Personification of Bag-of-features Dataset for Real Time Activity Recognition

Moses L. Gadebe and Okuthe P. Kogeda
Department of Computer Science
Faculty of Information Communication Technology
Tshwane University of Technology
Private Bag X680, Pretoria 0001, South Africa
Gadebeml@tut.ac.za, KogedaPO@tut.ac.za

Abstract*—Personalization of activity recognition is possible and important, when existing public dataset collected from large group of subjects can be tailored and be used as training and testing dataset for new users (subjects) who have similar personal traits. However, due to shortage of personalized dataset and techniques to tailor public dataset for new users weakens the personalization of human activity. To address shortage of personalized dataset, we propose a personification algorithm that extracts and tailor-make bag-of-features dataset to support new users using publicly available Human Activity Recognition dataset (PAMAP2 and USC-HAD). Studies indicate that BMI can be used to profile user's weight as either normal weight or overweight or obese, which could be used to predict cardiovascular diseases. For that purpose our personification algorithm uses height, weight and BMI to generate human activity bag-of-features. The personification algorithm is implemented in Scala and Java programming languages and is deployed on Apache Spark Server. We validated our algorithm, by running three set trials of experiments for each 5 K threshold values using 2 randomly selected new user's profile against two publicly available Human Activity Recognition dataset PAMAP2 and USC-HAD. The results indicate that it is possible to tailor bag-of-features from public dataset. Overall performance of our algorithm shows precision, recall and F-score of 0.70%, 0.50% and 0.60% respectively.

Keywords-personalization; BMI, dataset; bag-of-features; subject; new user; Human Activity Recognition; Euclidean distance; K Nearest Neighbour; personification algorithm

I. INTRODUCTION

Studies conducted in cardiovascular diseases in particular stroke single-out physical activity as a main contributor to improvement of glucose intolerance, cholesterol, Body Mass Index (BMI), blood pressure, pulse rate, vital capacity and oxygen intake [1][2][3]. A commitment of 30 minutes moderate physical activity for five days a week can produce long-term health benefits, and thus diminish high morbidity and mortality [1].

Personification and monitoring of human activity is possible using devices that are affordable and available 24/7 at a closer proximity of their owners. Fortunately smartphones service owners at proximity of 0 centimeters to 60 centimeters daily and they are equipped with powerful motion sensors such as accelerometers, gyroscopes, ambient light sensors, temperature sensors, Global Positioning Systems, barometer, orientation sensors and storage capability [4][5][6]. Smartphone accelerometer and gyroscope sensory data have great potential in the identification of human activity to reduce chance of obesity. Normally Human Activity Recognition (HAR) in pattern recognition is divided into three steps: first, sensory data collection from accelerometer and gyroscope sensors, and labeling of collected sensor data as human activity; second, feature extraction to reduce classification features from dataset also known as pre-processing [7]. Third and last, is the classification step based on machine learning algorithms such as K Nearest Neighbor, Support Vector Machine, Hidden Markov model and others [5][7]. Sensory dataset for benchmarking real time HAR systems activity is still lagging behind due to the shortage of personalized dataset [7][8]. Recently researchers in HAR collected and published their datasets online to allow other researchers to benchmark their HAR systems [9]-[12]. However, some of the datasets lack personal attributes (e.g., age, height, weight, body mass index) and are in high dimensional space, which require high computational power and large memory [13]. Recent HAR models are mindful of resource constraint smartphone visa-vie classification requirements [7][14][15]. Nevertheless, due to the shortage of personalized datasets most of the proposed models produce poor. Hence in this paper, we propose personification algorithm to generate reduced bag-of-features dataset from public available datasets to tackle shortage of personalized datasets based on weight and BMI of individuals. Because, BMI is an efficient measure to categorize individuals as either normal weight or overweight or obese [2][3].

The rest of this paper is structured as follows: In Section II, we present the related work. In Section III, we present the methodology followed to generate personalized bag-of-features. In Section IV, we present the results and discussion. Lastly, in Section V, we provide the conclusion and future work.

II. RELATED WORK

Human Activity Recognition (HAR) has been studied for decades, but there is still limited number of publicly available dataset for machine learning. Researchers in [9] investigated and established that standard benchmarking HAR dataset was lacking. To close the gap, they collected and published dataset called PAMAP2 collected from 9 subjects. They used 3 Colibri wireless Inertial Measurement

978-1-5090-3696-7/16 $31.00 © 2016 IEEE
DOI 10.1109/ISCMI.2016.27

Unit (IMU) devices attached on the dominate arm wrist, ankle and one on the chest of each participant at a frequency rate of 100 hertz. Researcher in [16] collected and published their dataset called USC-HAD from 14 subjects wearing MotionNode device on their waist, which is connected to the laptop at a frequency rate similar to that in [9]. On the other hand the authors in [10] and [12] found that public dataset is incrementally introduced, but there is still a lack of personalized Smartphone dataset. To crack this problem the researchers in [10] and [12] published datasets collected using Samsung Galaxy S II Smartphone from 30 subjects at a frequency rate of 50 hertz respectively. However, their dataset lacks personal attributes (e.g., age, height, weight, body mass index) and their datasets are in high dimensional space and require large memory and high computational power [13].

Authors in [14] investigated how the selection of frequency sampling rate and classification features affect the energy overheads introduced by high computational power. Their investigation was influenced by the investigation done in [16]. Based on their findings they proposed an activity sensitive strategy called Adaptive Accelerometer-based Activity Recognition (A3R) for continuous activity recognition. A3R selection of sampling rate and classification features adapts in real time as the individuals perform daily activities. They validated their A3R strategy, by annotating human activity dataset collected from 4 selected subjects. To investigate A3R naturalistic lifestyle-driven they employed two additional dataset collected from 6 users for a period of 6-8 weeks. Lastly to evaluate energy savings potential of A3R they monitored 2 additional users who were engaged in their everyday lifestyle activities. However, due to limited dataset employed A3R produced poor results.

Researchers in [7] investigated and found that normal HAR model often learns from accelerometer data collected from one person and distributed to other persons to recognize the same activities instead of generating a different model. They also found that such HAR models are restricted by instability of accuracy especially in cross-people. They proposed and implemented a personalization algorithm that selects confident samples from one person in real time to update existing models of other individuals. However, their technique raised security issues due to usage of generated model from subject *A* being forwarded to subject *B* and *C*. Moreover, the usage of one confident sample data negatively impacts the reliability of their results and leads to poor performance in cross-people prediction.

Most recently and similar to our technique is the work in [15], where authors found that learning new activities to adapt to new users needs is challenging due shortage of annotated dataset. The researchers proposed the Feature-Based and Attribute-Based learning that leverages the relationship between existing and new activities to compensate for the shortage of annotated dataset [15]. They evaluated the technique and found that their technique outperforms other traditional HAR models in recognizing new activities using limited training dataset. However, the technique does not address the shortage of personalized dataset; it only detects new activities from existing dataset. In this paper, we address the shortage of personalized dataset by tailoring bag-of-features dataset for new users (based on similar personal traits such as height, weight and BMI), by taking advantage of publicly available HAR dataset that contains personal attributes such as age, weight and body mass index.

III. Methodology

In this section, we address the shortage of personalized dataset for real-time HAR systems. We first present the system architecture shown in Fig. 1. User A, B and C accesses our personification algorithm deployed on Apache Spark Server using their personal Smartphone via the Internet. Users enter their personal profile attributes (e.g., weight, height, age and gender).

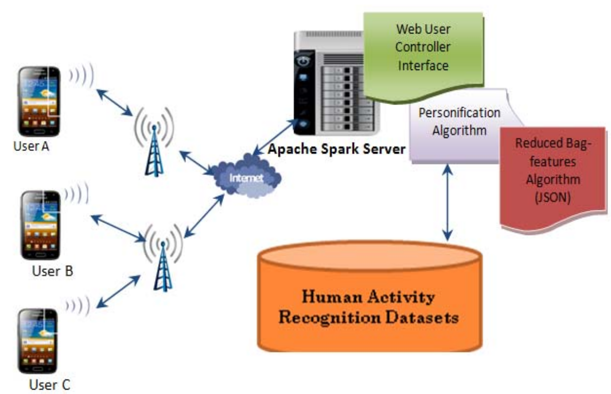

Fig.1. Personification HAR System Architecture

The personification algorithm implemented in Scala will accept and respond to users requests through the Web User Controller Interface (WUCI) developed in Java. The WUCI accepts new user inputs (including weight, height, age, gender) and pass them to the personification algorithm, which computes new user's BMI. Then, the algorithm tailors bag-of-features by matching new user's demographic profile using two publicly available Human Activity Recognition (HAR) PAMAP2 and USC-HAD HAR datasets. Thereafter raw personalized bag-of-features is reduced by extracting and returning lightweight time domain features. We used a lightweight JavaScript Object Notation (JSON) to send reduced time domain features back to new users. Next we present dataset selection process, pre-processing process and the personification algorithm.

A. Dataset Selecton Process

In this process, we selected the PAMAP2 and USC-HAD HAR datasets published in [9] and [17]. Both [9] and [17] are widely used datasets in HAR that have personal attributes such as age, gender, weight and height as recommended in [8]. The dataset published in [9] comprises of 9 multiple subjects' files labeled subject101.dat to subject109.dat. The dataset provided in [17] was collected from 14 subjects performing 5 trials of 12 low-level human activities stored in MATLAB files. The next process details how the selected datasets are pre-processed.

B. Pre-processing

We selected only 5 first trials of USC-HAD dataset for each subject. Then we converted the selected files from MATLAB (.mat) format into Comma Separated Values (CSV) files. Thereafter, we merged all the CSV files into one master file, and computed Body Mass Index (BMI) using equation (1) for each subject.

$$BMI = \frac{\boldsymbol{weight}}{\boldsymbol{height^2}} \tag{1}$$

We reduced attributes of the dataset provided in [17] from 54 to 13 attributes by extracting tri-axial sensor (3D Accelerometer and 3D gyroscope) data, activity number and human activity label. Then for each subject we included subject's number, age, height and weight using the provided *subjectInformation.pdf* file provided in [17]. Thereafter, all 9 subject files were merged into one single master file. Lastly, we computed BMI using equation (1) for each subject in the master file.

C. Personification Algorithms

We present personification algorithm based on K Nearest Neighbor (KNN) algorithm to generate personalized bag-of-features dataset using two pre-processed master files. KNN is more effective and suitable in classifying objects from largest dataset. The proposed personification algorithm 1 computes personification distance between new user's personal traits and subject's personal profile (i.e., using height, weight and BMI) using personification equation (2) based on Euclidean distance.

```
initialize  HashMap personalizedBOFeature Vector
newUser : {weight, height, bmi ← (weigh / (height)^2)}
count ← 0
inputCSV: files : {pre_USCdata, prePAMAP2}
ForEach  inputCSV DO
 ForEach subject_i in InputCSV file Do
   PersonalDist  ← EuclideanDistance (newUser, subject_i)
  IF PersonalDist ≤ Threshold KValue THEN
      svm = sqrt(( subject_i .x^2) + ( subject_i .y^2) + ( subject_i .z^2))
      personalizedBOFeature.put(count, ( subject_i, mag ))
 END IF
 count  ← count + 1
 End do
 End Do
```

Algorithm 1. Personification Algorithm

$$\boldsymbol{p_{user,subjects}} = \sqrt{\sum_{j=1}^{J}(w_j - h_j)^2} \tag{2}$$

Where $\boldsymbol{p_{user,subjects}}$ represents the personification distance between new user and subjects, w_j represent new user demographic attributes (i.e., BMI and weight variables) and h_j denote subject's variables. Authors in [2] and [3] indicated that BMI is correct measure to profile user as normal weight, overweight and obese. Hence, our personification algorithm tailors bag-of-features for new users using weight and BMI. Moreover, we selected both BMI and weight because they change more frequently than human height. The algorithm 1 uses equation (3) to determine closest K neighbours based on the computed personification distance.

$$\boldsymbol{p_{user,subjects}} \leq \boldsymbol{K} = \{\boldsymbol{true}, \boldsymbol{false}\} \tag{3}$$

The computed personification distance is compared with the similarity threshold as given in equation (3). The K value is the estimate threshold measurement used to determine the closest neighbors. All neighbors within the K threshold value are stored into a personalized bag-of-features. Each neighbor values comprise of subject's weight, height, BMI, 3D accelerometer and gyroscope tri-axial (X-axis, Y-axis and Z-axis) sensor data and activity label. The algorithm 1 also computes and includes Signal Magnitude Vector (SMV) using accelerometer tri-axial axes for each stored neighbor using equation (4).

$$SMV = \sqrt{x^2 + y^2 + z^2} \tag{4}$$

Where x, y, and z are the accelerometer tri-axial axes. We selected SMV signals because it is insensitive to device position orientation [19][20].

D. Reduced bag-of-features Algorithm

We removed redundant features using Algorithm 2 to extract lightweight time-domain classification features using SynchronizedDescriptiveStatistics package of Apache Common Math Library (APCML) [21].

```
input personalizedBOF HashMap
intialize reducedPersonalizeBOF ArrayList
ForEach  activityGroup_k in personalizedBOF
   mean = SynchronizedDescriptiveStatistics. Mean(personalizedBOF)
  mode = Mode(personalizedBOF)
  median = SynchronizedDescriptiveStatistics.Median(personalizedBOF)
  min = SynchronizedDescriptiveStatistics.Min(personalizedBOF)
 max=SynchronizedDescriptiveStatistics.Max(personalizedBOF)
  add (mean, median, mode, min, max) into reducedBOF
 End do
 JSON_file = JSON.converter(reducedBOF)
```

Algorithm 2. Reduced of bag-of-features Algorithm

The SynchronizedDescriptiveStatistics package is thread safe and maintains input data in memory. More importantly it has the capability to produce rolling statistics computed for specific "window" [21]. The algorithm extracts the mean, median, mode, minimum and maximum according to SMV at rolling window of 20 rows per grouped labeled activity. That is for each 20 rows of subject records we extract time domain features for each labeled grouped activity (see in Table III). All the extracted time domain features per labeled grouped activity are then stored as reduced personalized bag-of-features. Finally the algorithm converts the reduced bag-of-features into JSON file. However, the generated JSON file is too large and needs to be compressed.

IV. Evaluation of the Algorithm

In this section, we evaluate our implemented personification and reduced bag-of-features algorithms. We used 2 pre-processed PAMAP2 and HSC-HAD dataset samples that were presented in our previous pre-processing process. PAMAP2 master file contains about 2871534 rows of 9 merged subjects' records; whereas the pre-processed HSC-HAD dataset file contains 314100 rows of all 14 subjects. We evaluated our algorithms based on three metrics F-Score using equation (5), precision using equation (6) and recall using equation (7):

$$\textbf{\textit{Fscore}} = \frac{2\,.precison\,.recall}{precision+recall} \qquad (5)$$

$$\textbf{\textit{Precision}} = \frac{True\ Positive}{True\ Positive+False\ Positive} \qquad (6)$$

$$\textbf{\textit{Recall}} = \frac{True\ Positive}{True\ Positive+False\ Negative} \qquad (7)$$

The datasets in [9] and [17] contain various subject profiles portrayed in Tables I and II respectively and set of activities depicted in Table III.

TABLE I. PAMAP2 DATASET SUBJECTS PROFILES [9]

Subject No	*Age*	*Height in CM*	*Weight in KG*	*BMI*
101	27	182	83	25.06
102	25	169	78	27.31
103	31	187	92	26.30
104	24	194	95	25.24
105	26	180	73	22.53
106	26	183	69	20.60
107	23	173	86	27.73
108	32	179	87	27.05
109	31	168	65	23.03

TABLE II. USC-HAD SUBJECTS PROFILES [17]

Subject No	*Age*	*Height in CM*	*Weight in KG*	*BMI*
1	27	164	43	15.99
2	26	185	75	21.91
3	31	169	68	23.81
4	23	168	52	18.42
5	35	170	63	21.8
6	27	164	50	18.59
7	32	160	75	29.3
8	22	180	76	23.46
9	30	171	60	20.52
10	28	170	75	25.95
11	34	165	48	17.63
12	36	170	80	27.68
13	21	178	71	22.41
14	49	166	68	24.68

TABLE III. HUMAN ACTIVITY LABELS

Features	*Source*
• Laying, Sitting, Standing • Walking, Running, Cycling • Nordic-walk • Watching-TV, Computer-work • Upstairs, Downstairs • Vacuum clean, Ironing, Folding-laundry, House cleaning • Playing soccer, Rope jumping • Other not labelled	[9]
• Walk forward, Walk left, Walk right • Upstairs, Downstairs, Run forward • Jump up and down • Stand, Sleep/laying	[17]

We randomly selected two users' profile (A and B) with different weight, height and BMI attributes as shown in Table IV and K threshold variables depicted in Table V to determine the robustness of our algorithm. For each user's profile we ran three trials of experiments for each K threshold values.

TABLE IV. NEW USER ATTRIBUTES PROFILES

User	*Height in CM*	*Weight in KG*	*BMI*
A	175	68	22.20
B	169	70	24.50

TABLE V. K VARIABLE THRESHOLD

K Value	*Threshold*
V 1	2.0
V 2	2.5
V 3	3.0
V 4	3.5
V 5	4.0
V 6	4.5

V. Results

The personification algorithm was evaluated using 2 widely used HAR dataset USC-HAD and PAMAP2. We ran 3 trial sets of experiments for each *K* threshold variables. The personification algorithm consistently produced the same results as summarised in Table VI and Table VII. As we change the *k* threshold value, the number of subjects meeting the Euclidean personification distance also increased, because frequent examples tend to dominate the classification due to majority voting of KNN algorithm using un-weighted distance functions. The algorithm for each *K* threshold managed to extract different number of raw bag-of-features and reduced to only relevant light-weight time domain features indicated by the TP (True Positive) in both tables. FP indicates False Positive activities and whilst FN represents False Negative activities. Both Tables VI and VII are summarised in Fig. 2 and Fig. 3 representing three metrics F-Score, precision and recall.

TABLE VI. USER A PERSONALIZED BAG-OF-FEATURES

K	*Subject*	*Euclidean Distance*	*Total rows*	*Reduced Rows*	*Total Activities*	*TP*	*TN*	*FN*
V1	3, 106	1.6,1.9	375874	12780	28	15	5	8
V2	3,14,106	1.6,2.4,1.9	396275	13450	38	20	10	8
V3	3, 14,106	1.6, 2.4,1.9,	396274	13450	38	20	10	8
V4	3,13,14,106,109	1.6,3.0,2.4,1.8,3.1	434457	14685	66	26	10	24
V5	3,13,14,106,109	1.6,3.0,2.4,1.8,3.1	434456	14685	66	26	10	24
V6	3,13,14,106,109	1.6,3.0,2.4,1.8,3.1	434456	14685	66	26	10	24

TABLE VII. USER-B PERSONALIZED BAG-OF-FEATURES

K	*Subject*	*Euclidean Distance*	*Total Rows*	*Reduced Rows*	*Total Activities*	*TP*	*TN*	*FN*
V1	0	0	0	0	0	0	0	0
V2	3,13,14	2.1,2.3,2.0	64600	1905	30	13	8	9
V3	3,13,14	2.1, 2.3, 2.0	64600	1905	30	13	8	9
V4	3,13,14	2.1, 2.3, 2.0	64600	1905	30	13	8	9
V5	3,13,14, 105	2.1, 2.3, 2.0, 3.5	439071	15485	48	25	9	14
V6	3,13,14, 105,106	2.1, 2.3, 2.0, 3.5, 4.0	800545	27970	66	37	10	19

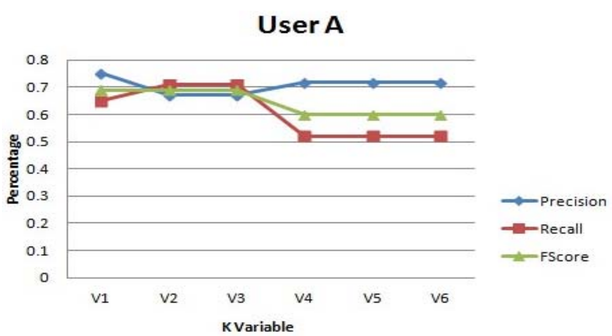

Fig.2. User A performance matrix

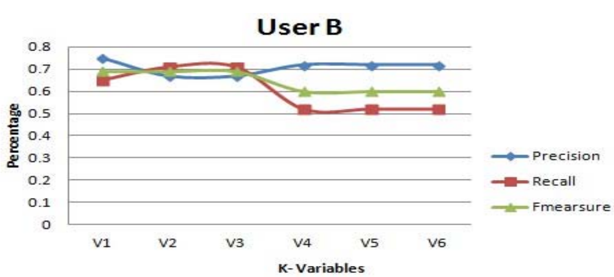

Fig.3. User B performance matrix

On both user A and user B, the personification algorithm produced overall precision of above 0.70% and recall of below 0.60% and balanced measure of above 0.60%. The results indicate that it is possible to extract personalized bag-of-features dataset from the diverse related dataset using personal attributes. We are therefore confident that the algorithm produces satisfactory results precision is above 0.70%.

VI. CONCLUSION AND FUTURE WORK

In this paper, we proposed personification algorithm based on *K* Nearest Neighbor algorithm implemented in Scala and Java to extract personalized bag-of-features dataset (comprising of 3D sensor data and human activity label) to compensate for a shortage of personalized dataset. The personification algorithm uses user's personal attributes i.e., weight, height and body mass index to sift through public dataset to establish and extract related records based on computed personification distance. The results of the experiments are promising and indicative that limited publicly available dataset can be exploited and leveraged to compensate for a shortage of personalized dataset. In future, we intend to test our personification algorithm to generate personalized bag-of-features dataset to train and test classification algorithm in real time; to determine whether we can effectively identify new user's daily human activities on resource constrained Smartphone. We will enhance our KNN personification by employing weighted distance functions. Lastly we intend to compare our K Nearest Neighbor algorithm with other machine learning algorithms.

ACKNOWLEDGMENT

I would like to thank the Faculty of Information Communication Technology and department of Computer

Science, Tshwane University of Technology for financial and other logistical support for the success this project. This project is ethically approved by the Faculty Committee of Research Ethics of Tshwane University of Technology (FCRE ICT Ref#2016=06=001(2)=GadebeML).

REFERENCES

[1] H. Ellekjær, J. Holmen, E. Ellekjær, E. and L. Vatten. “Physical activity and stroke mortality in women ten-year follow-up of the Nord-Trøndelag Health Survey, 1984–1986”. Stroke vol. 31, pp. 14-18, 2000.

[2] F.B. Hu, M.J. Stampfer, G.A. Colditz, A. Ascherio, K.M. Rexrode, W.C. Willett and J.E. Manson,“Physical activity and risk of stroke in women”. Jama vol. 283(22), pp. 2961-2967, 2000.

[3] I.M. Lee, K.M. Rexrode,. N.R. Cook, J.E. Manson and J.E. Buring, , “Physical activity and coronary heart disease in women: Is no pain, no gain passé?”. Jama vol. 285, pp. 1447-1454, 2001.

[4] S. Zhang, “Smartphone Based Activity Recognition System ”. The Ohio State University (Doctoral dissertation) 2001.).

[5] J.W. Lockhart, T. Pulickal and G.M. Weiss, “Applications of mobile activity recognition”. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, pp. 1054-1058, September 2012.

[6] X. Su, H. Tong and P. Ji, . “Activity recognition with smartphone sensors”. Tsinghua Science and Technology, vol. 19, pp.235-249, 2014.

[7] Q.V. Vo, M.T. Hoang and D. Choi. “Personalization in mobile activity recognition system using-medoids clustering algorithm”. International Journal of Distributed Sensor Networks, 2013.

[8] J.W. Lockhart and G.M. Weiss, “Limitations with activity recognition methodology & data sets”. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, pp. 747-756, September 2014.

[9] A. Reiss and D. Stricker, “Introducing a new benchmarked dataset for activity monitoring”. In IEEE 16th International Symposium on Wearable Computers, pp. 108-109, June 2012

[10] D. Anguita, A. Ghio, L. Oneto, X. Parra and J.L. Reyes-Ortiz, “A Public Domain Dataset for Human Activity Recognition using Smartphones”. In ESANN, April 2013.

[11] O. Banos, R. Garcia, J.A. Holgado-Terriza, M. Damas, H. Pomares, I. Rojas, A. Saez and C. Villalonga, “mHealthDroid: a novel framework for agile development of mobile health applications”. In International Workshop on Ambient Assisted Living, pp. 91-98, December 2014. Springer International Publishing.

[12] J.L. Reyes-Ortiz, L. Oneto, A. Samà, Parra, X. and D. Anguita, “Transition-aware human activity recognition using smartphones”. Neurocomputing, vol. 171, pp.754-767 2016.

[13] C.A. Martins, M.C. Monard and E.T. Matsubara, “Reducing the dimensionality of bag-of-words text representation used by learning algorithms”. In Proceedings of The Third IASTED International Conference on Artificial Intelligence and Applications (AIA 2003), Benalmádena, Espanha vol. 38, 2003 in press.

[14] Z. Yan, V. Subbaraju, D. Chakraborty, A. Misra and K. Aberer, “Energy-efficient continuous activity recognition on mobile phones: An activity-adaptive approach”. In IEEE 2012 16th International symposium on wearable computers,pp. 17-24.

[15] L.T. Nguyen, M. Zeng, P. Tague and J. Zhang, “Recognizing new activities with limited training data”. In ACM Proceedings of the 2015 ACM International Symposium on Wearable Computers, pp. 67-74, September 2015.

[16] H. Junker, P. Lukowicz and G. Troster, “Sampling frequency, signal resolution and the accuracy of wearable context recognition systems”. In Wearable Computers, 2004. ISWC 2004. IEEE Eighth International Symposium on, vol. 1, pp. 176-177, October 2004.

[17] M. Zhang and A.A. Sawchuk, 2012, September. USC-HAD: a daily activity dataset for ubiquitous activity recognition using wearable sensors. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, pp. 1036-1043, October 2004.

[18] D. Figo, P.C. Diniz, D.R. Ferreira and J.M. Cardoso, “Preprocessing techniques for context recognition from accelerometer data”. Personal and Ubiquitous Computing, vol.14, pp.645-662, 2010..

[19] F. Miao, Y. He, J. Liu, Y. Li and I. Ayoola, “Identifying typical physical activity on smartphone with varying positions and orientations”.Biomedical engineering online, vol. 14, p.1 , 2015.

[20] A. Mannini, S.S. Intille, M. Rosenberger, A.M. Sabatini and W. Haskell, “Activity recognition using a single accelerometer placed at the wrist or ankle”. Medicine and science in sports and exercise, vol.45, p.2193, 2013.

[21] Apache Common Math. 2016. Available at: http://commons.apache.org/proper/commons-math/userguide/stat.html