International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 5, ISSN No. 2455-2143, Pages 286-287
Published Online September 2021 in IJEAST (http://www.ijeast.com)

FRAMEWORK FOR TRACKING FITNESS USING DATA ANALYTICS

Suma Shruthika M
PSG College of Technology
Coimbatore, Tamilnadu, India

Abstract— **Today, all the devices around are built with the capacity to produce and store data in its most relevant form. Unless and until interesting insights and data which are meaningful can be extracted, the data stored will be of no use. Fitness tracking smart watches like fitbit track and store data like the number of steps walked, the quality of sleep, heart rate (beats per minute), the number of calories burned in a particular day and many other various activities. From this health tracking system, the fitbit application gathers a huge amount of data and allows us to analyze the fitness data collected by the application. By employing techniques like data exploration, modelling, deploying and integrating, we will be able to arrive with very useful insights.**

Keywords— **Fitness Tracking, Data Exploration, Deploying-Integrating**

I. INTRODUCTION

Wearable devices or in other words, fitness trackers are compact and are worn by users through the entire day for collecting data and storing them on the respective devices. However, since these devices have very less memory, they should be regularly synced with another device to offload the data that is collected. Most of these devices pair a smartphone as a syncing device and both these devices exhibit data exchanging by employing wireless communication protocols like Bluetooth. Once the data is sent to a paired device like the smartphone, the data is uploaded via an Internet connection onto a Cloud service in order to process it further for the purpose of tracking, and visualization. The mobile device also visualizes the physiological data to the user. The collected raw physiological data from wearable smart watches is exchanged between the wearable device and the smart phone which can contain personal information about the users which includes information about health. Manufacturers of these wearable fitness trackers are always concerned with energy consumption and often overrule the security protocols.

II. PROPOSED ALGORITHM

A. Data Exploration –

Accessing and downloading the data from the fitbit application is pretty simple. We can easily access and download the data from the website's dashboard in many formats including the csv format. We can select the data that we specifically want to deal with and then export the data set. Data exploration being the key aspect of data analytics comprises the process of data cleaning and preprocessing.
The training data set will incorporate of a list of records which will be each entered on separate lines in the csv file. Every record will itself contain the physiological data generated by the wearable device for an activity performed along with the values and labels attributed to the user who recorded the activity.

Date	Activity	Steps	Distance	Duration	Calories
Today, 06:43	Workout	2,245	N/A	57:54	337 cals
Sep 6, 07:07	Workout	1,651	N/A	1:02:14	343 cals
Sep 4, 07:10	Run	3,426	2.79 kilometers	22:23	256 cals
Sep 3, 07:14	Workout	2,010	N/A	1:23:34	413 cals
Sep 2, 06:57	Workout	3,194	N/A	1:34:54	540 cals
Aug 22, 19:55	Walk	3,014	N/A	34:59	184 cals
Aug 20, 19:14	Walk	2,196	N/A	25:36	134 cals
Aug 16, 07:12	Workout	2,609	N/A	1:11:40	367 cals
Aug 15, 07:31	Run	3,376	2.16 kilometers	6:14:37	662 cals
Aug 14, 07:02	Workout	2,936	N/A	1:02:24	410 cals

Fig.1. Activity History of Fitbit Application

B. Preparing the Data –

Since the data collected can involve garbage, inconsistent and irregular values, it can lead to many difficulties. In order to be accurate in the analysis and make the right predictions, it is essential that we clean the garbage data through the process of data cleaning. Data cleaning is the process of removing incorrectly formatted, duplicate, or incomplete data existing within a dataset. Through such a process, data is validated.
as the see value is available at the receiver end to the authorized users.

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 5, ISSN No. 2455-2143, Pages 286-287
Published Online September 2021 in IJEAST (http://www.ijeast.com)

Fig. 1. Data Cleaning

C. Plotting the Data –

The main purpose to plot scientific data is for the purpose of visualizing the variations or to show relationships that exist between variables, but not every data set requires representation in the form of a plot. In case of existence of only one or two points, examining the numbers in a direct manner is very easy, and very little or almost nothing can be obtained from plotting them as graphs. In the same way, if there are no variations in the data set, it is very easy for us to observe or assert facts without employing a graph of any form. By using the plotly library, we can create plots which are interactive with very minimal effort. The relation between the number of workouts performed on a monthly basis is plotted below.

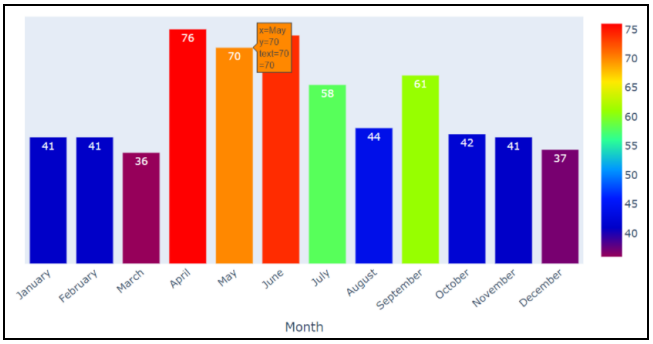

Fig. 2. Count of Workout in Each Month

For illustrating numerical proportions in data, we make use of a type of data visualization called pie chart. 'matplotlib' is a python library that provides various useful tools for plotting insightful visualizations, like pie charts.

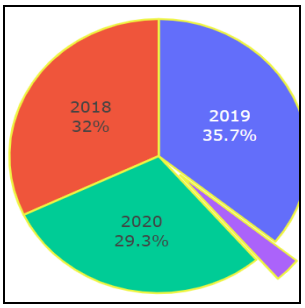

Fig. 3. Year's Contribution to Total Workouts

III. EXPERIMENT AND RESULT

There could be many possible visualizations from the dataset available but this research paper focuses on the bar chart and the pie chart. It has been further demonstrated how the data sets can possibly be obtained and analyzed by employing machine learning algorithms for identifying individuals and tracking their corresponding behavior. The futuristic step in this research will be to incorporate a wider range of individual persons and to differ the fitness activities for generating a bigger training data set in order to gain more insights.

IV. REFERENCE

[1] Neter, Wasserman, and Whitmore (1993). Applied Statistics, 4th Edition, Allyn and Bacon, Boston, MA.
[2] E. Bauer and R. Kohavi, An empirical comparison of voting classification algorithms: Bagging, boosting, and variants, Machine Learning 36(1–2) (1999), 105–139.
[3] S.A. Khan, E. Leppäaho and S. Kaski, Bayesian multi-tensor factorization, Machine Learning 105(2) (2016), 233–253.
[4] Trevor Hastie, Robert Tibshirani, and Jerome Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 58(1-2) (2013), 110-130.
[5] Britt Cyr, Webb Horn, Daniela Miao, and Michael Specter. Security analysis of wearable fitness devices (fitbit). Massachusets Institute of Technology, page 1, 2014.
[6] R. Bouhenguel, I. Mahgoub, and M. Ilyas. Bluetooth security in wearable computing applications. In 2008 International Symposium on High Capacity Optical Networks and Enabling Technologies, Penang, pages 182–186, 2008.
[7] G. Press. Internet of things (iot) predictions predictions from forrester-machina-research wef, gartner idc. Forbes Magazine, January 2016.