Human activity recognition using Smartphone

Raju Karmuri

CS Graduate

CSCI5710 - Introduction to Machine Learning

Abstract

Understanding and improving the quality of human lives has been a challenge for scientists and experts over the centuries. In this project, we develop a multi-class classification model that can predict human actions like sitting, standing, laying, and walking based on the data collected from smartphones. The proposed system has the accelerometer and gyroscope signals generated by 30 volunteers from age group 19-48 years wearing a smartphone on their waist. The dataset is randomly divided into 70% training and testing 30% and the signals are pre-processed using noise filters and sampled in fixed-width sliding. The classification models Logistic Regression, Support Vector Machine, Random Forest, KNN, Decision trees, and Neural networks have been used to classify the activities of six categories. The results of the experiments show excellent accuracy scores in all the classifiers that help us to accurately predict human actions based on smartphone data.

I.Introduction

Improving the quality of human life has been a major concern for health care scientists, experts, and researchers. With the help of Machine Learning, Deep Learning, and Neural networks techniques we can improve the quality of human lives. These techniques will be very beneficial for the elderly to maintain their health in rehabilitation centers, diabetic centers, and also save a huge number of resources with sensors that can help caretakers record and monitor the patients all the time and report automatically when any abnormal behavior is detected. Smartphones have increasingly become a needful part of daily lives over the last decade and it has been growing even now. So using smartphones and smart devices is a very reliable and handful asset to detect the motion of human bodies.

All the smartphones come with an accelerometer and gyroscope sensor as default which will be helpful to detect not only the motion but also the angles at which the body moves from all dimensions. The developed system takes the body acceleration from the accelerometer in triaxial acceleration dimensions (XYZ) and body angles from the gyroscope in triaxial angular velocity dimensions(XYZ). The processed signals are used to extract data points that show the activity labels and features for a dataset. Overall, we got 561 features and 6 labels from the training sets and testing datasets. Using this data many classification algorithms have been developed to experiment and test the accuracy of the body sensor signals. Also, neural network models have been developed to predict body actions. As this is a multi-class classification model, a confusion matrix has been used to find the accuracy, precision, and recall scores of all the classifiers.

All the classifiers performed well with the given datasets. Neural networks have been developed using sigmoid and softmax activation functions. The accuracy results of neural networks are also decent and feature importance analysis using random forest has been developed. With these results from various experiments, the main goal of detecting human interactions will be accurate and improve the fitness of human bodies.

II. Related Work

Many types of research have been developed with Human activity recognition. Some of them have used accelerometer signals alone to detect the body motions, while others have used both accelerometer and gyroscope sensor signals. One of the interesting experiments was using Dimensionality Reduction, which produced the most accurate results using LDA analysis(*Human Activity Recognition using Smartphone by Amin Rasekh, Chien-An Chen, Yan*

Lu from Texas A&M University). Another interesting experiment was using deep learning recurrent neural networks, it is fascinating to see that convolution neural networks outperform not only the traditional machine learning models but also yield better performance than other CNN techniques like deep belief networks(*Deep Recurrent Neural Networks for Human Activity Recognition by Abdulmajid Murad and Jae-Young Pyun, Department of Information Communication Engineering, Chosun University, 375 Susuk-dong, Dong-gu, Gwangju 501-759, Korea*.).

III. Data set pre-processing and visualization

The data for this project is collected from Kaggle, which is originally extracted from UCI Machine Learning Repository. The features for this database are captured from the signals generated by the accelerometer and gyroscope. Triaxial accelerations and triaxial angular velocities are denoted using 'XYZ'. All features are divided into time signals denoted by 't' and frequency signals (i.e., Fast Fourier Transform signals) denoted by 'f'. Additionally, some features of the signals have been obtained by averaging the signal samples to get the mean, median, standard deviation, and other additional vectors. Filters were used to remove background noise.

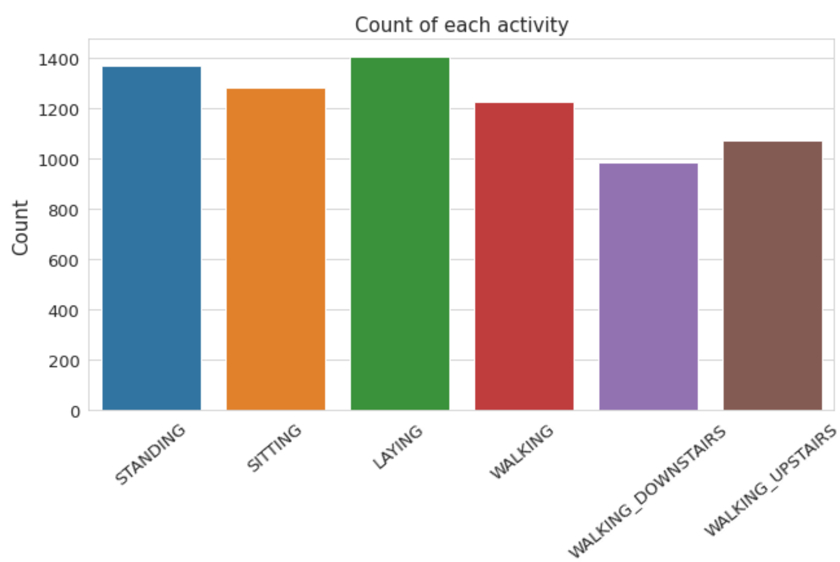
Both the train and test datasets consist of 561 features, 1 subject, and 1 output label called activity. There are six categories of output labels they are STANDING, WALKING_DOWNSTAIRS, WALKING_DOWNSTAIRS, LAYING, and STANDING. All output labels are categorical.

```
train_df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7352 entries, 0 to 7351
Columns: 563 entries, tBodyAcc-mean()-X to Activity
dtypes: float64(561), int64(1), object(1)
memory usage: 31.6+ MB
```

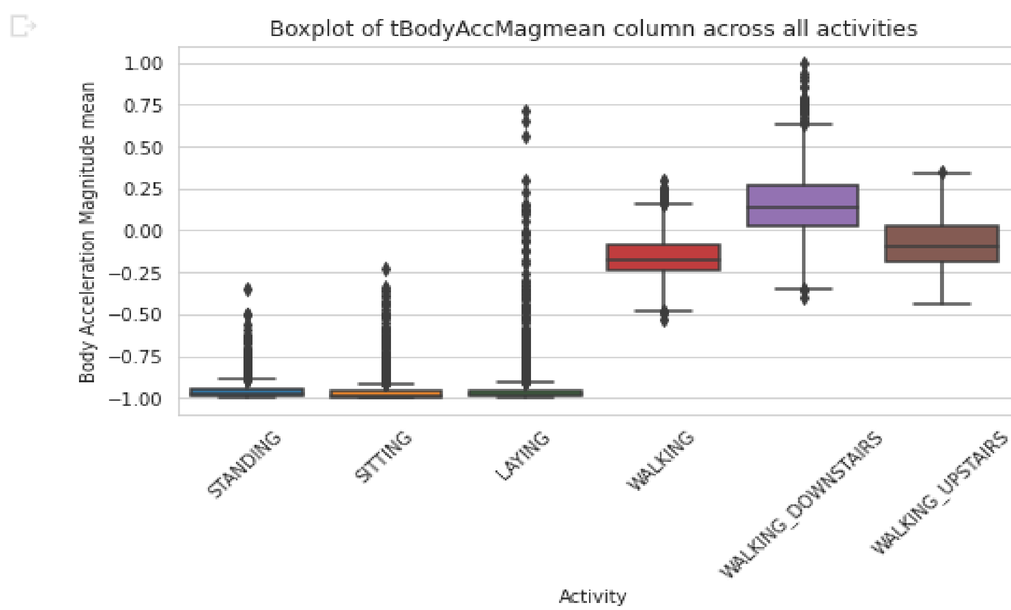
The data is already divided into train and test sets by UCI, therefore there was no need to split it. There were no outliers as the data points are between -1 to 1, there were no missing values, no duplicated values, and the label encoding has been done to the target variables, as they are categorical.

The data is visualized using the matplotlib and seaborn packages. As there were multiple outputs for the features, initially it was difficult to understand and plot them. Therefore, a bar plot has been used to get the count of all output. All subjects produced an equal amount of data.



IV. Experiments

The Box plot helped a lot in understanding how the trend of the signals changes in the output with all the features. From the box plot, we can see clearly that the static moment signals like standing, lying, and sitting are different from the dynamic motion signals like walking.



We can also see from the box plot that WALKING_DOWNSTAIRS is misclassified from other dynamic signals and the LAYING signal is different from the other static signals. Therefore, it is difficult to apply binary classification on these output labels and that is why we are using multi-class classification methods.

The main goal of this project is to predict the action of the subject and there are six outcomes. Based on the outcome, we can know the health and fitness of a person. Different classifiers have been used to model the data, and all the classifiers produced amazing results.

Logistic regression is used to predict the accuracy of all the labels, as it involves six outcomes, maximum iterations had to be set to 6000, otherwise, the results were not expected to be shown. The KNN algorithm is used with the nearest neighbor set to 10, an SVM algorithm is run with gamma set to 0.01, and C set to 10. Random forest algorithm and decision trees are used with the random state set to 42, and neural network is designed using the sigmoid and softmax activation functions. Other activation functions like Relu or LeRelu produced very bad results, therefore we used softmax. The accuracy score on the neural network was 0.98 and other models also produced above 90% of accuracy score.

The main challenge was to find the precision and recall score based on the confusion matrix for all the models. As it is a multi-class classification dataset, there were 6 classes, 6 precision scores, and 6 recall scores that were calculated manually for all models using the confusion matrix. Also, I couldn’t succeed in developing CNN networks to find the accuracy.

Results

The results of all the classification models are as follows:

	Accuracy
Logistic Regression	0.9613165931455717
K-nearest neighbor	0.9066847641669494
SVM	0.9558873430607397
Random Forest	0.9260264675941635
Decision Trees	0.8622327790973872
Neural networks	0.9872

Discussion

The results are very accurate and the precision and recall scores are also decent. The neural networks and deep learning algorithms outperformed all other machine learning models with an accuracy rate of 98%. Other classification models also did well but I think the time series data will produce better results with the CNN and deep learning algorithms. In contemporary society, all use smartphones and get real-time data. The time-series data like Human activity recognition is rapidly growing because of smart devices and it gives a lot of information about human health and fitness. This research is a continuous process, getting the right predicted models to improve the quality of human life.

References

1. *https://github.com/anas337/Human-Activity-Recognition-Using-Smartphones.github.io*
2. *Morris, M., Lundell, J., Dishman, E., Needham, B.: New Perspectives on Ubiquitous Computing from Ethnographic Study of Elders with Cognitive Decline. In: Proc. Ubicomp (2003).*
3. *Deep Recurrent Neural Networks for Human Activity Recognition by Abdulmajid Murad and Jae-Young Pyun, Department of Information Communication Engineering, Chosun University, 375 Susuk-dong, Dong-gu, Gwangju 501-759, Korea.*
4. *https://www.researchgate.net/publication/260003860_Human_Activity_Recognition_using_Smartphone*
5. *https://valueml.com/human-activity-recognition-using-smartphone-data-python/*