Bridging Technology and Health:
A Comprehensive Analysis of Fitbit and Apple Watch in Enhancing User Engagement and Activity Efficiency

Alex Choi, Pranav Patil, Will Corbin

Introduction

Wearable technologies such as Apple Watch and Fitbit are able to track information about our health and fitness down to frequencies of seconds that are aggregated to provide minute-level details. It is of interest to better understand how effective and reliable these devices are in monitoring and improving individuals' physical health.

The objective is to analyze data collected from these technologies to
- Understand how the device type (Apple Watch vs. Fitbit) influences the prediction of heart rate across the different types of activities
- Classify the type of user (age group) based on physical metrics
- Analyze if efficiency of physical activities vary by individual characteristics or device type

Analysis and Methods: The Data

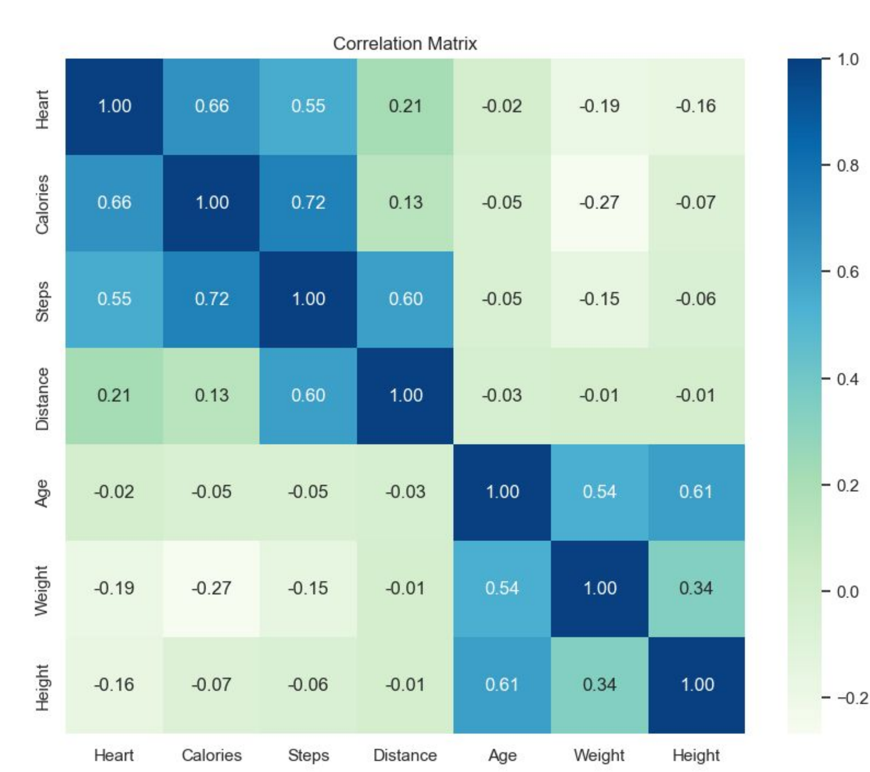
Overview:
The dataset contains recorded measurements from participants using the Apple Watch and Fitbit devices. The data is captured at 1Hz across several health metrics. The dataset reflects observations from two groups, with Apple Watch recording 57,097 minutes and Fitbit recording 21,489 minutes of data.

The attributes of the data:
1. **Heart Rate (bpm)**: Measured per minute; data interpolated using linear methods to address gaps.
2. **Calories (kcal)**: Energy expended as recorded by each device.
3. **Steps (count)**: Total steps counted during the recording period.
4. **Distance (km)**: Total distance covered as measured by the wearables.
5. **Age (numerical)**: Age of participants.
6. **Gender (M/F)**: Male or Female.
7. **Weight (lb)**: Body weight in pounds.
8. **Height (ft)**: Height in feet.
9. **Activity Labels (Sleep/Sedentary/Light/Moderate/Vigorous)**: Activity categories based on intensity, labeled using GENEActiv device data.

Data Cleaning / Preparation:

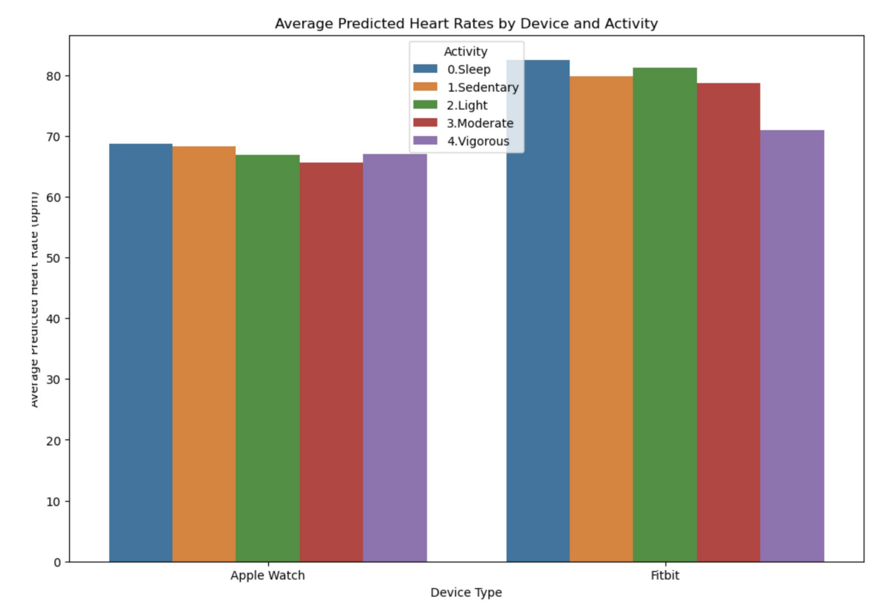
Fitbit.csv + AppleWatch.csv → Devices.csv (78586 rows, 10 columns)

Exploratory Data Analysis:

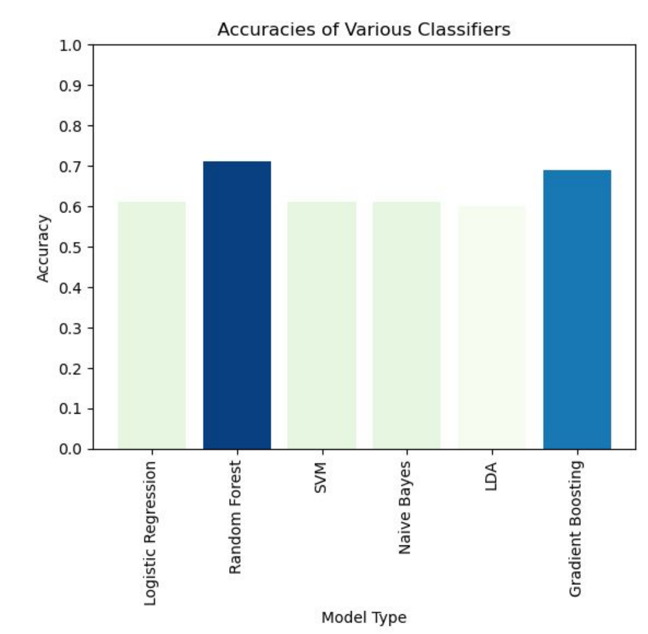


Analysis and Methods: Methodology

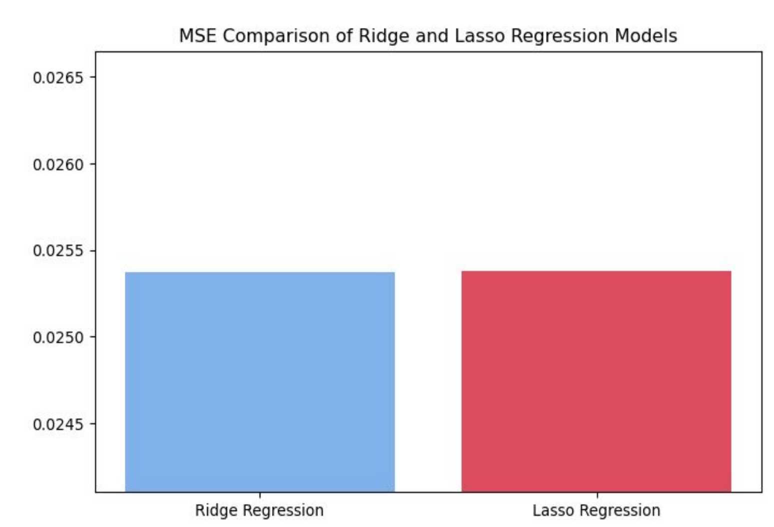
Ordinary Least Squares Regression
- Analyzing influence of device type, activity levels, and step count on heart rate measurements
- **Data Encoding**: 'Activity' and 'Device' were encoded to prepare for the analysis
- **Resampling Methods**: Addressing imbalance in 'Activity' by reducing bias on majority
- **Interaction Feature:** Between 'Device_encoded' and 'Activity_encoded' to capture combined effects of both features on heart rate
- **K-fold Cross Validation:** Evaluated the model's performance and ability to generalize to the test set (20% of dataset)
- Predictors in Model: 'Steps', 'Activity_encoded', 'Device_encoded', and 'Interaction' (using StatsModels)



Multi-Class Classification
- Analyze the different type of users based on the numeric metrics of the data as well as encodings of the categorical metrics.
- Goal was to predict age groups (20-30,30-40,40+) of the user.
- **Classification Techniques:**
  - Logistic regression, SVM, Naive Bayes, LDA
  - Random Forest, Gradient Boosting
- Each classifier was ran on the same random state of a train test split.
- Resampling to address class imbalance.
- Compared classification reports and confusion matrices between the models.



Ridge Regression and Lasso Regression
- **Advanced Regression Techniques:** Used Ridge and Lasso models to evaluate how personal traits and device type influence calorie burn per step.
- **Data Preprocessing:** Applied StandardScaler and OneHotEncoder to standardize and encode data, introducing 'Calories per Step' for a targeted analysis of activity efficiency.
- **Model Insights and Validation:** Employed K-fold cross-validation for robust testing, revealing that age and heart rate enhance efficiency, while higher weight and Fitbit usage decrease it.

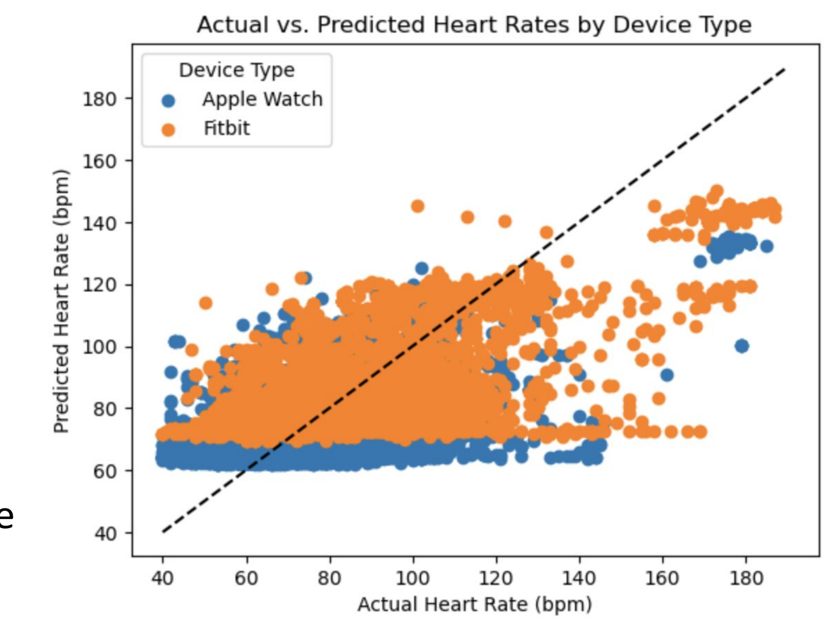


Results

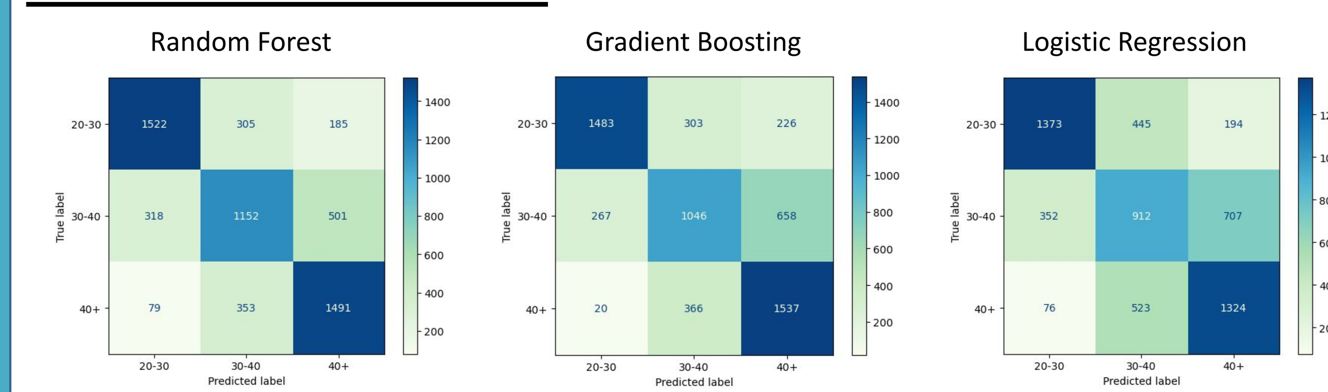
Ordinary Least Squares Regression

Model Summary:
- R-squared: 0.372
- Coefficient of Interaction Term: -0.1682
- p-value of Interaction Term: 0.340
- p-value < 0.05 for other predictors

Cross Validation:
- Mean RMSE: 14.6332
- Standard Deviation of RMSE: 0.104
- Consistency in the model's performance across different samples



Multi-Class Classification



- The ensemble models of Random Forest and Gradient Boosting had the best overall performance.
- Other models had difficulty separating the classes due to the close nature of the data and class imbalance.

Ridge Regression and Lasso Regression

Ridge Regression MSE: 0.025371075916977296
Lasso Regression MSE: 0.02537662074589587

These small values suggest that both models perform similarly and adequately in predicting the efficiency of physical activity (calories burned per step), with a low average squared error across predictions.

Conclusions

This data exploration, methodology, and analysis has given us deeper insight into the types of data collected by wearable fitness technologies and their applications.
- Steps taken and the type of activity are significant predictors of heart rate according to the model, while the interaction between device type and activity does not significantly affect heart rate predictions. Apple Watch and Fitbit show similar capabilities in predicting heart rate during different activities, but they may not vary in prediction quality.
- We analyzed classifiers of varying levels of strength, but found decent success in predicting the age of a user based on their reported metrics. Issues that prevented stronger models included class imbalance and the granularity of the data.
- Both Ridge and Lasso regression models performed effectively in predicting activity efficiency, with low mean squared errors indicating high accuracy. These models have also helped in understanding how variations in device types and individual characteristics impact overall fitness outcomes.