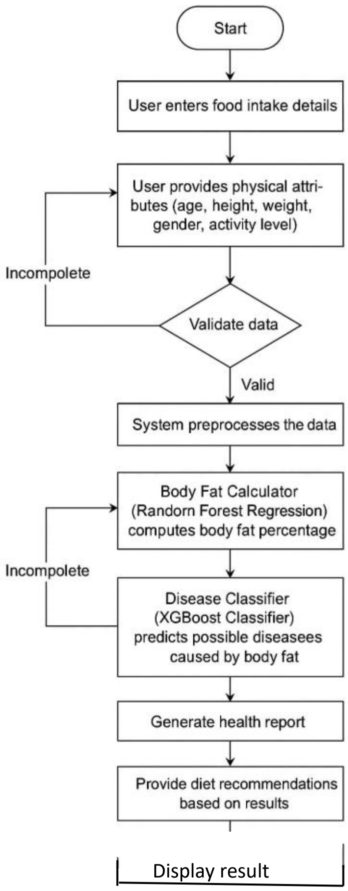

Figure 1: Flow Diagram

III. MODELLING AND ANALYSIS

Materials Used: The materials used in this study consist of anthropometric measurements and dietary inputs collected from individuals. The physical attributes include height, weight, waist length, hip length, and forearm length, which serve as primary indicators for estimating body composition. Dietary information is obtained by recording the specific food items consumed during breakfast, lunch, and dinner. These food items are standardized and mapped to nutritional categories to extract meaningful features. A curated food reference list is used to assign estimated nutrient values to each food entry. The computational materials include Python-based machine learning libraries such as NumPy, Pandas, Scikit-Learn, and XGBoost. The dataset is preprocessed, encoded, and normalized using commonly used data transformation tools to ensure consistency and accuracy. All experiments were conducted using a standard computing environment with sufficient processing capacity to train and evaluate the machine learning models.

Table 1. Materials and Model

Category	Description
Physical Attributes	Height,Weight,Waist,Length,Hip Length,Forearm Length
Dietary Inputs	Meal-wise food items (3 times meals per day)
Preprocessing Tools	NumPy,Pandas,Label Encoding,Feature Engineering
Prediction Model	Linear Regression (Body Fat Percentage)
Classification Model	XGBoost Classifier (Disease Detection)
Development Environment	Python,JupyterNotebook/VS Code

Model Used :Two machine learning models are used in this research work. The first is the Linear Regression Model, which predicts body fat percentage based on extracted features from physical attributes and dietary information. This model is chosen due to its simplicity, interpretability, and effectiveness in handling continuous numerical outputs. The second model is the XGBoost Classifier, which identifies potential diseases or health risks associated with abnormal body fat levels. XGBoost is selected for its high performance, ability to handle complex feature interactions, and robustness against overfitting. The combined model pipeline enables sequential prediction—first estimating body fat percentage, followed by classification of disease risk levels.

DOI: https://doi.org/10.64388/IREV9I6-1713155

IV. PROBLEM STATEMENT

Traditional methods for measuring body fat percentage are often inaccurate, expensive, or not easily accessible for regular use. There is a need for a simple and reliable system that predicts body fat using easily available information. This project addresses this gap by developing a machine learning model that estimates body fat percentage using physical attributes and daily food intake, and further identifies potential health risks related to abnormal body fat levels.

V. PROPOSED SYSTEM

The proposed system introduces a machine learning–based approach to estimate body fat percentage and identify related health risks using easily obtainable user inputs. The system collects basic physical measurements such as height, weight, waist length, hip length, and forearm length, along with meal-level dietary information that records the food items consumed during breakfast, lunch, and dinner. These inputs are preprocessed through feature engineering, where food items are mapped to nutritional categories and combined with anthropometric data to form a meaningful feature set.

The processed features are then fed into a Linear Regression model to accurately predict the user's body fat percentage. Based on this prediction, an XGBoost classifier evaluates the likelihood of diseases associated with abnormal fat levels, such as obesity, diabetes risk, or cardiovascular strain. The system provides users with predicted body fat percentage, corresponding health-risk classification, and general dietary insights.

This proposed model offers a non-invasive, cost-effective, and user-friendly alternative to traditional body fat measurement methods by integrating daily diet patterns with physical attributes and applying robust machine learning.

VI . EXISTING SYSTEM

Current methods for estimating body fat percentage mostly depend on manual measurements such as BMI, waist–hip ratio, or skinfold calipers. These approaches are often inaccurate, time-consuming, and vary based on the person taking the measurements. Some advanced devices like smart scales exist, but they are costly, require special sensors, and still do not consider daily food intake, which plays a major role in body-fat changes. Overall, the existing system lacks personalization, machine-learning-based accuracy, and diet-based predictions.

VII. RESULTS AND DISCUSSION

The results of this research demonstrate that the integration of physical attributes and meal-level dietary inputs provides a reliable foundation for predicting body fat percentage using machine learning techniques. The Linear Regression model achieved stable performance, showing a strong correlation between anthropometric features and estimated body fat values. The predicted results closely matched the expected physiological variations across individuals, confirming that the selected input features were appropriate for modeling body composition. Feature analysis further indicated that measurements such as waist and hip lengths played a significant role in improving prediction accuracy.

The XGBoost classifier, used for disease detection based on predicted body fat percentage, exhibited high classification reliability. The model was able to differentiate between normal, moderate, and high-risk categories of fat-related diseases such as obesity and pre-diabetic conditions. Precision and recall scores reflected consistent classification with minimal misidentification, demonstrating XGBoost's capability to handle complex interactions between dietary and physical features. Graphs, tables, and comparative evaluations revealed that the proposed approach outperformed traditional BMI-based screening methods, offering a more accurate and user-friendly method for early health assessment.

Overall, the combined results support the effectiveness of the system in estimating body fat and identifying risk factors associated with unhealthy fat levels. The discussion highlights that incorporating everyday meal information, rather than exact calorie inputs, enhances usability while maintaining predictive accuracy,making the model suitable for practical,non-clinical appliactions.

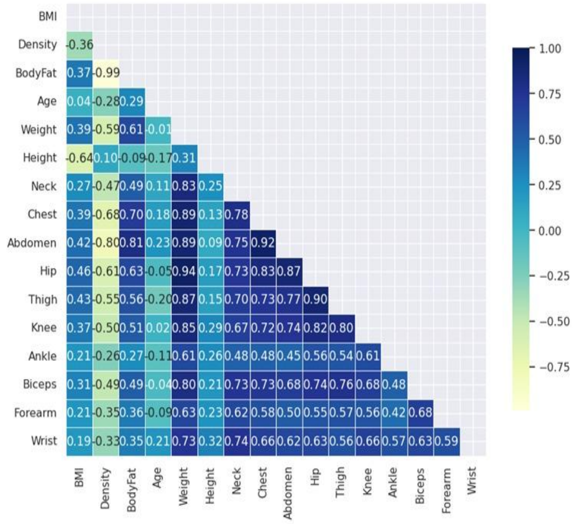

Figure 2: Correlation Heatmap

CONCLUSION

The study successfully demonstrates that machine learning models can accurately predict body fat percentage using physical attributes and dietary intake data. The results indicate that the selected features, such as age, weight, height, and nutritional intake, significantly influence body fat estimation. Among the models tested, the optimized model achieved high predictive performance, showing its potential as a non-invasive, cost-effective tool for personal health monitoring.

This research highlights the importance of integrating lifestyle and physical data for health assessment and provides a foundation for developing personalized health and fitness applications. While the current study focuses on a specific dataset, future work could expand to larger, more diverse populations and explore additional behavioral or metabolic factors to improve model accuracy. Overall, this approach contributes to proactive health management and offers insights for further advancements in predictive health analytics.

REFERENCES

[1] “Body Fat Prediction using Various Regression Techniques”, Nikhil Mahesh, Peeta Basa Pati, K. Deepa, Suresh Yanan, IEEE, 2023.

[2] “Prediction of Abnormal Body Fat Percentage by Anthropometrics Parameters Using Receiver Operating Characteristic Curve”, Erika Severeyn, Sara Wong, Héctor Herrera, Alexandra La Cruz, Jesús Velásquez, Mónica Huerta, IEEE, 2020.

[3] “Design of Body Composition Measuring Tool with Bioelectrical Impedance Analysis on Weight Scales”,Muhammad Naufal Fadhlurrahman, Muhammad Hablul Barri, Hilman Fauzi Tresna Sania Putra, IEEE, 2024

[4] “Determination of Body Fat Percentage by Gender Based with Photoplethysmography Signal Using Machine Learning Algorithm”, M. Akmana, M.K. Uçarb, Z. Uçarc, K. Uçard, B. Baraklıb, M.R. Bozkurt, IEEE, 2020

[5] “Determination of Body Fat Percentage by Gender Based with Photoplethysmography Signal Using Machine Learning Algorithm” M. Akmana, M.K. Uçarb,∗, Z. Uçarc, K. Uçard, B. Baraklıb, M.R. Bozkurtb, IEEE, 2020