# CHAPTER 62

# Smart Health Forecasting: Predictive Health Scoring and Classification using Random Forest on Wearable Device Data

*Supriya Sapa* and Sharvari Tamane***

## ABSTRACT

Wearable devices like smartwatches and fitness bands collect a large amount of raw health data, but this data is often difficult for users to understand on its own. This research focuses on processing that raw data and applying Machine Learning (ML) techniques to give users a clear insight into their health. In this research Random Forest algorithm is used for both classification and regression. The results from Random Forest help users know their health status in a simple way. For example, in the classification problem, if someone falls into the "poor health" category or has a health score below 40 in the regression problem, they can be advised to consult a doctor. The data was collected from 556 users for 9 months to one year using devices such as Apple, Samsung, Xiaomi, and Huawei smartwatches and fitness bands. The data included information like step count, distance walked, flights climbed, calories burned, type of exercise and many more. After applying feature engineering, four key features were selected to predict the health score. The study shows that Random Forest can successfully turn raw wearable data into meaningful health insights for users.

**Keywords:** Random forest; Wearable devices; Health score; Classification; Regression.

## 1.0 Introduction

Wearable devices such as smartwatches and fitness bands have rapidly emerged as essential tools for personal health monitoring and digital healthcare. These devices are equipped with advanced sensors capable of continuously capturing physiological and behavioural data, including heart rate, step count, blood oxygen saturation ($SpO_2$), calories burnt, and sleep activity.

**Corresponding author; Assistant Professor, I.T. Department, IMSCD&R, Ahilyanagar, Maharashtra, India (E-mail: sapasupriya@gmail.com)*
***HOD, UDICT, MGM University, Ch. Sambhajinagar, Maharashtra, India*
*(E-mail: hoditjnec@mgmu.ac.in)*

*DOI: 10.17492/JPI/ICISI2025/251262* *ISBN: 978-93-49790-69-8*

The growing popularity of wearables is driven by their accessibility, ease of use, and ability to provide real-time feedback, enabling individuals to take a more active role in managing their health. Wearables can be broadly classified into categories such as activity trackers, smartwatches, medical-grade wearables, and smart clothing, depending on their primary purpose and functionality. While consumer-grade devices (e.g., Fitbit, Apple Watch, Garmin) focus on general fitness and wellness, clinical-grade wearables are designed for medical monitoring and disease management. Regardless of classification, the reliability of wearable data is highly dependent on correct device usage. For instance, improper placement of a smartwatch, loose fitting of fitness bands, or inaccurate skin contact can lead to significant deviations in recorded measurements, thereby affecting the accuracy of health assessments.

In addition to accuracy, data privacy and security remain major concerns in wearable technology. These devices continuously collect sensitive health information, which, if improperly stored or transmitted, can pose risks of unauthorized access, identity theft, and misuse. Ensuring secure data transmission, encryption, and compliance with privacy regulations is therefore critical in wearable-based healthcare research.

Given the volume and complexity of data generated by wearables, traditional analysis methods are often insufficient to uncover meaningful health patterns. Machine learning techniques, particularly ensemble methods such as Random Forest, provide a robust framework for handling high-dimensional data, managing noise, and improving prediction accuracy. In this study, wearable data are analysed and processed using the Random Forest algorithm to predict user health scores and classify individuals into health categories. By integrating wearable sensor data with machine learning, this research aims to provide more reliable and interpretable insights into personal health status.

### 1.1 Problem statement

Wearable devices such as smartwatches and fitness bands generate vast amounts of physiological and activity-related data, including heart rate, step count, calories burnt, and blood oxygen saturation. While these digital biomarkers have significant potential for health monitoring, the raw data are often be noisy, inconsistent, and difficult for users to interpret meaningfully. Many existing studies analyse wearable data in isolation or with limited machine learning methods, which restricts the accuracy and reliability of health predictions. Moreover, challenges such as device placement accuracy, variations across brands, and concerns regarding data privacy and security further complicate the effective use of wearable technology in healthcare. There is a clear need for a robust, reliable, and interpretable approach that can integrate multiple features from wearable devices to predict an individual's health status more accurately.

DOI: 10.17492/JPI/ICISI2025/251262 ISBN: 978-93-49790-69-8

## 2.0 Literature Review

*Wearable activity trackers, accuracy, adoption, acceptance and health impact: A systematic literature review,* Authors: Grace Shin, Mohammad Hossein Jarrahi, Fei Y, Amir Karami, Nicci Gafinowitz, Ahjung Byun, Journal: *Journal of Biomedical Informatics* Volume & Article ID: Volume 93, Article 103153 (May 2019). Wearable activity trackers (WAT) are electronic monitoring devices that enable users to track and monitor their health related physical fitness metrics including steps taken, level of activity, walking distance, heartrate, and sleep patterns. Objective of the study is to understand the rich human-information interaction that is enabled by WAT adoption.

In the study Topic modeling methods were used to identify six key themes of WAT research, namely Technology Focus, Patient Treatment and Medical Settings, Behavior Change, Acceptance and Adoption, Self monitoring Data Centered, and Privacy. This work raises interdisciplinary awareness about the current landscape of WAT use and the related diversity of interesting research opportunities and challenges.

The study suggests that WAT devices are multi-dimensional technologies with complex impacts. Understanding WAT and their technological and non-technological aspects requires various research perspectives. This multi-dimensional framework highlights that WATs are not merely fitness devices but are embedded in complex human-information interactions. The review emphasizes the need for interdisciplinary approaches to fully understand the technological, behavioral, and ethical dimensions of WATs, and opens avenues for future research in both health informatics and user-centered design.

Wearable Sensors for Remote Health Monitoring, Authors**:** Sumit Majumder, Tapas Mondal and M. Jamal Deen - Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada; Journal**:** *Sensors* (an open-access journal published by MDPI)-2017**.** Remote health monitoring, based on non-invasive and wearable sensors, modern communication and information technologies offers an efficient and cost-effective solution that allows the elderly to continue to live in their comfortable home environment instead of expensive healthcare facilities.

These systems will also allow healthcare personnel to monitor important physiological signs of their patients in real time, assess health conditions and provide feedback from distant facilities. In this paper, they have presented and compared several low-cost and non-invasive health and activity monitoring systems. Finally, compatibility of several communication technologies as well as future perspectives and research challenges in remote monitoring systems are discussed.

Machine Learning for Healthcare Wearable Devices: The Big Picture, Authors**:** Farida Sabry, Tamer Eltaras, Wadha Labda, Khawla Alzoubi, Qutaibah Malluhi, Journal:

Journal of Healthcare Engineering (Hindawi)-2022. The paper highlights the Machine Learning Techniques used, the different modalities used, and the available datasets and the different challenges facing machine learning applications on wearable devices like deployment alternatives, power consumption, storage and memory, utility and user acceptance, data availability and reliability, communication, security and privacy were discussed while identifying possible solutions found in the literature. The objective of the study is to highlight the various ML techniques and the challenges in the deployment of wearable devices.

The methods used are applied on datasets available for human activity recognition. K-NN, SVM, LR, Tree-based, Deep learning models are used for analysis of the data. Further research concerning data availability, reliability, and privacy to enable effective and efficient learning from data generated by wearable devices. The wearable devices are used for remote patient monitoring and detection of any irregularities with the human body Consumers' and Physicians' Perceptions about High Tech Wearable Health Products, Authors**:** Suphan Nasira, Yigit Yurdera- Istanbul University, Faculty of Economics, Journal/Conference**:** Procedia – Social and Behavioral Sciences, Volume**:** 195 (2015). This study aims to explore and compare the perceptions of both consumers and healthcare professionals (physicians) toward high-tech wearable health technologies, which include smartwatches, fitness trackers, and similar digital health-monitoring devices. Recognizing the growing popularity of wearable devices in personal healthcare, the authors seek to understand the underlying factors that influence the acceptance and adoption of such technologies among different user groups.

To achieve this, the researchers extend the Technology Acceptance Model (TAM)—a widely used theoretical framework in technology adoption studies. While the original TAM focuses on perceived usefulness and perceived ease of use as the primary determinants of technology acceptance, this study enhances the model by incorporating two additional constructs: perceived risk and compatibility. Perceived risk refers to the concerns users may have regarding the reliability, privacy, or potential misuse of health data collected by wearable technologies. Compatibility assesses the degree to which wearable health technologies align with an individual's lifestyle, habits, and healthcare needs. By integrating these two constructs into TAM, the study offers a more comprehensive framework to examine how both psychological and practical considerations influence the decision-making process.

The study employs empirical data collection through surveys conducted with both consumers and physicians. The responses are analysed to identify patterns in perception and adoption behaviour.

*DOI: 10.17492/JPI/ICISI2025/251262* *ISBN: 978-93-49790-69-8*

## 3.0 Literature Gap

Several studies have explored the role of wearable devices in health monitoring, demonstrating their potential in tracking physical activity, cardiovascular parameters, and overall well-being. Prior research has primarily focused on step count analysis, heart rate variability, and sleep monitoring to provide general health recommendations. Machine learning models have also been applied to wearable data, but many works emphasize either a single parameter (e.g., heart rate or step count) or employ conventional regression approaches with limited predictive accuracy. While ensemble learning methods like Random Forest are recognized for their robustness and ability to manage noisy, high-dimensional data, their application in comprehensive health score prediction from multi-parameter wearable datasets remains underexplored. This gap highlights the need for research that not only integrates multiple digital biomarkers (e.g., heart rate, $SpO_2$, step count, calories) while applying advanced machine learning techniques for improved prediction and classification of health outcomes.

## 4.0 Research Objectives

The main objectives of this research are:

- To process and analyse multi-parameter data from wearable devices (including heart rate, step count, calories burnt, and $SpO_2$) for comprehensive health assessment.
- To apply Random Forest machine learning technique for predicting health scores and classifying users into health categories (e.g., poor, average, good).
- To develop a framework that provides users with meaningful health insights derived from wearable data, enabling better awareness and preventive healthcare.

## 5.0 Research Methodology

This research adopts a supervised machine learning-based predictive and classification methodology to evaluate and classify the users into a health category by forecasting an individual's yearly health score based on physiological and activity data derived from wearable devices.

### 5.1 Data collection

The dataset was collected over a continuous period of one year, for 556 individuals, each monitored monthly. The data includes:

- Step Count (monthly total)

*DOI: 10.17492/JPI/ICISI2025/251262* *ISBN: 978-93-49790-69-8*

- Walking Speed
- Exercise Time
- Flights Climbed
- Calories Burnt (monthly total)
- Average Heart Rate
- Blood Oxygen Saturation
- Demographics**:** Age and Gender
- Health Score (target variable)

The values were recorded consistently each month, creating a detailed time series per participant. Data was consolidated in CSV format and later pre-processed for machine learning analysis.

### 5.2 Feature engineering

From the raw monthly data, 4 features were considered and their aggregated features were derived:

- Avg_Heart_Rate: Mean heart rate across all months
- Avg_Blood_Oxygen_Saturation: Average oxygen level across the year
- Total_Step_Count: Sum of step counts from July to June
- Total_Calories_Burnt: Annual total calorie expenditure
- Age and encoded Gender

These engineered features formed the independent variables (X), while Health Score was treated as the dependent variable (y).

### 5.3 Data pre-processing

The data was cleaned and normalized where needed. No significant missing values were reported. Categorical encoding was applied to Gender for model compatibility. The dataset was then split into training (80%) and testing (20%) subsets using stratified random sampling to preserve data balance.

### 5.4 Model selection and training

To predict the *Yearly Health Score (out of 100),* the regression model used was Random Forest for Regression and then 3 Categories of health were generated poor, average and good based on the ranges in health score like below 50 poor, in between 50 to 80 average and above 80 as good. Random Forest Classifier was used for this

### 5.5 Random forest

Random Forest is an ensemble learning method that builds multiple decision trees and combines their outputs to improve prediction accuracy and reduce overfitting. Unlike

DOI: 10.17492/JPI/ICISI2025/251262 ISBN: 978-93-49790-69-8

sequential methods such as Gradient Boosting, Random Forest builds trees independently and in parallel, and aggregates their results through majority voting (for classification) or averaging (for regression).

### 5.6 Algorithmic steps

1. From the training dataset, generate multiple bootstrap samples (random sampling with replacement).
2. For each bootstrap sample, grow a decision tree:
   - At each node, a random subset of features is selected.
   - The best split is chosen only from this subset, introducing randomness.
3. Each tree is grown fully without pruning, making them high-variance, low-bias learners.
4. Aggregate predictions from all trees:
   - Classification: use majority voting.
   - Regression: use the average of all predictions.

*Mathematical Objective Function:*

For regression:

$$\frac{f_{RF}}{x}(x) = \frac{1}{M}\sum_{m=1}^{M} T_{m})$$

Where:

- $f_{RF}(x)$ is the Random Forest prediction for input x
- M is the total number of trees in the forest
- Tm(x) is the prediction from the m-th decision tree

For classification:

$$= \text{mode}\left\{ h_1(x),\, h_2(x) \right.$$

$$\hat{y} = \text{mode}\{h_1(x), h_2(x), \ldots, h_T(x)\}$$

Where:

- T = total number of trees
- $h_t(x)$ = prediction from the t $^{th}$ decision tree
- y^ = final Random Forest prediction

### 5.7 Expected outcomes of random forest

- High accuracy due to averaging across many trees.
- Robustness to noise and overfitting (better generalization than a single tree).
- Handles high-dimensional data well, including datasets with many features.
- Less interpretable compared to a single decision tree.
- Computationally intensive if too many trees are used.

DOI: 10.17492/JPI/ICISI2025/251262 ISBN: 978-93-49790-69-8

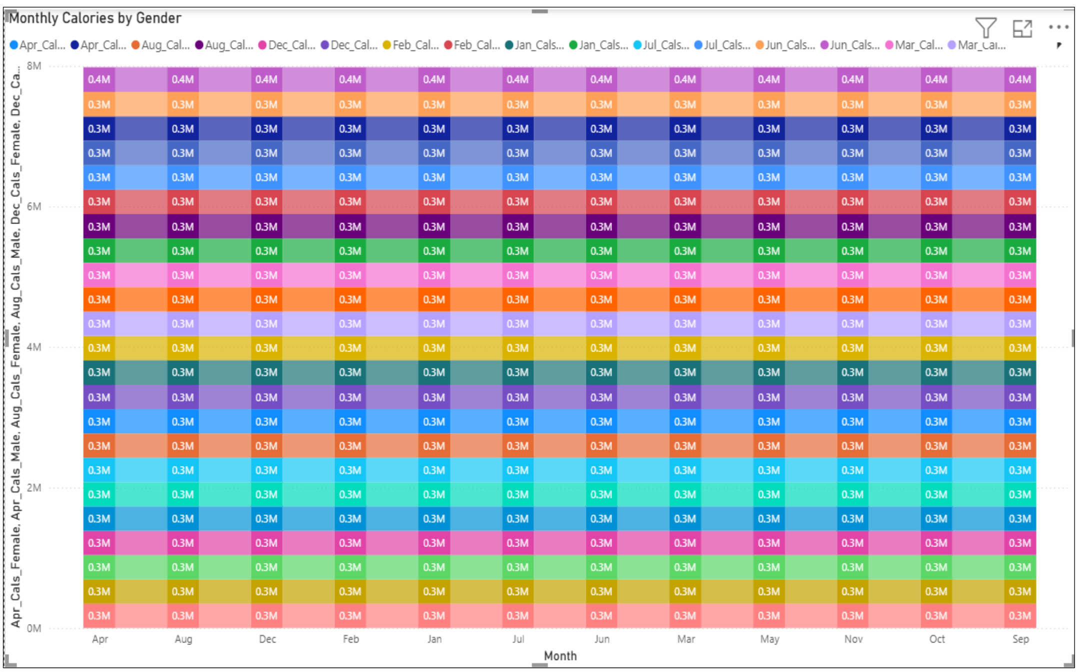
### 5.8 Data overview

The dataset used in this research was collected over a continuous 9 months to one-year period, covering health-related metrics of 556 individuals**.** Each individual's data was recorded monthly from July to June, ensuring consistency and completeness across all entries. The primary aim was to capture longitudinal trends in physical activity and vital signs that contribute to overall health.

The dataset includes key physiological and activity-based indicators such as monthly step count**,** calories burnt**,** average heart rate**,** and blood oxygen saturation. Demographic details like age and gender were also included to study potential variations in health score patterns across different population segments.

The central focus of the analysis is the Yearly Health Score, which acts as the target variable for prediction. To prepare the data for analysis, the raw monthly readings were aggregated to create meaningful features such as total annual step count**,** average heart rate across the year**,** cumulative calories burnt**,** and mean blood oxygen levels.

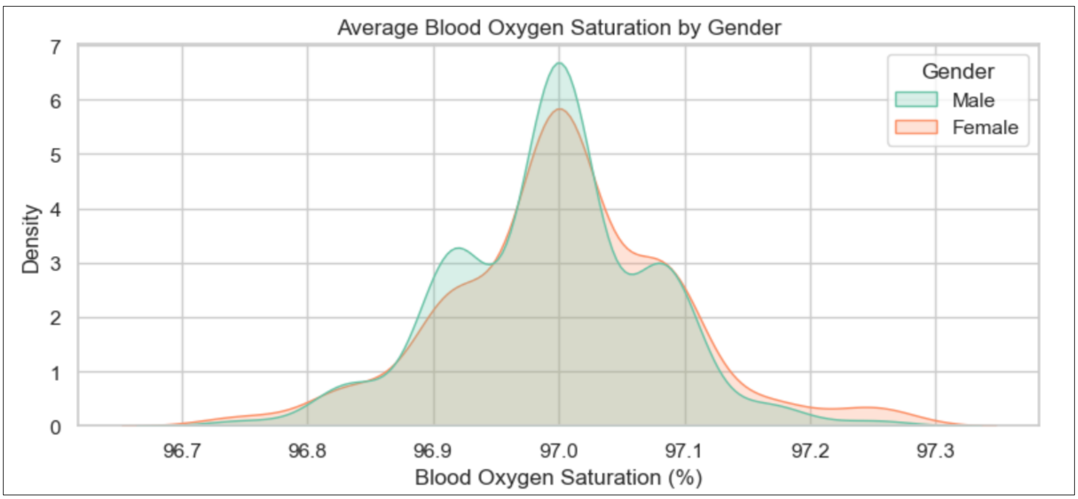
This aggregation ensured dimensional consistency and improved the quality of the feature set used for machine learning models. Categorical encoding was applied to the gender variable to make it model-compatible. The dataset exhibited a balanced gender distribution and a healthy spread across various age groups, making it suitable for generalizable modelling.



No major missing values or anomalies were detected, minimizing the need for imputation. The dataset's time-bound and structured nature made it highly appropriate for supervised machine learning tasks like regression, providing a reliable foundation for prediction model and classification for classifying the users into different health categories.

### 5.9 Exploratory data analysis (EDA) with visualizations

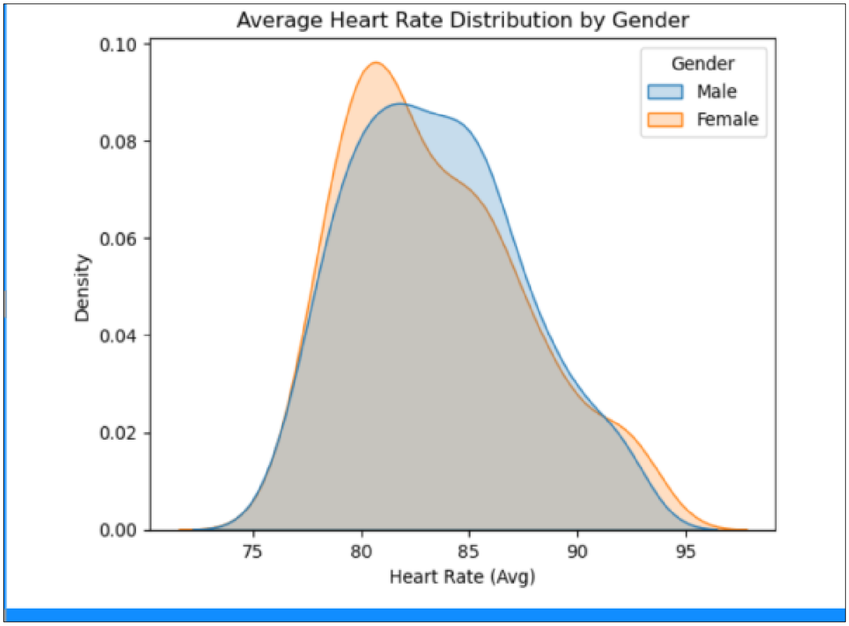
Exploratory Data Analysis (EDA) was conducted to understand the distributions, patterns, and relationships among the features before applying machine learning models. It played a vital role in identifying key predictors for the yearly health score, checking for data quality issues, and preparing visual insights that could be compared across gender and age groups. The first step of EDA involved a correlation heatmap to examine interdependencies among variables. It was observed that Avg_Heart_Rate and Total_Calories_Burnt showed moderate positive correlation with the Yearly_Health_Score, indicating their significance in predicting overall wellness. On the other hand, Avg_Blood_Oxygen_Saturation remained relatively stable, contributing less directly to score variance but still relevant as a physiological marker.

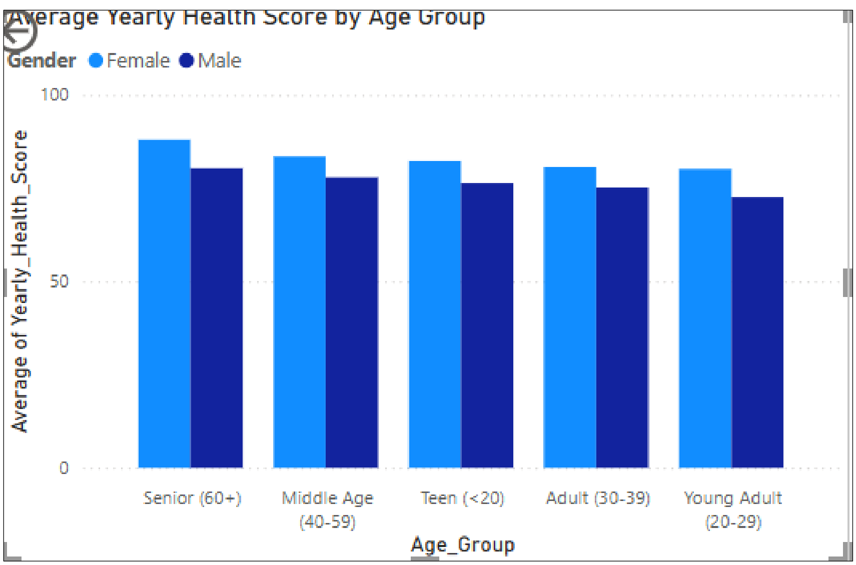



To explore gender-wise patterns, Kernel Density Estimate (KDE) plot (also called a density plot) were created for Average Blood Oxygen Saturation (%) and Average Heart Rate. These visuals highlighted that males had a slightly wider range and higher blood oxygen saturation compared to females, whereas heart rate distributions remained similar across both groups. Age distribution was visualized using a histogram, showing that most participants fell within the 20–40 age range.

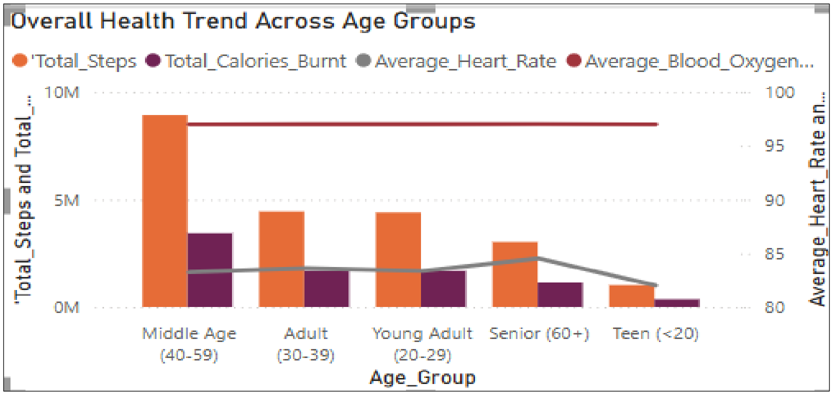
DOI: 10.17492/JPI/ICISI2025/251262 ISBN: 978-93-49790-69-8

Further EDA focused on age-based trends, where a line chart was plotted to show changes in Avg_Heart_Rate and Avg_Blood_Oxygen across age intervals (e.g., 20–30, 31–40, etc.). The analysis revealed that average heart rate tends to decline slightly with age, while oxygen saturation remains consistently high, regardless of age group.




DOI: 10.17492/JPI/ICISI2025/251262 ISBN: 978-93-49790-69-8




These all visualizations were done using Power BI for a clearer interpretation of the data and their inter relationships.

### 5.10 Model development

The machine learning models selected for this study were Random Forest Regressor and Classifier. These models were chosen to compute the health score using ensemble-based tree methods. The goal was to accurately predict the Yearly Health Score using a set of engineered features derived from one year's worth of physiological and activity data and then classify the users based on this health score into different classes.

Each model was trained using the pre-processed dataset split into 80% training and 20% testing sets. Feature columns included: Total_Calories_Burnt, Total_Step_Count, Avg_Heart_Rate, Avg_Blood_Oxygen_Saturation, Age, and encoded Gender. These features were selected based on insights from EDA that showed moderate correlation and explanatory potential regarding the target variable. Hyperparameter tuning was applied where necessary using techniques such as grid search to enhance model performance and reduce overfitting. Random Forest used decision tree ensembles to sequentially minimize residual errors. The model was evaluated using regression metrics including $R^2$ Score**,** Mean Squared Error (MSE)**,** Mean Absolute Error (MAE), and a custom Accuracy % metric. The classification metrics included Accuracy, Confusion Matrix, Precision, Recall and F1 Score.

### 5.11 Results and outcome

The predictive performance of Random Forest Model was evaluated using standard regression metrics: Mean Squared Error (MSE)**,** $R^2$ Score, and a custom-defined Accuracy (%). RandomForestRegression - MSE: 6.48, $R^2$: 0.92.

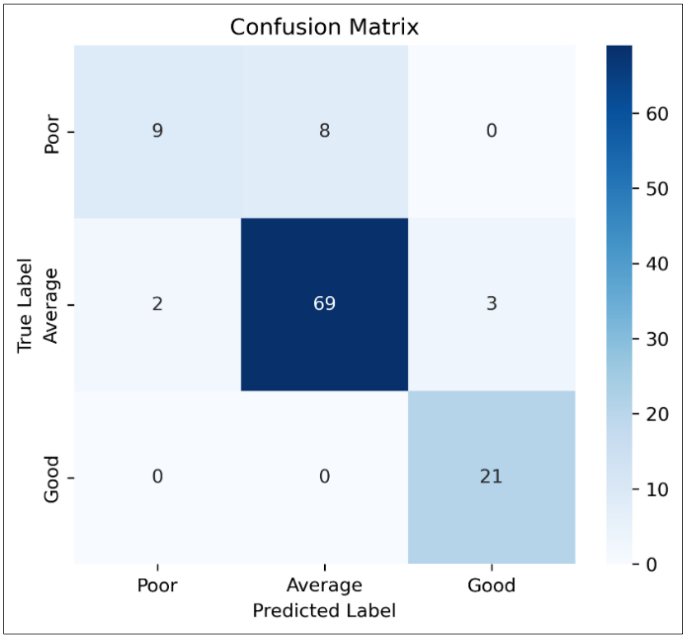
Overall Accuracy: RandomForestRegression - 97.47%

Random Forest Classification Model was evaluated using standard metrics: Accuracy, Confusion Matrix, Precision, Recall and F1 Score.

*DOI: 10.17492/JPI/ICISI2025/251262* *ISBN: 978-93-49790-69-8*

Accuracy: 0.8839285714285714
Confusion Matrix:
[[ 9 8 0]
[ 2 69 3]
[ 0 0 21]]
Precision: 0.8803194573283859
Recall: 0.8839285714285714
F1-Score: 0.8764081294769562

**Figure 1: Confusion Matrix**




Row 1 (Actual = Poor): [9, 8, 0]

- 9 correctly predicted as Poor (True Positives for Poor).
- 8 misclassified as Average.
- 0 predicted as Good.

Out of 17 actual Poor samples, only 9 were correct → accuracy for this class is moderate.
Row 2 (Actual = Average): [2, 69, 3]

- 69 correctly predicted as Average (True Positives for Average).
- 2 misclassified as Poor.
- 3 misclassified as Good.

Most Average samples are classified correctly → strong performance for this class.
Row 3 (Actual = Good): [0, 0, 21]

- 21 correctly predicted as Good.
- 0 misclassified as Poor or Average.

Perfect classification for Good class.
First 10 Predictions:

| Actual | Predicted | Accuracy (%) |
|---|---|---|
| 76.08 | 78.49 | 96.83 |
| 84.62 | 83.63 | 98.82 |
| 68.42 | 60.10 | 87.85 |
| 70.71 | 70.61 | 99.86 |
| 59.96 | 62.14 | 96.37 |
| 67.79 | 66.15 | 97.58 |
| 77.67 | 77.68 | 99.98 |
| 59.96 | 60.28 | 99.46 |
| 83.50 | 80.94 | 96.94 |
| 82.75 | 81.69 | 98.72 |

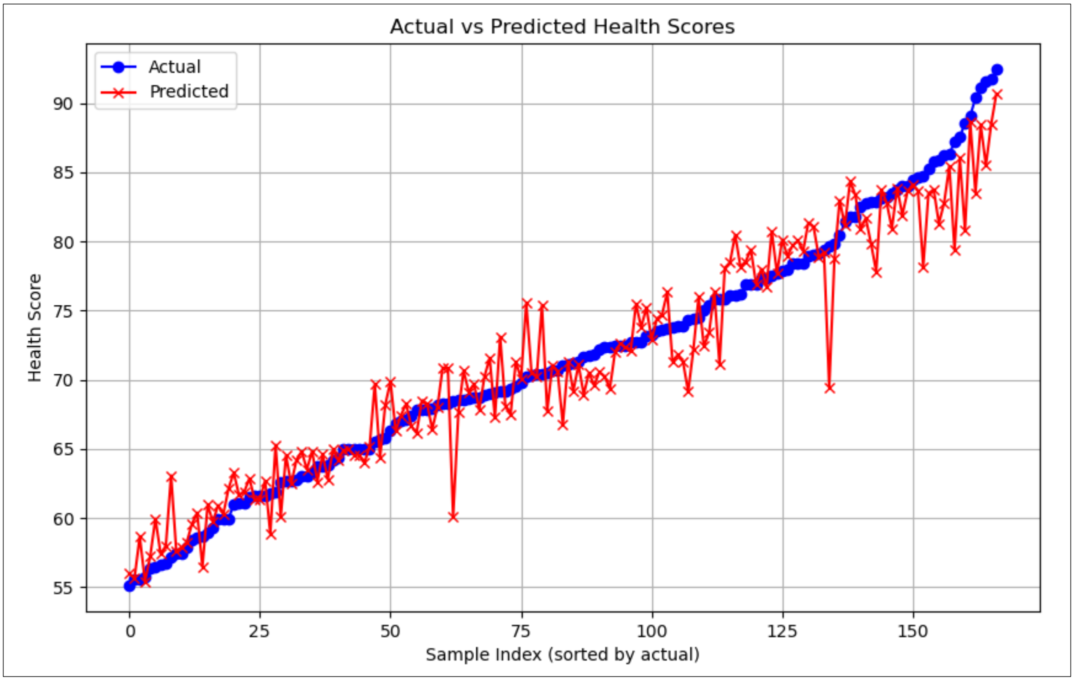
Last 10 Predictions:

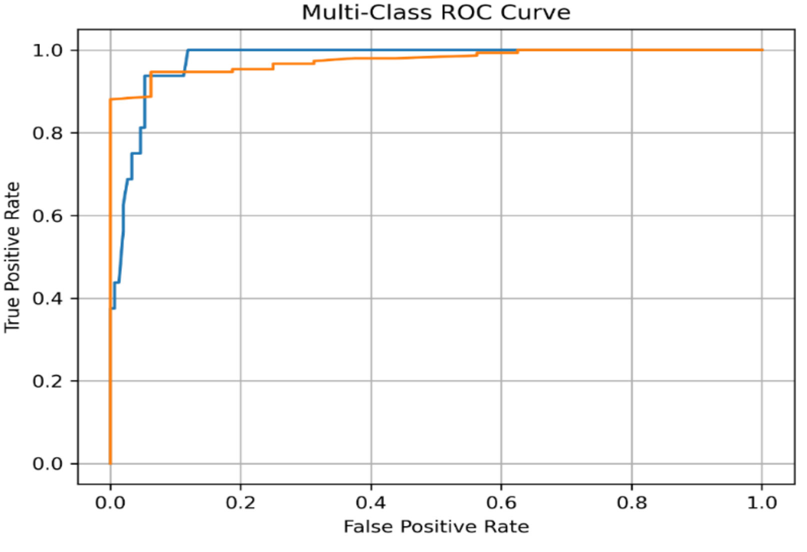
| Actual | Predicted | Accuracy (%) |
|---|---|---|
| 70.25 | 75.55 | 92.46 |
| 76.92 | 79.42 | 96.75 |
| 83.08 | 83.72 | 99.23 |
| 83.21 | 82.80 | 99.51 |
| 74.29 | 69.15 | 93.08 |
| 72.17 | 70.57 | 97.79 |
| 78.42 | 79.30 | 98.87 |
| 79.62 | 69.45 | 87.22 |
| 77.33 | 78.01 | 99.12 |
| 87.25 | 79.37 | 90.96 |

Table 1 displays the first and last 10 predictions made by the regression. Each row includes the actual value, the values predicted by the model, and their respective accuracy percentages. The table reveals that the model perform strongly, consistently shows accuracy above 97%. The model demonstrates robust predictions, though with slightly more variation in accuracy, particularly in the first 10 records. This suggests the model is reliable, and provides slightly more stable performance on this dataset.

*DOI: 10.17492/JPI/ICISI2025/251262* *ISBN: 978-93-49790-69-8*

The graph below illustrates the actual values against the predicted values , clearly highlighting their accuracy levels. The plotted data points confirm that the models closely follow the actual trends, exhibiting higher predictive stability across the observed range.




DOI: 10.17492/JPI/ICISI2025/251262 ISBN: 978-93-49790-69-8

In the above Multi-Class ROC Curve

- TPR (True Positive Rate) = Recall = Sensitivity
  TPR = TP / TP+FN
  Measures how many actual positives are correctly identified.
- FPR (False Positive Rate)
  FPR=FP / FP+TN
  Measures how many actual negatives are incorrectly classified as positives.
- ROC Curve (Receiver Operating Characteristic) **-** Plots TPR vs FPR at different threshold values. A curve closer to the top-left corner indicates better classification performance.
- Both curves rise steeply toward the top-left, which means High TPR (Sensitivity) and Low FPR
- This shows that the classifier has high discriminative power for separating classes. The ROC curve confirms the model is strong (likely with Area Under Curve close to 1).

## 6.0 Conclusion

This study demonstrates the effectiveness of Random Forest models for analysing wearable device data to assess individual health status. The Random Forest Regression achieved a high predictive performance (MSE = 6.48, $R^2$ = 0.92, Accuracy = 97.47%)**,** indicating its strong ability to estimate continuous health scores with minimal error. Similarly, the Random Forest Classification model showed robust results (Accuracy = 88.39%, Precision = 0.88, Recall = 0.88, F1-score = 0.87), confirming its suitability for categorizing health into discrete groups of poor, average, and good. The results suggest that Random Forest models are capable of capturing complex, nonlinear patterns in physiological and activity-related features such as heart rate, blood oxygen saturation, step count, and calories burned. By leveraging these digital biomarkers, the models provide accurate and reliable predictions that can support personalized health monitoring and early detection of health risks**.** In conclusion, Random Forest proves to be a powerful machine learning approach for both predictive and diagnostic applications in digital health, highlighting its potential to contribute to smarter, data-driven healthcare solutions.

## References

1. Ali, J., Khan, R., Ahmad, N., & Maqsood, I. (2012). Random Forests and Decision Trees. *ResearchGate*. https://www.researchgate.net/publication/259235118

2. Chen, Y., Wang, J., Yu, C., Gao, W., & Qin, X. (2021). *FedHealth: A Federated Transfer Learning Framework for Wearable Healthcare* (arXiv:1907.09173). arXiv. https://doi.org/10.48550/arXiv.1907.09173
3. Chiang, P.-H., Wong, M., & Dey, S. (2021). Using Wearables and Machine Learning to Enable Personalized Lifestyle Recommendations to Improve Blood Pressure. *IEEE Journal of Translational Engineering in Health and Medicine*, *9*, 1–13. https://doi.org/10.1109/JTEHM.2021.3098173
4. Neshitov, A., Tyapochkin, K., Kovaleva, M., Dreneva, A., Surkova, E., Smorodnikova, E., & Pravdin, P. (2023). Estimation of cardiorespiratory fitness using heart rate and step count data. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-43024-x
5. Odhiambo, C., Wright, P., Corbett, C., & Valafar, H. (n.d.). *MedSensor: Medication Adherence Monitoring Using Neural Networks on Smartwatch Accelerometer Sensor Data*.
6. Prasad Raju, N. V. D. S. S. V., & Devi, P. N. (2024). A Comparative Analysis of Machine Learning Algorithms for Big Data Applications in Predictive Analytics. *International Journal of Scientific Research and Management (IJSRM)*, *12*(10), 1608–1630. https://doi.org/10.18535/ijsrm/v12i10.ec09
7. Sabry, F., Eltaras, T., Labda, W., Alzoubi, K., & Malluhi, Q. (2022). Machine Learning for Healthcare Wearable Devices: The Big Picture. *Journal of Healthcare Engineering*, *2022*, 1–25. https://doi.org/10.1155/2022/4653923
8. Salman, H. A., Kalakech, A., & Steiti, A. (2024). Random Forest Algorithm Overview. *Babylonian Journal of Machine Learning*, 2024, 69–79. https://doi.org/10.58496/BJML/2024/007
9. Yusuf, A., Jaber, T. A., & Gordon, N. (2025). Comprehensive Health Tracking Through Machine Learning and Wearable Technology. *Journal of Data Science and Intelligent Systems*. https://doi.org/10.47852/bonviewjdsis52023588

*DOI: 10.17492/JPI/ICISI2025/251262* *ISBN: 978-93-49790-69-8*