International Journal of Latest Trends in Engineering and Technology
Special Issue SACAIM 2017, pp. 006-012
e-ISSN:2278-621X

GROUP BASED HUMAN PHYSICAL ACTIVITY CLASSIFICATION (GBHAC) USING SHIMMER2 WEARABLE SENSOR DATA SETS

Doreswamy[1] & Yogesh K M[2]

Abstract - Human activity recognition is one of the most emerging fields of research in pervasive computing. It has realized more interest in several research communities given that understanding user activities and behavior help to deliver proactive and personalized services. In wearable computing scenarios, activities such as climbing stairs, walking, sitting and relaxing, and lying down can be inferred from data provided by shimmer2 acceleration sensors. In such scenarios, most approaches use a single or three dimensional features, regardless of which activity to be recognized. This paper describes how to recognize certain types of human physical activities using tri-accelerometer's three dimensional data generated by a shimmer2 wearable sensor device and described how recognition rates can be improved by careful selection of individual group based classification such as gender, age, height and weight for each physical activity. We present a systematic analysis of features computed from an accelerometer shimmer2 mHealth data set and different classifiers are studied on group based classification.
Keywords – Activity Recognition, Wearable Sensor, Accelerometer, Classification

1. INTRODUCTION

Activity Recognition is an emerging field of research, born from the larger fields of ubiquitous computing, context-aware computing and multimedia. Recognizing everyday life activities is becoming a challenging application in pervasive computing, with a lot of interesting developments in the health care domain, the human behavior modeling domain and the human-machine interaction domain. Even if first works about activity recognition used high dimensional and densely sampled audio and video streams, in many recent works activity recognition is based on classifying sensory data using one or many accelerometers [2]. Accelerometers have been widely accepted due to their compact size, their low-power requirement, low cost, non-intrusiveness and capacity to provide data directly related to the motion of people. In recent years, several papers have been published where accelerometer data analysis has been applied and investigated for physical activity recognition. Nevertheless few of them override the difficulty to perform experiments out of the lab.

The condition to perform experiments out of the lab creates the need to build easy to use and easy to wear systems in order to free the testers from the expensive task of labeling the activities they perform. The identification of human activities has attracted very much interest lately. Typically, wearable sensors are used to register body motion signals that are analyzed by following a set of signal processing and machine learning steps to recognize the activity performed by the user. Most of the existing works in this area contribute with diverse models that normally yield very high recognition capabilities. However, a major part of these solutions have only been validated in controlled environments and through online evaluations. More importantly, there is a lack of papers covering the whole design process for the development of a system that can actually recognize human activity in realistic settings. This paper aims to identification of body motion. Body motion can vary from one person to another. Motion differs from group and sub-group classification methods. The motion various from different ages of people, different gender, and also it's varies from different weights. In our works, we aim at capturing the motions of all the parts of the body for a thorough study of the activity recognition problems.

2. RELATED WORK

Human activity recognition using wearable sensors is a very widespread research subject. Phone-based accelerometers to perform activity recognition, a task which involves identifying the physical activity a user is performing. To implement this system, authors was collected labeled accelerometer data from twenty-nine users as they performed daily activities such as walking, jogging, climbing stairs, sitting, and standing, and then aggregated this time series data into examples that summarize the user activity over 10-second intervals. Authors then used the resulting training data to induce a predictive model for activity recognition. Earlier work by Kwapisz J R et al. [3], introduce a pushed forward estimation of soft biometric information from inertial sensor. By solving different classification tasks like age, weight and height based on the motion data of human walking steps represented by accelerations and angular velocities. Data were recorded by one sensor placed at various locations on the human body, namely the chest, the lower back, the wrist and the ankle. The results show that these classification tasks can be solved well by using accelerometers and/or gyroscopes at any of the given locations. The classification rates were the highest

[1] Department of Computer Science Mangalore University, Mangalore University, Mangalore, Karnataka, India
[2] Department of Computer Science Mangalore University, Mangalore University, Mangalore, Karnataka, India

for sensors located at the lower back and chest in each of the experiments, but still realistically high when the sensor is attached to the wrist or ankle. The experiments have made clear that there is not one feature mainly responsible for any of the distinctions necessary for a classification. However, the feature importance in each of the classification gave pointers as to what combination of features produces the best results. The most important findings were that angular velocities did not perform better than accelerations. Qaiser Riaz Akker et al. [5] and earlier work by Incel et al. [6] survey activity recognition research using smart phones. However, most research described therein still involves offline processing of the data collected on the smart phone. Kunze et al. [7] studied how acceleration and gyroscope signals are affected by sensor displacement.

2.1 Feature Generation and Data Transformation

The Standard classification algorithms cannot be directly applied to raw time-series accelerometer data. Instead, we first must transform the raw time series data into examples duration ED. To accomplish this we divided the data into 10-second, 20-second, 30-second, 40-second, 50-second and 60-second segments and then generated features that were based on the 512 instance. We frequently chose a 10, 20, 30, 40, 50, and 60 second because we felt that it provided sufficient time to capture several repetitions of the motions involved in some of the twelve activities.

2.2 Activity set

The accelerometer shimmer2 sensor data sets contains 12, 15,749 (in Lacks) instances. All sensing modalities are recorded at a sampling rate of 50 Hz, which is considered sufficient for capturing human activity, 0.02/ms time was take for generation one instance. We can calculate. Time for one instance = 1/50 = 0.02/ms. The user was taken 1/min time for each activities, except waist bend forward, front elevation of arms, knee bending and jump front & back for those activity user was taken 20/sec only. We can calculate time taken for each activity using the below formula.

$$T = T_A = \frac{E_t * I_t}{t}$$

T_A = Total time taken for each activity
E_t = Time taken for each instance
I_t= Total number of instance in an individual activity
t = Time in sec
For an instance for activity, standing still is given below:

$$Running = \frac{0.02 * 3072}{60/\sec} = [61.44]$$

Standard classification algorithms can be directly applied to raw time-series accelerometer data, but it leads computational complexity and decline the performance of a classifier. Therefore, data transformation is deployed on the raw time series window samples. To accomplish this data is divided into varying windows sizes that are corresponding to 10, 20, 30, 40, 50 and 60-second segments containing 512, 1024 1536, 2048, 2560 and 3072, readings respectively. In this study we consider 12 human physical activities. We selected daily activities because they are performed regularly by many people in their daily routines. Furthermore, most of these activities involve repetitive motions and we believe this should also make the activities easier to recognize

Sl.No	**Human Physical Activity(HPA)**	**Actual Time Taken in seconds**	**≈ Time**
1	Standing still	61.44	1 min
2	Sitting and relaxing	61.44	1 min
3	Lying down	61.44	1 min
4	Walking	61.44	1 min
5	Climbing stairs	61.44	1 min
6	Waist bends forward	**20 sec**	**20 sec**
7	Front elevation of arms	**20 sec**	**20 sec**
8	Knees bending	**20 sec**	**20 sec**
9	Cycling	61.44	1 min
10	Jogging	61.44	1 min
11	Running	61.44	1 min
12	Jump front & back	20 sec	20 sec

Table 1 Activity set

3. CLASSIFICATION

Classification is the process of building a model (or functions), which describes and distinguishes data classes or concepts, for the purpose of being able to use the model to predict the class of objects whose class label is unknown. The classification models are constructed based on the analysis of $3/4^{th}$ of training data whose class label is known and tested on $1/4^{th}$ of actual instances as testing samples. The advanced Data Mining Techniques deployed for classification of mHealth data sets are tabulated in table 3. Human activity recognition systems rely on advanced data mining algorithms to predict an individual's activity during a certain period of time. In addition, it has been emphasized that different classification methods could be used in HAR systems, depending on the specific characteristics of each scenario (e.g., the set of activities, the type of sensors, and so forth). Miguel A. Labrador Oscar D.et al. [8] elaborate on the advantages of implementing a completely mobile HAR system in terms of reliability, scalability, and energy consumption, just to mention a few. But such a task entails the evaluation of a classification model on the Smartphone, which brings about an additional challenge that is implementing each and every classifier under the Android platform. This could be very time consuming given the underlying complexity in the implementation of machine learning algorithms, along with the computational constraints present in mobile devices. The focus of this paper, therefore, is to classified shimmer2 accelerometer data based implementations of a number of machine learning classification methods provided by WEKA to enable classifier evaluation in order to build the accuracy.

3.1 Analysis of results

Acceleration-based physical monitoring algorithms can be validated in identifying different postures and movements using precision, recall and F-measure. The precision or positive predictive value (PPV) is defined as the proportion of instances that belongs to a class (TP: True Positive) by the total instances, including TP and FP (False Positive) classified by the classifier as belong to this particular class.

$$\mathrm{Precision} = \frac{TP}{TP+FP}$$

The recall or sensitivity is defined as the proportion of instances classified in one class by the total instances belonging to that class. The total number of instances of a class includes TP and FN (False Negative).

$$\mathrm{Recall} = \frac{TP}{TP+FN}$$

The F-measure is the combination of precision and recall and is defined as

$$F-Measures = \frac{2 \times \mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}$$

IV. MECHINE LEARNING CLASSIFIERS

3.2 J48 algorithm :

J48 classifier is a simple C4.5 decision tree for classification. It creates a binary tree. The decision tree approach is most useful in group based classification. With this technique, a tree is constructed to model the classification process. Once the tree is built, it is applied to each tuple in the database and results in classification for that tuple. In our experiments it was seen that J48 Decision trees performed really well. In most cases, their performance was almost at par with that of Support Vector Machines. In fact in several cases, it was seen that J48 Decision Trees also had more than 85% of accuracy than either Naïve Bayes, or other classifiers.

3.3 Random Forest

Random Forests [70] consists of a combination of decision-trees. It improves the classification performance of a single-tree classifier by combining the bootstrap aggregating (bagging) method and randomization in the selection of partitioning data nodes in the construction of decision tree. The assignment of a new observation vector to a class is based on a majority vote of the different decisions provided by each tree constituting the forest. However, RF needs huge amount of labeled data to achieve good performances. In this paper authors proposed a group based classification methodology to recognize human physical activity, using shimmer2 acceleration sensor data sets, different classes of motions, such as running walking, jogging, climbing stairs and twelve different types of activities, by comparing different machine learning techniques (Random Forests, RT, MLP and NB). The authors showed that in this group based classification method Random Forest algorithm provides the highest average 99% accuracy outperforming the MLP and the Naive Bayes classifiers.

3.4 Decision Trees

Decision trees are one of the common algorithms for classification problems such as Human Activity Recognition. First model was built using J48 decision tree for classification. Decision Trees are easy to understand. However, if there is a non-linear relationship between predictors and outcome, in this paper we consider group based classification methods decision tree (DT) were got more than 95% of classification accuracy values.

The general algorithm for building decision trees is

1. Check for the above base cases.
2. For each attribute a, find the normalized information gain ratio from splitting on a.
3. Let a best be the attribute with the highest normalized information gain.
4. Create a decision node that splits on a best.
5. Recur on the sub lists obtained by splitting on a best, and add those nodes as children of node.

3.5 Multi-layer perceptron

Multi-layer perceptron (MLP) consists of multiple layers with nodes using weighted connections. Each layer is fully connected to the next one. Between the input and output layers, there can be one or more hidden layers. Weights measure the degree of correlation between the activity levels of neurons that they connect. Moreover, a training algorithm needs to be used to adjust the weights. The most popular, the Back propagation is composed by two phase. It is the training or learning algorithm or multilayered perception. The network is first initialized by shaking off all its weight to small random values. Next the input data is applied and corresponding output is calculated. The calculation gives an output which is completely different from what we want because all the weights are taken randomly. We then calculate the errors which is equal to the target output which is the actual output. The error is then used to change the weight in such a way that the error will get the minimum value until the desired output is obtained for group based classification multi layer perceptron (MLP) classifiers were got 68% of accuracy values, its vary low value compare to random forest tree (RFT), j48 algorithms and decision trees (DT)

4. EXPERIMENTAL RESULTS

4.1 Group and sub-group based classification

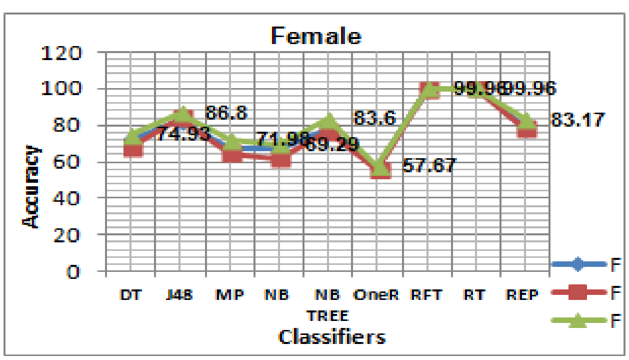
Training and validation data were prepared for Accelerometer chest sensor using the features extracted. Four types of group classification tasks were performed: (1) gender classification, (2) Height classification, (3) age classification. Furthermore; and (4) Weight classification training and validation data were also prepared for classification within participant subgroups for height and age classification. In table the characteristics of the population within different classification tasks are presented. Forage and height classification, we choose classes based on the available data.

Group Classification Tasks	Classes	N
Gender Classification	Male	3
	Female	7
Age Classification	> 25	5
	< 25	5
Height Classification	> 170 cm	3
	< 170 cm	7
Weight Classification	> 70 kg	5
	< 70 kg	5

Table: 2 Participate with different classification task

4.2 Gender Classification

Our goal was to show that classification tasks regarding the gender classification of the trial subject can be performed sufficiently well by using the shimmer2 accelerometer sensors data sets. The gender can be identified by motion recordings of any of the employed sensors. The graph represents in figure 4 individually.



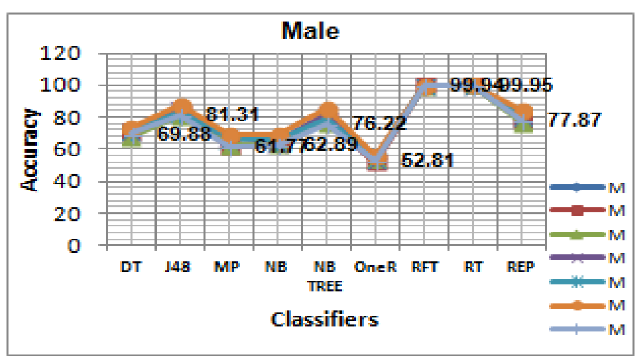


Fig: 1 Classification accuracy between Male and Female Participates

4.3 Height Classification

Another goal was height classification from accelerometer shimmer2 sensor data sets. The body height can be identified by motion recordings of any of the employed sensors the results of the naïve bayes advanced classifier were got more accuracy rate are shown in Figure 4.

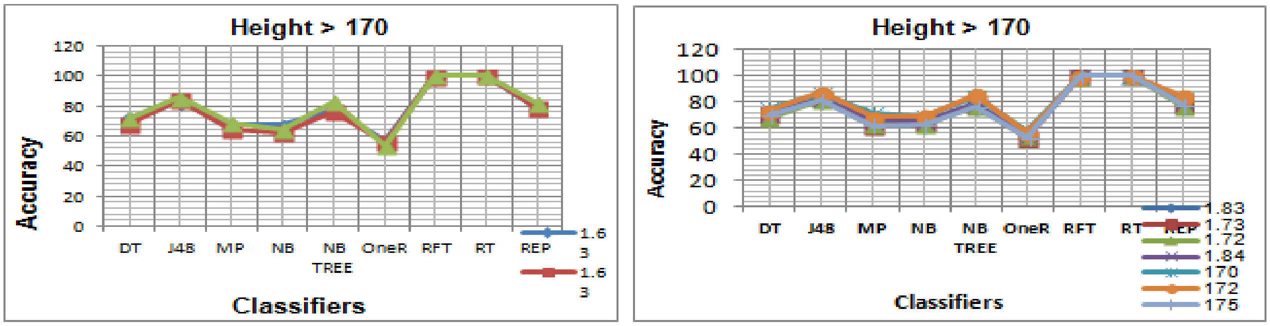


Fig: 2 Classification accuracy between heights of the person between Height >170 and < 170

4.4 Age Classification

Another goal was age classification from shimmer2 accelerometer sensor data sets. On the basis of age can be identified by motion recordings of any of the employed sensors the results of the classification accuracy between age of the person between less than 25 and greater than 25 are shown in figure5.

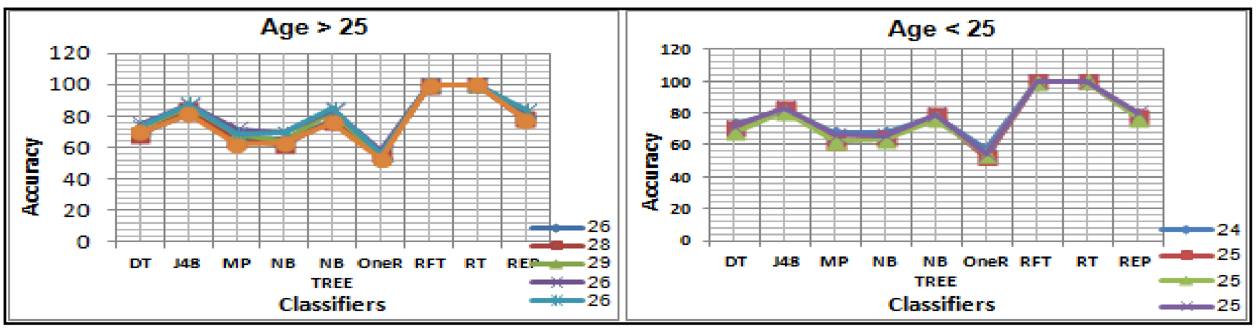


Fig: 3 Classification accuracy between Age of the person between Age > 25 and < 25

4.5 Weight Classification

Another goal was weight classification from shimmer2 accelerometer sensor data sets. On the basis of weight can be identified by motion recordings of any of the employed sensors the results of the classification accuracy between weight of the person between less than 70 and greater than 70 are shown in figure5.

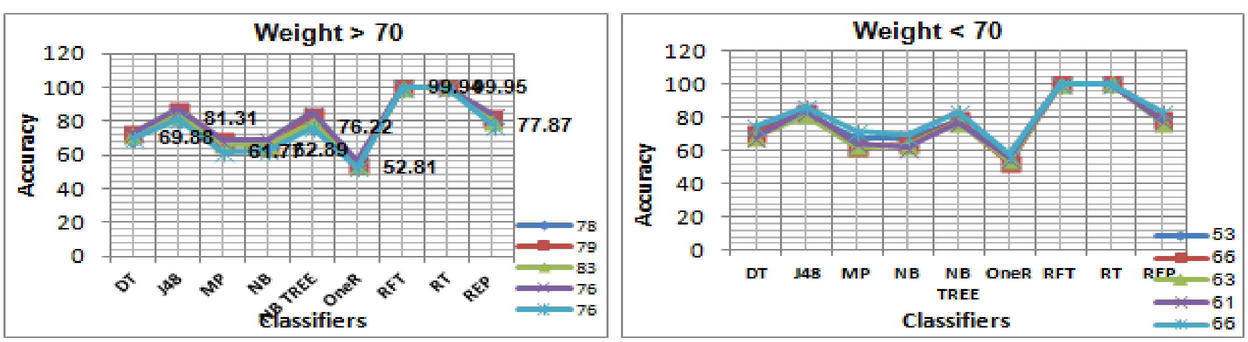


Fig: 6 Classification accuracy between Weight of the person between Weight > 70 and < 70

Once the model is constructed and evaluated to ensure the performance of classifiers on testing samples, classification accuracy alone is typically not enough information to make this decision. Further, the different advanced classification methods such as Naïve bayes, multilayer perceptron, decision Table, OneR, J48, Random Forest, Random Tree, Reduced Error Pruning (REP) Tree and naive Bayes Tree were experimented and evaluated on group based data. These are respectively called as Group based classification (GBD. The experimental results of each one are described in the following section.

GROUP	DT	J48	MLP	NB	NBT	OneR	RFT	RT	REP
MALE	70.79	83.33	65.26	65.02	79.21	53.97	99.96	99.96	79.75
FEMALE	71.87	84.39	68.04	66.12	79.37	56.72	99.96	99.96	79.9
H >=172	70.52	82.93	64.7	65.09	78.63	54.06	99.97	99.97	79.33
H<172	72.01	84.71	68.19	65.73	80.2	55.9	99.95	99.95	80.49
W>=70	71.37	83.81	66.45	65.53	79.95	54.3	99.96	99.96	80.66
W<70	70.86	83.48	65.73	65.17	78.56	55.28	99.96	99.97	78.93
AGE>=26	71.42	84.38	67.04	65.51	80.16	55.06	99.95	99.96	80.74
AGE < 26	70.66	82.55	64.68	65.11	77.89	54.4	99.97	99.97	78.38

Table3 Experimental Result

The classification problem, given a set of simple trees and a set of random predictor variables, the random forest method defines a margins function that measures the extent to which the average number of votes for the correct class exceeds the variable. This measure provides us not only with a convenient way of making predictions, but also with a way of associating a confidence measure with those predictions. Table 3 show the group based classification performance in the above table shows that final classification accuracy of all different types group classification. In this group classification Random forest tree (RFT) and Random tree (RT) were got more than **98%** of accuracy rate in all types of group based classification.

5. CONCLUSION

The Using machine learning classifiers, recognition accuracy of over 90% in particular classifiers like Random forest tree (RFT) and Random tree (RT) on a variety of 12 different everyday activities was achieved using shimmer2 variable sensor data acquired with supervision from 10 subjects. This work shows acceleration sensor data can be used to activity recognition based on group based classification method. Using weka tools for standard 10-flod cross-validation, the classification rates have been for group based classification. Next we are planning to add more human physical activities to our work like biking, riding a car or bus, smoking eating etc. The number of participants is to be increased as well to collect an extensive set of human activity data set. Moreover, it is less sensitivity to orientation and position unlike motion sensor in the wearable device. Further investigation is focused on different group and sub group classification with their reduced feature data sets by applying neural network and deep learning model.

6. ACKNOWLEDGEMENT

The authors wish to acknowledge the financial support of SERB-DST, New Delhi vide SB/EMEQ-137/2014, dated 21-03-2016, and also gratefully acknowledge the unanimous reviewers for their kind suggestion and comments for improving this paper.

7. REFERENCES

[1] Banos, O., Garcia, R., Holgado, J. A., Damas, M., Pomares, H., Rojas, I., Saez, A., Villalonga, “mHealthDroid: a novel framework for agile development of mobile health applications”. Proceedings of the 6th International Work-conference on Ambient Assisted Living an Active Ageing (IWAAL 2014), Belfast, Northern Ireland, December 2-5, (2014). Data set available at http://www.ugr.es/~oresti/datasets.html.

[2] Banos, Oresti, et al. "On the Development of a Real-Time Multi-sensor Activity Recognition System." Ambient Assisted Living. ICT-based Solutions in Real Life Situations. Springer International Publishing, 2015. 176-182.

[3] Kwapisz, Jennifer R., Gary M. Weiss, and Samuel A. Moore. "Activity recognition using cell phone accelerometers." ACM SigKDD Explorations Newsletter 12.2 (2011): 74-82.

[4] Banos, Oresti, et al. "Recognition of human physical activity based on a novel hierarchical weighted classification scheme." Neural Networks (IJCNN), the 2011 International Joint Conference on. IEEE, 2011.

[5] Q. Riaz, A. Vogele, B. Kruger, and A. Weber, “One small step for a man: estimation of gender, age and height from recordings of one step by a single inertial sensor,” Sensors, vol. 15, no. 12, pp. 31 999–32 019, 2015.

[6] O. D. Incel, M. Kose, and C. Ersoy, “A review and taxonomy of activity recognition on mobile phones,” BioNanoScience, vol. 3, no. 2, pp. 145– 171, 2013

[7] K. Kunze and P. Lukowicz, “Dealing with sensor displacement in motion-based on body activity recognition systems,” in Proceedings of the 10th international conference on Ubiquitous computing. ACM, 2008, pp. 20–29.

[8] Lara, Oscar D., and Miguel A. Labrador. "A survey on human activity recognition using wearable sensors." Communications Surveys & Tutorials, IEEE 15.3 (2013): 1192-1209.

[9] Witten I. H., Frank E., Hall M. A. Author, “The explorer,” in Data Mining Practical Machine Learning Tools and Techniques, 3rd ed. United States: Morgan Kaufmann, 2011, ch.11, p. 445-479

[10] Dash, Yajnaseni, Sanjay Kumar, and V. K. Patle. "A Novel Data Mining Scheme for Smartphone Activity Recognition by Accelerometer Sensor." Proceedings of the 4th International Conference on Frontiers in Intelligent Computing: Theory and Applications(FICTA) 2015. Springer India, 2016.

[11] Banos, Oresti, et al. "Daily living activity recognition based on statistical feature quality group selection." Expert Systems with Applications 39.9 (2012): 8013-8021.

[12] Patil, Tina R., and S. S. Sherekar. "Performance analysis of Naive Bayes and J48 classification algorithm for data classification." International Journal of Computer Science and Applications 6.2 (2013): 256-261.

[13] Catal, Cagatay, et al. "On the use of ensemble of classifiers for accelerometer-based activity recognition." Applied Soft Computing 37 (2015): 1018-1022.

[14] Bayat, Akram, Marc Pomplun, and Duc A. Tran. "A study on human activity recognition using accelerometer data from smartphones." Procedia Computer Science 34 (2014): 450-457.