Human Activity Recognition Using Smartphones

Vidhan Dholakia

Artificial Intelligence

Department of Electrical and Computer Engineering

Queens University

18vbd@queensu.ca

12/13/2020

ABSTRACT – Human Activity Recognition (HAR) is a person's activity that uses sensitive sensors that are influenced by human movement. These movements are often normal indoor activities such as standing, sitting, jumping, and going upstairs. Recognition can be done by taking advantage of the data collected from different sources, such as environmental or body-worn sensors. Both users and capabilities(sensors) of smartphones increase and users usually carry their smartphones with them. These facts make HAR more important and popular. The idea is that an intelligent computer system will then provide assistance once the subject's activity is recognized and known. Data obtained from the accelerometer and gyroscope sensors of smartphones are categorized in order to detect human activity. The results of the methods used are compared in terms of performance and accuracy.

KEYWORDS—*Machine learning, smartphone. Activity recognition, classification, gyroscope, accelerometer.*

I. INTRODUCTION

Human activity recognition is a significant but challenging research area with many applications in smart environments, home security, and healthcare. Smartphones are the most useful tools in our everyday lives and with advanced technologies, they are more capable of fulfilling consumer demands and desires day by day. [1] Designers add new modules and devices to the hardware to make these gadgets more usable and efficient. Sensors play a major role in increasing the functionality and knowledge of smartphones

Therefore, most smartphones have numerous embedded sensors in the environment and this makes it possible to gather large quantities of information about the everyday life and activities of the user. Among these instruments, we have accelerometer and gyroscope sensors.

The accelerometer has been standard hardware for almost all smartphone manufacturers. The accelerometer measures the change in speed; not the speed itself. Data retrieved from the accelerometer may be processed in order to detect sudden changes in movement. Another sensor in normal smartphone hardware is a gyroscope. A gyroscope that using gravity, measures orientation. Gyroscope-recovered signals can be analyzed to identify the location and orientation of the system. Since there is a major difference in characteristics between the knowledge. Many features could be produced from these sensor data to determine the behavior of the person carrying the device, obtained from these sensors. [5]

In this work, a dataset consists of signals from the accelerometer and gyroscope of a smartphone carried by different man and women volunteers while doing different activities are classified using different machine learning approaches. [10] The performance of different approaches is analyzed and compared in terms of precision and efficiency.

The paper is structured as follows. After discussing related work in Section 2 and explaining the description of experiments in Section 3, I have explained the methodology in Section 4 comprising of data collection, pre-processing, classification, and results. Finally, Section 5 concludes the paper, and individual tasks performed by me are specified in Section 6.

II. RELATED WORK

There have been various studies on the classification of smartphone-based on user behaviors. Bayat et al. [2] have examined the recognition of human behavior with accelerator signals. Attal et al. attempted to identify behavior according to various wearable devices i.e., Accelerometers and a gyroscope. In order to identify user movement using a mobile accelerometer and gyroscope, Ronao et al. [2] structured a coevolutionary artificial neural network. Kozina et al. [2] also experimented with the accelerometer on fall detection.

III. EXPERIMENT DESCRIPTIPON

The Human Activity Recognition database was created from the recordings of 30 study participants conducting daily living activities (ADL) while holding a waist-mounted smartphone with embedded inertial sensors the aim is to classify operations into one of the six tasks performed. The experiments were performed within an age range of 19-48 years with a group of 30 volunteers. Six activities(WALKING,WALKINGUPSTAIRS,WALKINGDOWNSTAIRS, SITTING, STANDING, LAYING) were carried out by each person wearing a smartphone (Samsung Galaxy S II) on their waist.

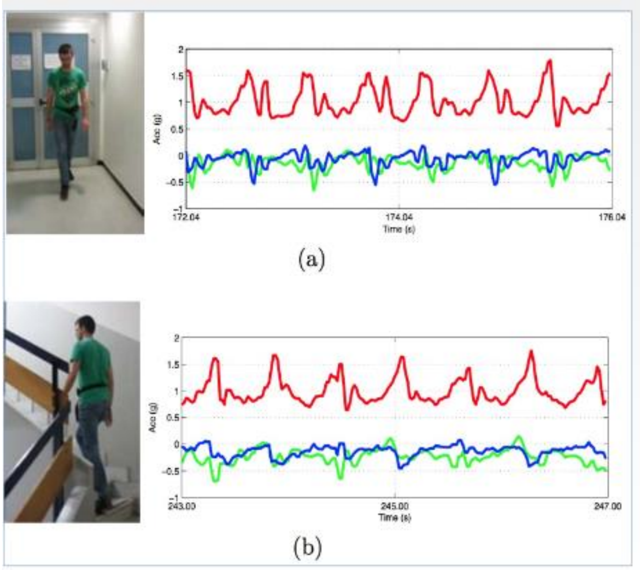

Figure 1. Human activity and analysis

At a constant rate of 50Hz, we captured 3-axial linear acceleration and 3-axial angular velocity using its integrated accelerometer and gyroscope. To mark the data manually, the experiments were video-recorded. The collected dataset was randomly divided into two groups, where 70% of the volunteers were chosen for the training data generation and 30% for the test data. By adding noise filters, the sensor signals (accelerometer and gyroscope) were pre-processed and then sampled in fixed-width sliding windows of 2.56 sec and 50% overlap (128 readings/window). [11]

Using a Butterworth low-pass filter, the sensor acceleration signal, which has gravitational and body motion elements, was split into body acceleration and gravity. It is believed that the gravitational force has only low-frequency components, so a filter with a cutoff frequency of 0.3 Hz was used. A vector of features was obtained from each window by measuring variables from the domain of time and frequency. [11]

Attribute information

For each record in the dataset the following is provided:

- Triaxial acceleration from the accelerometer (total acceleration) and the estimated body acceleration.
- Triaxial Angular velocity from the gyroscope.

- A 561-feature vector with time and frequency domain variables.
- Its activity labels.
- An identifier of the subject who carried out the experiment.

IV. METHODOLOGY

In this section, we have provided step by step implementation of the activity recognition.

A. Importing necessary files and data loading.

First, we import the necessary libraries as follows:

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

Next, we load the data using the following command:

pd.read_csv("r"C:\Users\vidha\OneDrive\Desktop\DATASET\train.csv"")

pd.read_csv(r"C:\Users\vidha\OneDrive\Desktop\DATASET\test.csv"

B. Data Pre-processing

B.(a) Checking for Duplicates.

print('Number of duplicates in train : ',sum(train.duplicated()))

print('Number of duplicates in test : ', sum(test.duplicated()))

```
Number of duplicates in train :  0
Number of duplicates in test :  0
```

Figure 2. Data Pre-processing

B(b) Checking for missing values

print('Total number of missing values in train : ', train.isna().values.sum())

print('Total number of missing values in train : ', test.isna().values.sum())

```
Total number of missing values in train :  0
Total number of missing values in train :  0
```

Figure 3. Data Pre-processing

B(c)Checking for class imbalance

plt.figure(figsize=(10,8))

plt.title('Barplot of Activity')

sns.countplot(train.Activity)

plt.xticks(rotation=90)

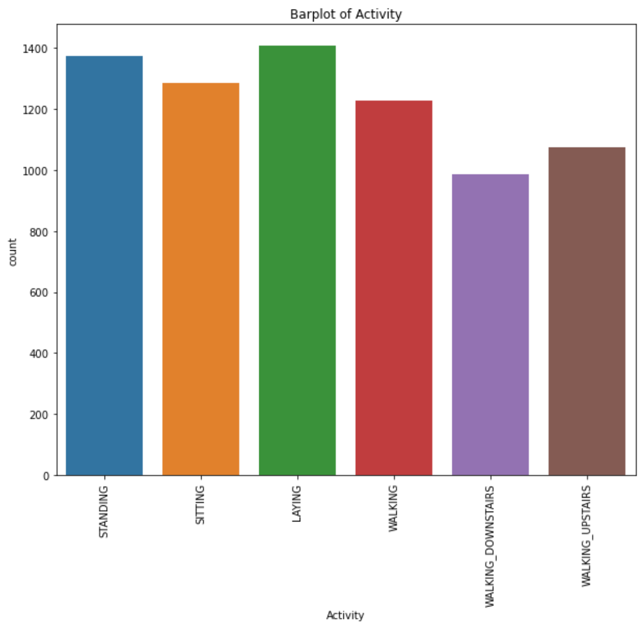

Figure 4. Bar plot of Activity

In all the six operations, there are almost the same number of findings, so this knowledge does not have a class imbalance problem.

C. Explanatory Data Analysis.

We may roughly position the activities in two groups based on the general nature of the activities.

a. Static Activities.

SITTING, STANDING, LAYING can be considered as static activities with no motion involved.

b. Dynamic Activities.

WALKING,WALKING_DOWNSTAIRS,WALKING_UPSTAIRS can be considered as dynamic activities with significant amount of motion involved.

We consider **tBodyAccMag-mean ()** feature to differentiate among these two broader sets of activities.

C(a) Analysing tBodyAccMag-mean feature

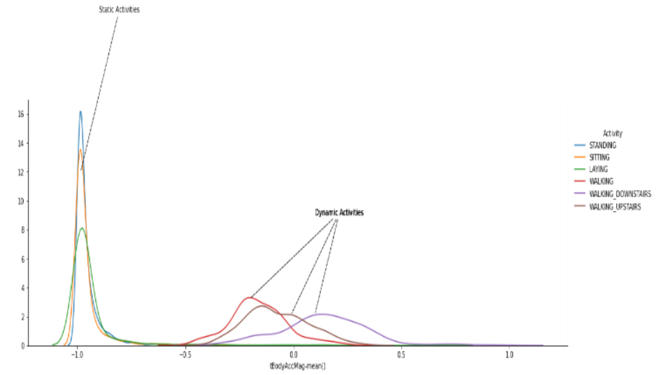

Figure 5. Static and Dynamic activities

We can easily come up with a condition to distinguish static activities from dynamic activities using the above density map.

Let's take a closer look at the PDFs under static and dynamic categorization for each activity.

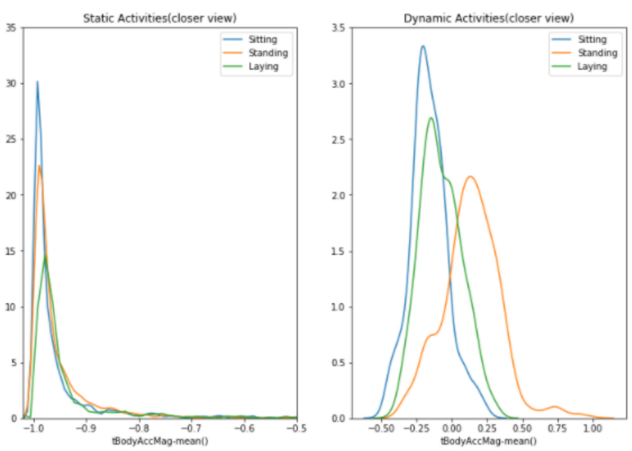

Figure 6. Analysis of activities

Box plots may also be used to depict the insights obtained through density plots. Let's plot the mean(tBodyAccMag-mean()) of Body Acceleration Magnitude boxplot through all six groups. [6]

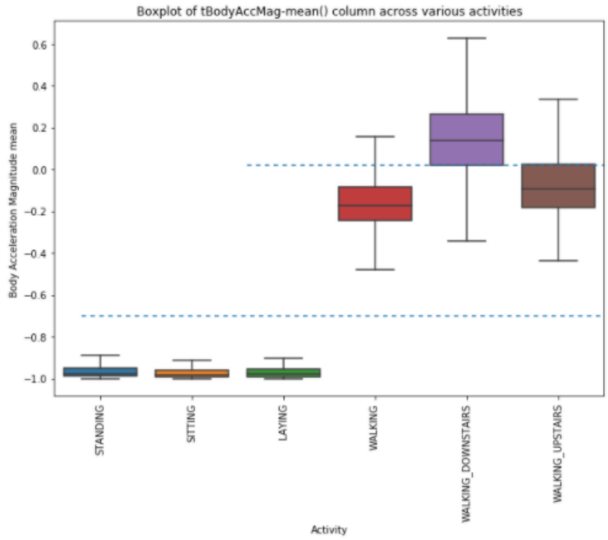

Figure 7. Box plot of tBodyAccMag-mean() column

We may come up with conditions to separate static operations from dynamic operations using boxplot again.

```
if(tBodyAccMag-mean()<=-0.8):
      Activity = "static"
if(tBodyAccMag-mean()>=-0.6):
          Activity = "dynamic"
```

In addition, using boxplot, we can easily distinguish WALKING DOWNSTAIRS behaviors from others.

```
if(tBodyAccMag-mean()>0.02):
  Activity = "WALKING_DOWNSTAIRS"
else:
      Activity = "others"
```

However, 25% of WALKING DOWNSTAIRS findings are still below 0.02, which are misclassified as others, so this situation makes a classification error of 25%.

C(b) Analysing Angle between X-axis and gravityMean feature

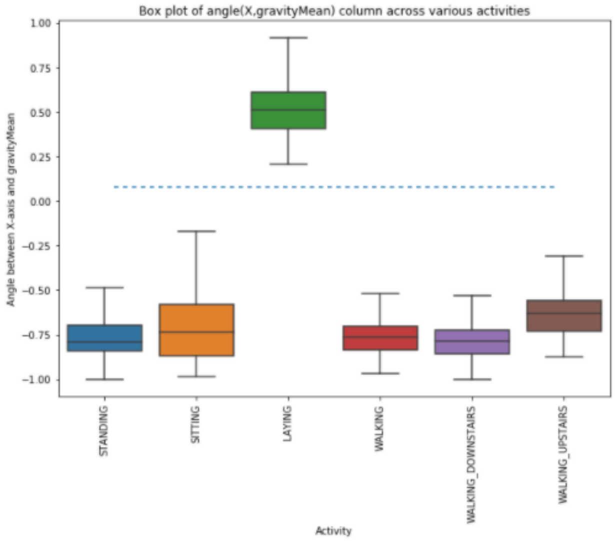

Figure 8. Box plot of angle(X.gravityMean) column

We can observe from the boxplot that angle(X,gravityMean) distinguishes LAYING perfectly from other activities.

```
if(angle(X,gravityMean)>0.01):
   Activity = "LAYING"
else:
   Activity = "others"
```

C.(c)Analysing Angle between Y-axis and gavityMean feature

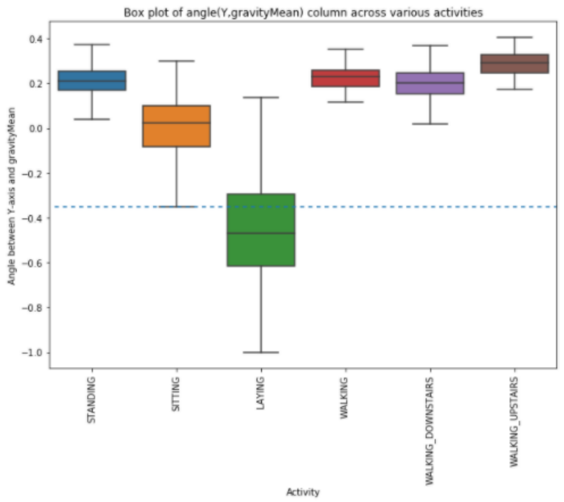

Figure 9. Box plot of angle(Y.gravityMean) column

Similarly, we can distinguish LAYING from other activities using Angle between Y-axis and gravityMean, but it leads to some error of misclassification again. [6]

C(d) Visualizing data using t-SNE

From an extremely high-dimensional space to a low-dimensional space, t-SNE data can be visualized and still retains a lot of real information. Provided that training data has 561 unique characteristics, let's visualize it in a 2D space using t-SNE. [7]

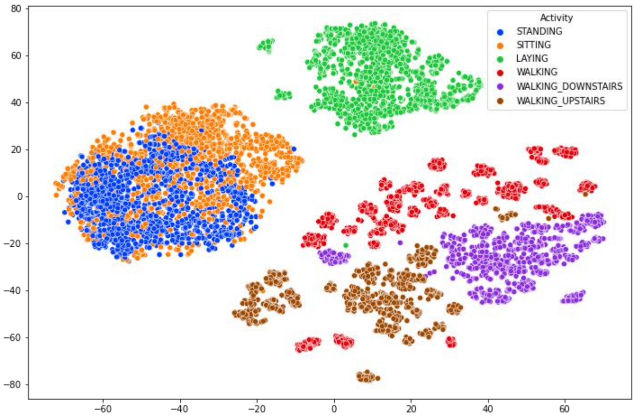

Figure 10. Visualization of activities using t-SNE in 2D space

We are able to visualize and separate all six activities in a 2D space using the two new components obtained through t-SNE.

D. IMPLEMENTATION and RESULTS

Getting training and test data ready

y_train = train.Activity
X_train = train.drop(['subject', 'Activity'], axis=1)
y_test = test.Activity
X_test = test.drop(['subject', 'Activity'], axis=1)
print ('Training data size: ', X_train.shape)
print ('Test data size: ', X_test.shape)

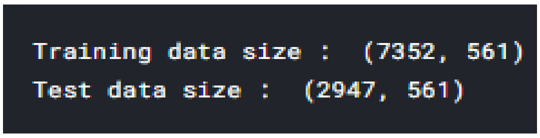

Figure 11. Training and Testing Data

Here, we have the size of the training data and testing data as shown in figure 11. Now, we will apply the usual machine learning classification algorithms on the dataset to get the best possible outcome in terms of performance accuracy.

D(a) Logistic regression model with Hyperparameter tuning and cross validation

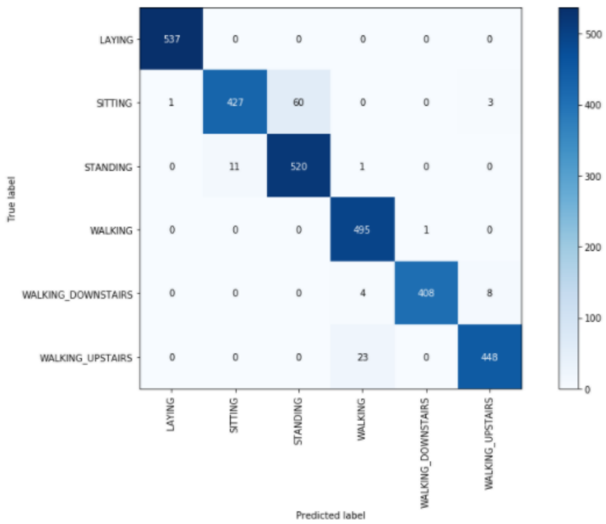

Figure 12. Confusion matrix for Logistic Regression

D(b)Linear SVM model with Hyperparameter tuning and cross validation

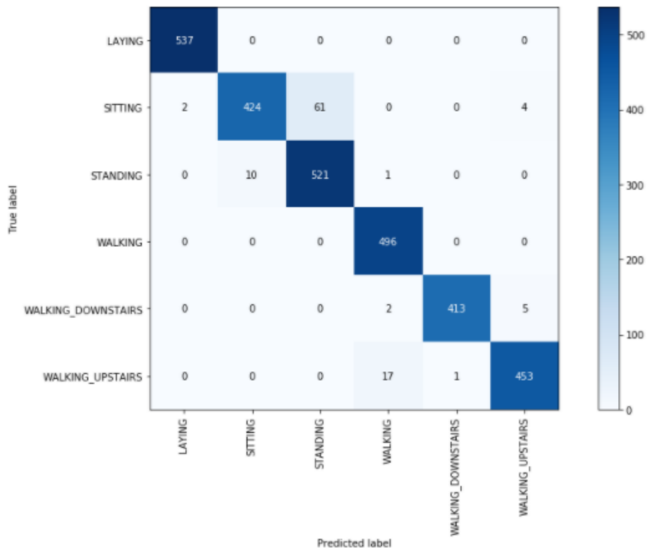

Figure 13. Confusion matrix for Linear SVM

D(c) Kernel SVM model with Hyperparameter tuning and cross validation

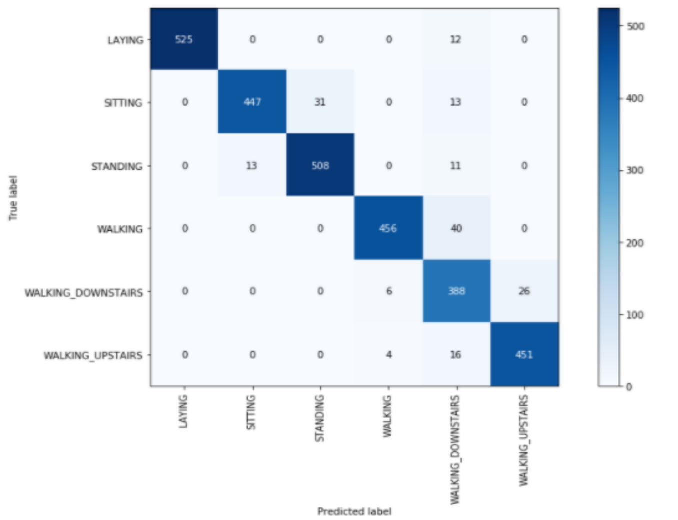

Figure 14. Confusion matrix for Kernel SVM

D.(d) Decision tree model with Hyperparameter tuning and cross validation

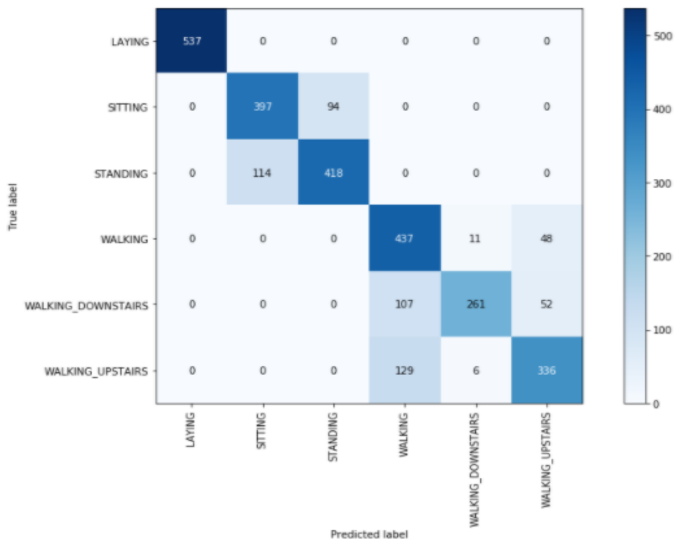

Figure 15. Confusion matrix for decision tree

D(e) Random forest model with Hyperparameter tuning and cross validation

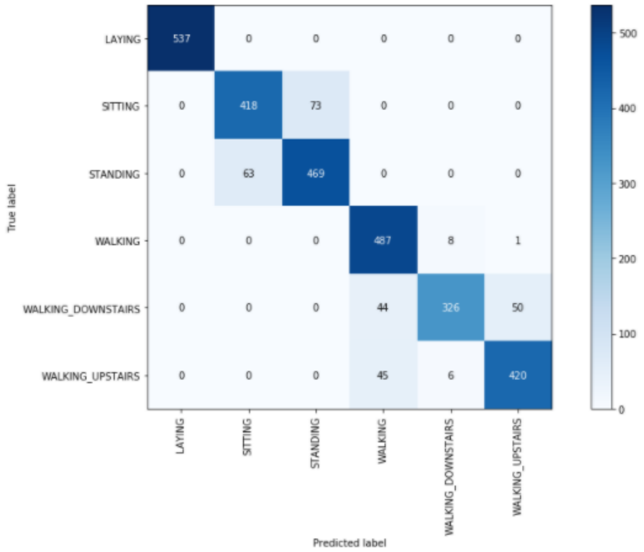

Figure 16. Confusion matrix for Random Forest

We have constructed several different models in this kernel using different classification algorithms. The precision reached by these models is as follows:

Logistic Regression	Linear SVM	Kernel SVM	Decision Trees	Random Forest
96.40	96.64	94.10	85.65	91.32

Table 1. Accuracy Performance

V. CONCLUSION

Human activity recognition has wide applications in medical research and human survey systems. The Human Activity Recognition database was created from the recordings of 30 study participants. Their daily living activities (ADL) are monitored while holding a waist-mounted smartphone with embedded inertial sensors. The aim is to classify operations into one of the six tasks performed. The experiments were performed within an age range of 19-48 years with a group of 30 volunteers. Six activities(WALKING,WALKINGUPSTAIRS,WALKINGDOWNSTAIRS,SITTING,STANDING,LAYING) were carried out by each person wearing a smartphone (Samsung Galaxy S II) on their waist. At a constant rate of 50Hz, we captured 3-axial linear acceleration and 3-axial angular velocity using its integrated accelerometer and gyroscope.

In this project, I used EDA by dividing the activities into Static and Dynamic categories. I used different ML models to train and test the data. While using multiple different models using various classification algorithms i.e., Logistic regression model, Linear SVM, Kernel SVM, Decision tree model and Random forest approaches, we can observe the best accuracy performance given by the Linear SVM (96.64%).

VI. INDIVIDUAL TASK

I did the whole project myself. To start with, I chose the idea of Human Activity recognition using smartphones from Kaggle and GitHub. I used the Kaggle dataset which is mentioned in the references section. I then performed the EDA on the dataset checking for duplicates etc., then used the tBodYAccMag-mean, and gravity mean feature for analyzing the angles between the axis. Furthermore, I used t-SNE for the visualization of the data. I used 5 different types of ML algorithms for the classification of datasets to get the performance accuracy for every algorithm. I penned the whole project report myself.

VII. ACKNOWLEDGEMENT

I would like to thank Prof. Ali Etemad for his continuous guidance and support to understand all the concepts used in the activity recognition framework.

VIII. REFERENCES

[1] Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine- International *Workshop of Ambient Assisted Living (IWAAL 2012)*. Vitoria-Gasteiz, Spain. Dec 2012.

[2] Energy Efficient Smartphone-Based Activity Recognition using Fixed-Point Arithmetic-*Journal of Universal Computer Science. Special Issue in Ambient Assisted Living: Home Care*. Volume 19, Issue 9. May 2013.

[3] Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine-4th International Workshop of Ambient Assited Living, IWAAL 2012, Vitoria-Gasteiz, Spain, December 3-5, 2012.

[4] Human Activity and Motion Disorder Recognition: Towards Smarter Interactive Cognitive Environments- *21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN* 2013. Bruges, Belgium 24-26 April 2013.

[5] Akram Bayat, Marc Pomplun, Dua A. Tran "A Study on Human Activity Recognition Using Accelerometer Data from Smartphones," in Elsevier B.V, Boston, USA, 2014.

[6] Deep Analysis for Smartphone-based Human Activity Recognition-2020 8th International Conference on Information and Communication Technology (ICoICT)

[7] Detailed Human Activity Recognition using Wearable Sensor and Smartphones - 2019 International Conference on Opto-Electronics and Applied Optics (Optronix)

[8] Human Activity Recognition Using Smartphones- 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)

[9] A Survey on Human Activity Recognition and Classification- 2020 International Conference on Communication and Signal Processing (ICCSP)

[10] https://www.kaggle.com/uciml/human-activity-recognition-with-smartphones

[11] https://github.com/sachingoyal2910/Human-Activity-Recognition-using-machine-learning