Are you doing your exercise regularly? Monitoring and Assessing the Gym Activity

Masnoon Nafees

Department of Information Systems

University of Maryland, Baltimore County

masnoon1@umbc.edu

ABSTRACT Smart Watch and electronics devices are becoming very common for people now a days. In market, so many options are available like Microsoft Band, Apple Watch, Galaxy Watch etc. These devices help us to monitor our daily activities and keep track of our health. In this project, we try to build a machine learning model where we can understand our daily Gym activities like Weightlifting, Running, Cycling etc. The main motivation is to make a novel idea on sensor data. Most of us just go to the Gym and do some random activities hence ignoring the pattern we follow in Gym every day. A machine learning model will give us more inner understanding of our pattern, monitor and analyze data. For this we used Microsoft Band and collect data for 3 days. Rest of the paper, we will try to state about the problem statement of our project, motivation, device, dataset, methodology, result and observation.

KEYWORDS

Gym, Sensor Data, MS Band, Sensor Data, Multiclass Classification

Motivation Most of us don't even realize we are doing the Gym activities properly or not. We just go into Gym and do some random work. But we have to understand our pattern of our and also keep in mind about if we missed any particular activities or not or doing one exercise more or less than our regular activities. So, in this project we decided to monitor the activities of a regular Gym going person where he will perform some of the task wearing the Microsoft Band. After collecting data, we need to process the data and cleaned it. For ground truth, we take the timestamp notes on every activity starting and ending time. Then we can label the data according to that. This project is very interesting and has some novelty especially counting sensor data on Gym activities.

Problem Most of the wearable devices doesn't give us the raw data. The problem we are trying to solve needs raw sensor data to build machine learning model. Already processed data cannot be applied into a machine learning model. In that case we chose Microsoft Band for data collection. It's mainly because MS Band can store raw sensor data into CSV format. After processed the data, we applied several classification models like KNN, LDA, Naïve Bayes, Decision Tree, PCA etc to observe classification accuracy. Later, we present our observations.

Data collection: For data collection we use MS band. One of our colleagues helped us to collect the data. To store the data into CSV format we had to use a third-party app name "Companion App for Android". This app connects the MS Band through Bluetooth. After connection all of the sensor data can be accessed from the user interface. There is an option for sampling frequency for accelerometer, gyroscope. We choose 128 sampling frequency for each sensor. That means, each sensor will produce data of 128 per second. After collecting the data for 3 days, there are 42 columns and thousands of rows of the data set. The next part is to identify important attribute and label the data.

Data Preprocessing Data preprocessing part is the most challenging one. There could be some missing values needed to be dealt with before going for calculation. We just cannot take all the attribute for calculation. Here is the list of attributes we consider into our model.

- Accelerometer (3 axis)
- Gyroscope (3 axis)
- Air Pressure
- Air Temperature
- Brightness
- RR Intervals

After that we have to Label the data according to the ground truth, we collected through notes in Gym activities. For labelling the data we use numerical values. We took three activities like Running, Weightlifting and Cycling.

- IDLE – 0
- Running - 1
- Weightlifting – 2
- Cycling – 3

While the user was idle between exercises, we put 0 into the timestamp and when was doing activities for that particular timeframe we label the data with numerical value. In this way, computer can learn the pattern and later can be applied for classification and anomaly analysis. We have to manually label the data by going through all the rows. Here is the complete flow of the methodology.

1. Collect data for 3-4 days
2. Record activities for Running, Cycling, Weightlifting and Idle time
3. Label the data and replace the label with numeric value
4. Train the mode
5. Test the model (25 percent of Dataset)

For test the model we use train test split to split the dataset for 25 percent of Test data from entire dataset.

Result: For experimental purposes, at first, we use decision tree. Decision has several attributes like tree size, max heap, max nodes etc.

Here is the result of the accuracy for decision tree. It seems if we decrease tree size the accuracy of the tree percentages for both train and test got increase.

Tree	Max Depth	Train	Test
100	3	80	78
50	5	95	93
Default	2	75	71

Figure 1: Decision Tree

But if we keep the tree default and decrease the max depth the accuracy got decreased.

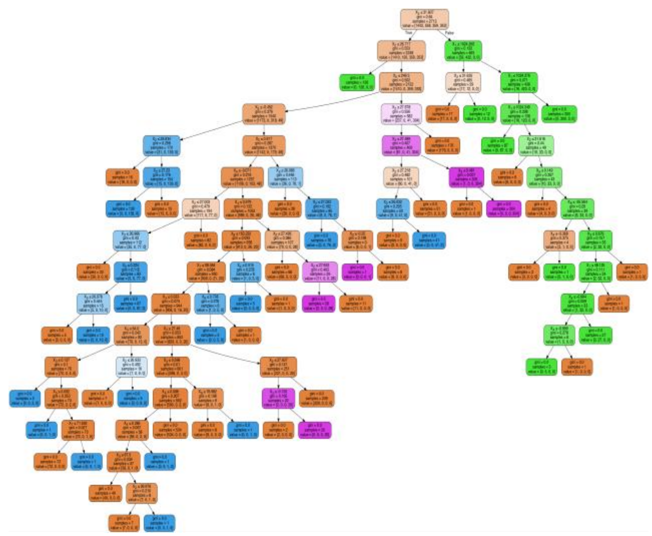

Figure 2: Decision Tree Visualization

From this figure we can see the entire decision tree has visualized by lowest Gini Entropy. The decision tree put Air Temperature is the root node for the tree.

The next experiment we used Naïve Bayes and Linear Discriminant Analysis. From the result we can see that Naïve Bayes and LDA has similar accuracy percentages for both train and test.

Prediction	Train	Test
Naïve Bias	77	81
LDA	75	77

Figure 3: Naïve Bayes and LDA

Next experiment is the K-nearest Neighbor Algorithms. KNN has parameters too like increase and decrease neighbor size.

Neigbour	Train	Test
7	78	74
9	77	73
11	76	73
13	75	73

Figure 4: KNN with Neighbors Size

From the table we can that if we increase the size of the neighbors, the accuracy got decreased. Even though the size of the neighbor got increased from 7 to 13, the accuracy got slightly decreased by 78 to 75 for train and 74 to 73 for test percenatges.

We also used the Principal Component Analysis to calculate the error.

Component	MSE Error Test
5	1.32
7	1.4
9	1.41

Figure 5: PCA for MSE Error

PCA for MSE showed that if we increase the size of the component, the test error got decreased by slightly 1.32 to 1.4 for increasing the component from 5 to 7.

The next part is we compute the pair wise plot for each of the attribute. Unfortunately, the pair is so details and small for visualizing in a paper to understand.

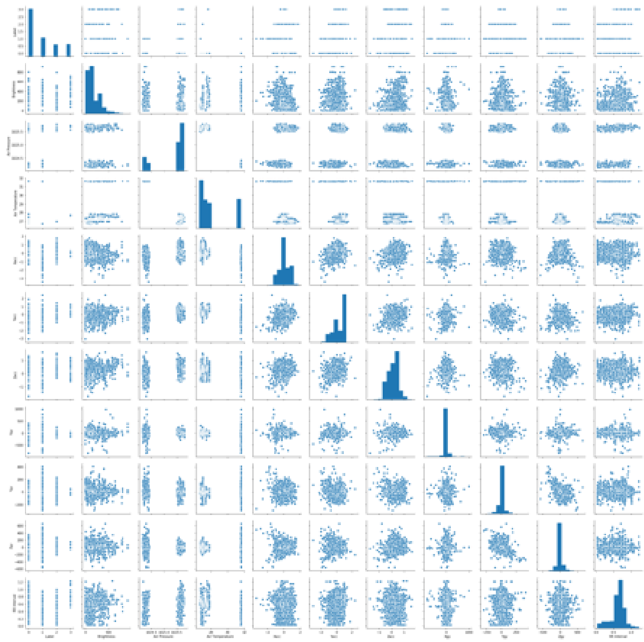

Figure 5: Pair Wise Plot

From the pair wise plot, we can see that all of the cluster and bar diagram of 11*11 pair of the attributes. The letters are small but if we zoom in we can get the understanding of any pair and how it is distributed with each other.

Later we analyze the feature importance based on decision tree classifiers.

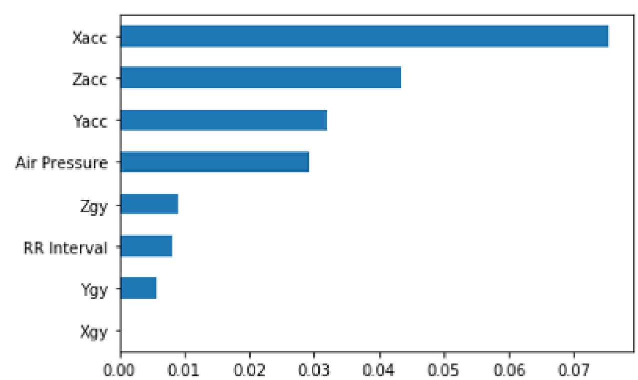

Figure 6: Feature Importance by Decision Tree

Decision tree classifier put Xacc (X-axis value for Accelerometer) attribute is the most feature importance for the model. Later followed by Z and Y axis value. Then we got Air Pressure, RR intervals and X, Y, Z axis value for Gyroscope.

After decision tree we used random forest classifier for feature importance.

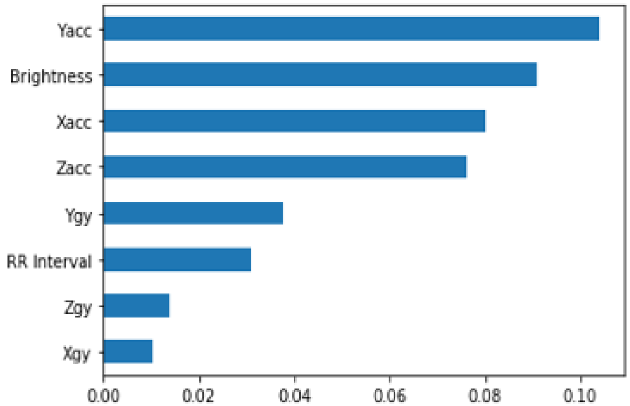

Figure 7: Feature Importance by Random Forest

For random forest, we found that Y-axis accelerometer has the most importance feature among the model. Interestingly we got Brightness in here and Air pressure in decision tree. These two are less importance for collecting sensor data. After observing both figure we can conclude that Accelerometer has the best feature importance for the model. The movement of the sensor hence body part has highly connected with Accelerometer and Gyroscope.

Lastly, we computed the precision and recall for the model. The precision and recall show that the model has very good accuracy.

	precision	recall	f1-score	support
0	0.988962	0.986784	0.987872	454
1	0.990196	0.990196	0.990196	204
2	0.964912	0.973451	0.969163	113
3	1.000000	1.000000	1.000000	134
accuracy			0.987845	905
macro avg	0.986018	0.987608	0.986808	905
weighted avg	0.987872	0.987845	0.987856	905

Figure 8: Precision, recall and F1 score

From the table we can see that IDLE (0) has precision 0.988, recall 0.986 and f1 score 0.987. Running (1) has precision 0.990, recall 0.990 and f1 score 0.990. weightlifting (2) has precision 0.96, recall 0.97 and f1 0.96 and lastly, Cycling has 100 percent for all three of them.

Observation: After all the experiment, we have few observations. Most important is all the classification has given good result and there is not much difference between them which suggest dataset we produce has very little inconsistency. Decreasing tree size, increasing max depth improve accuracy in Decision tree. And also Increasing neighbor leads to decrease accuracy in KNN. Another important observation is increasing component leads to increase MSE error in PCA. For testing, if we exclude any particular activities or increase or decrease the time frame the accuracy of the model decreases. While building the decision tree Air pressure got the root node which shouldn't have happened. It's mainly because Air pressure, temperature and brightness has fluctuated so much during the experiment.

Challenges: While completing the project, we have faced several challenges. The main problem was the connection issues between the MS band and the android application. We had to reset both the device several times to get connected. Some of the MS Band didn't even get connected hence we had to change the device. The most important issue was data loss in CSV file. While we are collecting data, it seems okay at that time but when we started to label the data, we realize there are some seconds got missing. It was very challenging for me to get the dataset going. We tried for the best and predicted some of the label according to their previous data distribution. As the dataset has thousands of rows, it was a hefty task to label the data manually. Data seems overfitted at some time of the experiment, so we had to find the best parameter to prevent the overfitting issues.

Future Work: In future we will try to overcome these challenges. Most importantly we don't need to use Air pressure, temperature and brightness. The value of this could be misleading and could imbalance the dataset. Also, instead of splitting the data into test and train, we could use separate dataset for testing the model.

Conclusion: In conclusion, the project was really interesting. I didn't have any previous experiment with anything about sensor data. Also, preparing dataset is the new experiment for me. I have learned many interest parts of machine learning area and surely it will help me in near future.

Acknowledgement: I would like to thank *Dr Nirmalya Roy* for helping me with device and necessary guideline. I am also grateful to *Mr Forhan Bin Emdad* for doing Gym activities for me.

Reference:

1. Cranney, L., Phongsavan, P., Kariuki, M., Stride, V., Scott, A., Hua, M., & Bauman, A. (2016). Impact of an outdoor gym on park users' physical activity: A natural experiment. *Health & place*, *37*, 26-34.
2. Al-Ali, Ammar. "Pulse oximetry sensor including stored sensor data." U.S. Patent No. 7,272,425. 18 Sep. 2007.
3. Jonsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE transactions on Geoscience and Remote Sensing*, *40*(8), 1824-1832.
4. Aggarwal, C. C. (Ed.). (2013). *Managing and mining sensor data*. Springer Science & Business Media.
5. Gruenwald, L., Chok, H., & Aboukhamis, M. (2007, October). Using data mining to estimate missing sensor data. In *Seventh IEEE International Conference on Data Mining Workshops (ICDMW 2007)* (pp. 207-212). IEEE.