Fitbit Data Analysis

Abhinav Gunti
B. Tech CSE
PES University, RR Campus
Bangalore, India
gunti.abhinav@gmail.com

Abhishek Singhi
B. Tech CSE
PES University, RR Campus
Bangalore, India
abhisinghi2002@gmail.com

Aniket Acharya
B. Tech CSE
PES University, RR Campus
Bangalore, India
aniket.acharya2001@gmail.com

978-1-7281-4384-2/20/$31.00 ©2020 IEEE

Abstract — **The Fitbit is a wearable wireless device used to monitor a user's daily activity and health stats such as sleep quality, steps walked, heart rate etc. This data provided by the device is used to predict overall fitness of the user along with finding out how day-to-day activities of the user are being affected such as – whether amount of activity affects user heart rate during a given time of day or how much BMI (Body Mass Index) affects calories burnt. The data is extracted from a dataset generated from Fitbit users by a distributed survey via Amazon Mechanical Turk between 03.12.2016-05.12.2016.**

Keywords—Fitness, Fitbit, Activity Tracking, Health, Data Extraction, Wearable Tracker, Health Statistics.

I. INTRODUCTION

Fitbits have become very sought after worldwide since the general public has become more aware regarding tracking their health. Similarly, they are also popular amongst fitness experts, physiologists and data scientists for the purpose of data collection. The Fitbit has been designed to provide as intricate of details of the user down to a given minute, making outliers in data easy to view and finding patterns within the given data.

The reason the Fitbit is in such demand is due to its easy-to-use interface (regarding the general public) and the vast insights that can be taken from a user's Fitbit data history (regarding academic experts). The study done below is for providing the user analysis of their daily stats along with predictions regarding health conditions and sleep quality.

II. DATASET

The Dataset used is Fitbit Data generated by respondents to a distributed survey via Amazon Mechanical Turk between 03.12.2016-05.12.2016. Thirty eligible Fitbit users consented to the submission of personal tracker data, including minute-level output for physical activity, heart rate, and sleep monitoring. The users are represented by unique Id's.

There is tracking of the users' distance travelled and time of travel at different activity levels. Activity levels are divided into – sedentary, lightly active, fairly active and very active. The calorie expenditure and step count for the total distance is also tracked. Calorie expenditure is also tracked on an hourly basis.

The sleep is tracked in terms of time slept in REM (Rapid Eye Movement) sleep, light sleep, awake in bed etc. MET (Metabolic Equivalent of Task) is also found, which gives an indication of the energy needed to complete a task in the form of a ratio. The weight and BMI of the user was also logged on a daily basis. However, it was observed that most of the users weren't logging the data regularly, therefore the weight data is sparse.

Data is stored in multiple CSV files, according to the frequency and type of the collected information.

III. DATA ANALYSIS

In the project, the aim was to inform the users about their activity stats and how their activity affects other health stats. Firstly, we determined how many hours users wore the device for and how any of those hours involved any activity.

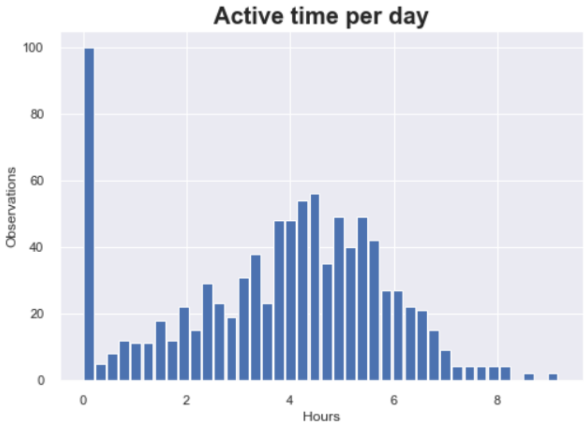

Figure 1: Hours per day spent in activity

Following this, we analyzed the days and hours of the week during which the user was most active based on their step count crossing a threshold for that given time period.

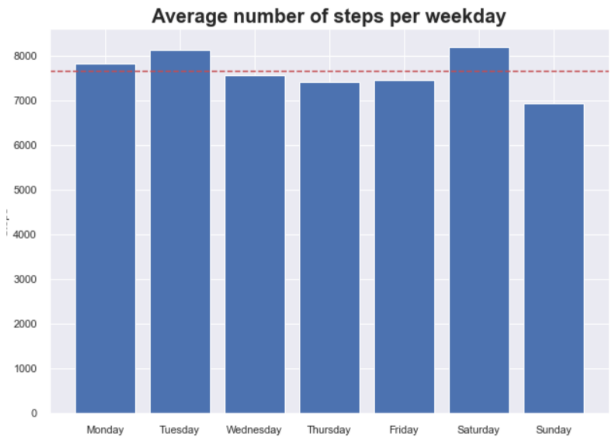

Figure 2: Activity Measure per day

Having found the days when the user is most active, we checked if the activity of that user during the day affected the heart rate at night along with how much time they spent in bed lying awake. We calculated daily activity based on the step count, calories burned and total activity time. [3] Our work indicates that the number of steps reported by different devices worn simultaneously could vary as much as 26%. At the same time, the variations seen in distance travelled, based on the step count, followed the same trends.

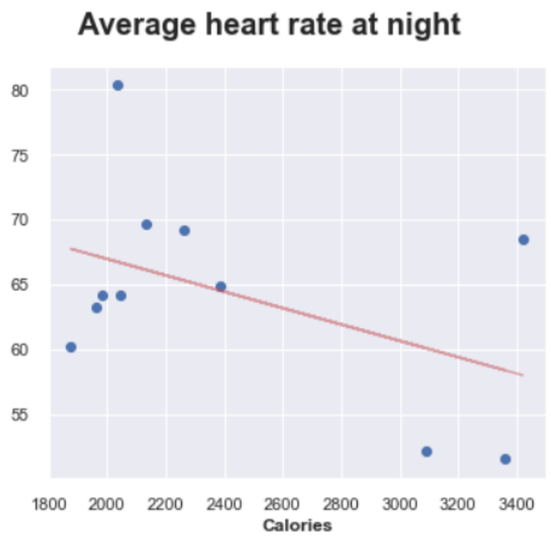

Figure 3: Night Heart Rate vs Calories Burned

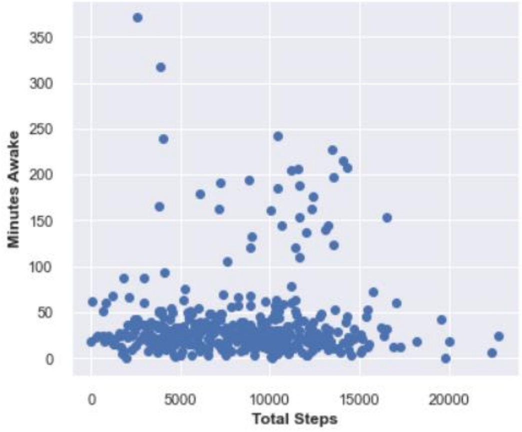

Figure 4: Minutes awake in bed vs Total Step Count

For correlation, it is determined between BMI of the user and the calories burned by them for a particular activity. We do this simply by logging the BMI of the user and checking calories expended per 100 steps.

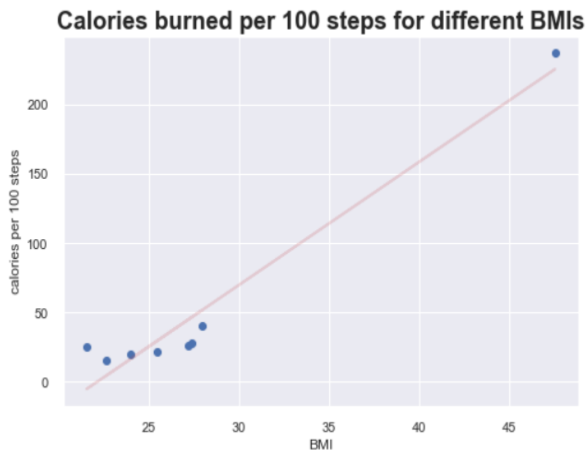

Figure 5: Calories burned vs BMI

In the next sections we have done predictions. Firstly, we are classifying cardiovascular disease based on attributes that can be acquired from the Fitbit such as – max Heart Rate and irregular heart rhythm. This data along with stats of the user such as cholesterol, resting blood pressure and fasting Blood Sugar is used to determine whether a user has a heart disease or not, using a KNN (K-nearest neighbors) model.

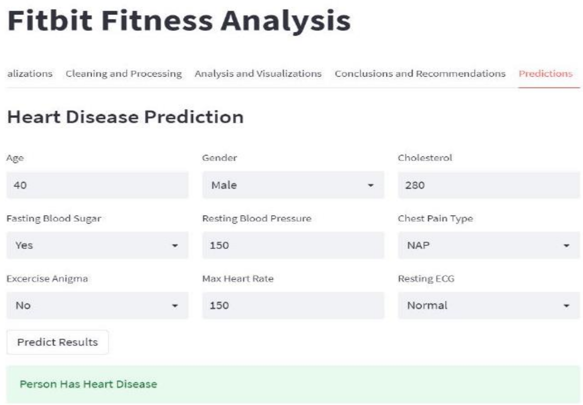

Figure 6: UI for heart disease prediction

[1] Through an interview study with 14 Fitbit users we identified three main sources of errors: (1) lack of definition of sleep metrics, (2) limitations in underlying data collection and processing mechanisms, and (3) lack of rigor in tracking approach. Therefore, we are using Multiple Linear Regression for determining a users' sleep score. This score gives information about the sleep quality of a user in the form a real number. It is calculated using attributes including- minutes of REM sleep, minutes of light sleep, minutes in bed etc. All this data is obtained from the Fitbit.

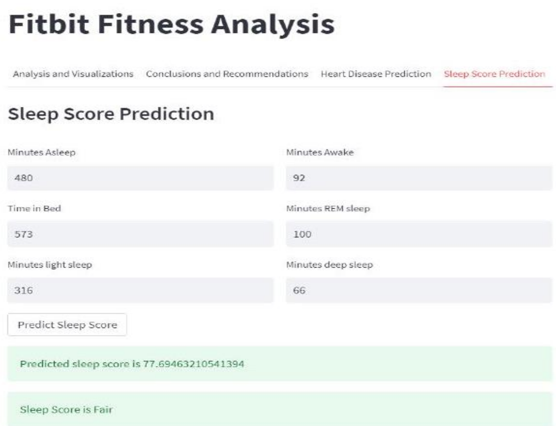

Figure 7: UI for sleep score prediction

IV. OBSERVATIONS

From visualizations, we can see that distance travelled at different activity levels follows a normal distribution. The calories burned by a user is directly, positively related to the distance covered by them. Similarly, heart rate has a strong correlation with time of user being active in the sedentary state.

The number of unique participants for 'weight' data is very limited. To draw reliable conclusions, deeper analysis is needed.

The users can be divided into four quarters for determining risk of heart disease can be divided on the basis of age and resting blood pressure. The division can also be done based on cholesterol level.

The sleep score can also have 'No of user Awakenings' as an attribute. However, since the data collected isn't for this isn't as accurate as the sleep time, it is dropped. [2] Exercise training may be more successful in subjects with existing sleep disturbances to improve sleep characteristics rather than in healthy older subjects.

V. CONCLUSION

The following conclusions have been made from the analysis, based on which various recommendations can be made to the user to improve their fitness and overall lifestyle:

1. Average users' activity time is 3.8 hours and wear time is 20.4 hours.

2. The users tend to be most active on Tuesday and Saturday and least active on Sunday.

3. Most of the users' steps are taken between 8 a.m. and 8 p.m. with a dip at 3pm, mostly due to it being lunch time.

4. There is negative correlation between night heart rate and number of calories burned. However, there is no correlation between night heart rate and number of average steps or active time.

5. The users spend mostly 2 hours awake in bed and there is almost no correlation between their activity level and time spent awake in bed.

6. Users with high BMI tend to burn more calories for the same activity level.

Regarding future scope of this project, we intend to implement a model to calculate calorie expenditure for different activities for a user for a given amount of time. This can be used to give the user info on efficient methods of calorie burning if they are looking to burn bodyfat via exercise. [4] Consistent evidence indicated that Fitbit devices were unlikely to provide accurate measures for energy expenditure in any testing condition. We are also looking into implementing a weight calculator which determines weight fluctuations that may occur in a user, having calculated the total calorie expenditure on a given day.

This info can be useful for users who are particular on tracking their daily calories and those looking to make healthy changes in their lifestyle.

REFERENCES

[1] Zilu Liang, Bernd Ploederer, Mario Alberto Chapa -Martell, PervasiveHealth '17: **Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare** https://doi.org/10.1145/3154862.3154897

[2] Oudegeest-Sander M.H.a, b · Eijsvogels T.H.M.a · Verheggen R.J.H.M.a · Poelkens F.a · Hopman M.T.E.a · Jones H.c · Thijssen D.H.J.a, c , ' **Impact of Physical Fitness and Daily Energy Expenditure on Sleep Efficiency in Young and Older Humans**'

[3] Chelsea G. Bender, Jason C. Hoffstot, Brian T.Combs, Sara Hooshangi, Justin Cappos, '**Measuring the fitness of fitness trackers** '

Published in: 2017 IEEE Sensors Applications Symposium (SAS)

[4] Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, Hamilton CB, Li LC

Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data

JMIR Mhealth Uhealth 2018;6(8):e10527

doi: 10.2196/10527

PMID: 30093371

PMCID: 61077