7/8/2023

Aarti Ramani

Bellevue University

Commercial Wearable Devices Predicting Physical Activity

Predictive Analytics Case Study

Contents

[**INTRODUCTION** 2](#_Toc139748390)

[Problem Statement: 2](#_Toc139748391)

[Problem Important to Solve: 2](#_Toc139748392)

[Data Collection: 2](#_Toc139748393)

[**METHODS** 5](#_Toc139748394)

[Data Preparation & Problem Solved: 5](#_Toc139748395)

[Modeling Techniques: 5](#_Toc139748396)

[Reason for Model Selection: 5](#_Toc139748397)

[Model Evaluation: 6](#_Toc139748398)

[**CONCLUSION** 6](#_Toc139748399)

[Results: 6](#_Toc139748400)

[Actionable Consequences: 7](#_Toc139748401)

[Learnings: 7](#_Toc139748402)

[Future addressing of the problem: 8](#_Toc139748403)

[**REFERENCES:** 9](#_Toc139748404)

# **INTRODUCTION**

Commercial Wearable devices have become an integral part of people's fitness journey. These commercial devices offer a range of tracking features, such as step counting, heart rate monitoring, sleep tracking, calorie expenditure, etc. For this study, data from two wearable devices, Apple Smartwatch Series 2, and Fitbit Charge HR2 was analyzed, as these devices had the highest market share among wearable devices at the time of study (2020).

## Problem Statement:

Apple Watch calculated active calories without considering the base metabolic rate, whereas Fitbit reports the energy expenditure per minute, which accounts for sedentary behavior. It is assumed that wearable devices can accurately track moderate to high/vigorous intensity activities but are not accurate in tracking low to sedentary behavior. The primary goal of this study is to examine if commercial wearable devices can accurately track low to sedentary behavior. Secondly, the study also examines the influence of the wearable device type on the classification results.

## Problem Important to Solve:

Individuals who prioritize healthy living often utilize wearable devices, whether consciously or unconsciously. Accurate measurements of features such as energy expenditure, heart rate, and steps are crucial in determining daily energy expenditure and achieving activity goals. Tracking low to sedentary behavior is also important especially important in elderly adults as they may not be capable of engaging in moderate or high-intensity workouts.

## Data Collection:

For the study, 49 individuals (23 males and 26 females) were recruited and provided with wearable devices. The devices were randomly distributed, with participants wearing one on each wrist. Participants were engaged in a total of 65 minutes of activities with 40 minutes of treadmill time and 25 minutes of sedentary time. To ensure the devices recognize changes in intensity, it is important to change activities.

The following sequence was followed for varied activity speeds and are captured under the feature *ActivityLevel*.

* 1. Sedentary activity such as lying, sitting, etc. for 5 minutes
  2. 10 minutes of self-paced walk on a treadmill
  3. 5 minutes of lying down
  4. 10 minutes of treadmill walking at a pace of 3 METs
  5. 5 minutes of lying down
  6. 10 minutes of treadmill walking at a pace of 5 METs
  7. 5 minutes of sitting
  8. 10 minutes of treadmill walking at a pace of 7 METs.

*One metabolic equivalent (*MET*) is defined as the amount of oxygen consumed while sitting at rest and is equal to 3.5 ml O2 per kg body weight x min* (Staying active. The Nutrition Source - 2023, February 2)*.*

Apart from the two wearable devices, participants were also given an iPhone 6S with a custom iOS app called PASS mobile (Physical Activity, Sleep, and Sedentary Behavior Mobile) which captures minute-by-minute data from both wearable devices. The captured data had around 3656 and 2608 minutes of Apple and Fitbit data, respectively. The final data consisted of six classes under the feature *ActivityLevel* – lying, sitting, self-paced walk, walking/running at 3 METs, walking/running at 5 METs, and walking/running at 7 METs. Apart from the existing features (steps, heart rate, energy expenditure, height, weight, age, gender, and distance),additional features such as heart rate entropy, step entropy (a measure of predictability of step count), the correlation coefficient between heart rate and steps, and intensity (Karvonen Formula - Intensity zone during activity), Product measure of the total amount of steps and distance covered in meters, and Standard deviation of normalized heart rate were developed. Additional feature development was required to standardize data across the devices.

The final dataset had the following features:

1. Heart Rate
2. Step Count
3. Energy Expenditure/Calories
4. Distance
5. Age
6. Gender
7. Weight
8. Height
9. Steps entropy – a measure of step variability
10. Resting Heart entropy – 10th percentile of heart rate data
11. Correlation Heart Rate and Steps – correlation coefficient between the 2 features.
12. Intensity (Karvonen Formula) – Intensity zone during a workout
13. Activity Level

# **METHODS**

## Data Preparation & Problem Solved:

The recorded timestamps in devices were converted to a generic format (%H:%M:%S) for consistency. Any columns with missing rows were dropped, as they were insignificant. Lean body mass of individuals and psychological factors such as age and health status have a significant effect on the MET of individuals. Hence, for stages with MET values, the energy expenditure was calculated with the VO2 to METs calculator. This was required because Apple watches do not account for calories burnt with a low to sedentary lifestyle. Missing rows in the features steps, heart rate, calories, and floors were filled using Linear interpolation steps (Linear Interpolation – (Zach - 2022, June 22).

## Modeling Techniques:

With the final dataset, the following four classification models were built with each dataset (Apple and Fitbit data) split into 70% training and 30% testing and validation, with *ActivityLevel* as the target.

* 1. Random Forest
  2. Rotation Forest
  3. Decision Tree
  4. SVM – Support Vector Machine

## Reason for Model Selection:

The SVM and Random Forest Models were chosen as they are the most common models in physical activity studies and Rotation Forest and Decision Tree models are similar to Random Forest.

## Model Evaluation:

Model accuracy was evaluated using accuracy, confusion matrix, feature ranking, sensitivity, and specificity.

* + 1. It is imperative to bear in mind that when using classification models, relying solely on accuracy as a metric to evaluate the effectiveness of the model is inadequate.
    2. A confusion matrix is a useful tool for determining accurate classification and identifying instances of misclassification.
    3. Sensitivity or recall is important when you are concerned with identifying positive outcomes and the cost of a false positive is low (Thieme, C. (2021, June 16)). Sensitivity in this study is important as we want to track as many low to sedentary activities as possible.
    4. Specificity is the ratio of true negatives to total negatives (Thieme, C - 2021, June 16). A high level of specificity is indicative of an accurate model that identifies the most negative results. Conversely, a low specificity implies that the model incorrectly labels numerous negative outcomes as positive.

# **CONCLUSION**

## Results:

Each model outcome for Apple and Fitbit data is as below:

|  |  |  |
| --- | --- | --- |
| **MODEL** | **FITBIT** | **APPLE** |
| Decision Tree | 62.34 | 41.39 |
| SVM | 56.66 | 50.87 |
| Random Forest | 90.80 | 81.95 |
| Rotation Forest | 89.26 | 82.60 |

After evaluating all models, the Rotation Forest model had the highest accuracy for all activities in both Apple and Fitbit devices.

Separately, the secondary goal was to check the significance of device type in the classification.

Chi-Squared feature ranking of the combined dataset (Apple and Fitbit) showed that device type ranked 13th in terms of feature importance and the rotation forest accuracy was around 85.9%, close to the accuracy on individual datasets, making this feature not significant for predictions. Having said that, it is worth noting that we only examined two devices.

## Actionable Consequences:

Between the two devices, Fitbit shows better accuracy than Apple Watch. However, the result may vary with a population-based sample. Also, this study is focused only on two devices. Device type although has a very low significance in the accuracy, might turn out to be of more significance when more devices are included with a population-based sample.

## Learnings:

The rotation forest model had the highest accuracy for classifying all types of activities.

Although this model was able to show a high accuracy, they are still slightly lower than previous research on wearable devices which ranged from 60 to 90%. This could be because, in this study, developed features such as normalized heart rate and intensity using the Karvonen formula are considered important for identifying all types of activities, and previous research considered features focused on identifying moderate activity (100 steps/minute).

Of the two devices, Fitbit has a higher accuracy than the Apple watch. Further analysis using Feature Ranking was done to identify the reason for the difference in accuracies between the devices. The results showed that Apple considered heart rate as the most important feature whereas Fitbit considered step count to be the most important feature.

## Future addressing of the problem:

This study focused on certain derived features while disregarding other features that have been deemed important in previous research for identifying moderate activity. However, the specific features that were considered in previous research are not mentioned. To fully understand the significance of these developed and ignored features, it would be beneficial to expand the study to include both and analyze their importance and potential findings.

The research can further be developed using other wearable devices in addition to Apple and Fitbit. One caveat of this method could be the metrics used across different device types. Different manufacturers use different metrics to measure heart rate, steps, calories, and distance. Further research can be done to account for these algorithmic differences across devices.

Previous research was performed on research-grade devices and not commercial devices. The amount of missing data in commercial devices is more in comparison to research-grade devices. Missing data in this study was dealt with by imputation. Future studies can be performed to identify the impact of imputed data on model accuracy.

# **REFERENCES:**

1. Daniel Fuller, Javad Rahimipour Anaraki, Bo Simango, Faramarz Dorani, Arastoo Bozorgi, Hui Luan, Fabien Basset - 2020, March 12. - *Using machine learning methods to predict physical activity types with Apple Watch and Fitbit data using indirect calorimetry as the criterion. (n.d.).* <https://assets.researchsquare.com/files/rs-17022/v1/d5923374-d56c-4fe7-a036-949ecf41917e.pdf?c=1631831698>
2. Bennet, D. T. - 2022, March 27. *Five fascinating case studies in predictive analysis.* Medium. <https://medium.com/@DeepthiTabithaBennet/five-fascinating-case-studies-in-predictive-analysis-b9bf15216ff7>
3. Staying active. The Nutrition Source -2023, February 2 - <https://www.hsph.harvard.edu/nutritionsource/staying-active/#:~:text=MET%20stands%20for%20the%20metabolic,may%20use%208%2D9%20METs>
4. Zach - 2022, June 22. How to interpolate missing values in R (including example). Statology. <https://www.statology.org/r-interpolate-missing-values/>
5. Thieme, C. - 2021, June 16. Understanding common classification metrics - titanic style. Medium. <https://towardsdatascience.com/understanding-common-classification-metrics-titanic-style-8b8a562d3e32>