# Analysis of Motion Patterns for Recognition of Human Activities

Majid Janidarmian[1], Atena Roshan Fekr[1], Katarzyna Radecka[1], Zeljko Zilic[1], Louis Ross[2]
[1]Electrical and Computer Engineering Department, McGill University, Montreal, Quebec, Canada
[2]Motion Engine Company, Montreal, Quebec, Canada
{majid.janidarmian, atena.roshanfekr}@mail.mcgill.ca,
{katarzyna.radecka, zeljko.zilic}@mcgill.ca, ross@motionengineinc.com

## ABSTRACT
Automatic recognition of human activity is one of the most important and challenging open areas of research in context-aware and physical training applications. The activity profiling systems normally use wearable sensors to record motion patterns over extended periods of time. The performance of these systems depends on the activity set, training data quality, extracted features and learning algorithms. In this paper, we describe the development of efficient activity recognition techniques using wearable motion sensors. An extensive evaluation is provided on three well-known classifiers with light-weight time series features to distinguish among thirty three different fitness activities. The effects of the segmentation, sensor placement as well as different sampling rates on the classification performance are discussed. The experimental results conducted with 17 subjects show an improvement of the classification accuracy compared with the previous work. In addition, the statistical analysis of the investigated models quantifies relative effects of the accelerometer axes reductions on the classification performance.

## Categories and Subject Descriptors
J.3 [Computer Applications]: Life & medical sciences

## General Terms
Algorithms, Measurement, Experimentation

## Keywords
Wearable sensors, activity recognition, machine learning, classification, performance optimization

## 1. INTRODUCTION
The maturity of pervasive sensing, wireless technology and data processing techniques enables us to provide an effective solution for continuous monitoring and promote individual's health. Today, the miniature sensors can be unobtrusively attached to the body or can be part of clothing items to observe people's lifestyle and behavior changes [1]-[3]. According to study presented in [4], on-body sensing proves to be the most prevalent monitoring technology for the gait assessment, fall detection and activity classification. Therefore, extensive research have been undertaken to select or develop reasoning algorithms to infer activities from the wearable sensor data. Sensor-based activity recognition which targets the automatic detection of people's activities is one of the promising research topics in different areas such as ubiquitous computing and ambient assistive living [5]. Accelerometer sensor is the most broadly used wearable sensor for activity recognition and could provide high classification accuracy of 92.25% [6], 96.2% [7], and 98% [8]. There is a large amount of work on the use of accelerometers for activity monitoring and behavior profiling. For example, there are surveys [9]-[11] which provide an outline of relevant research and applicable techniques. In fact, the performance of recognition models mainly depends on the activity set, training data quality, extracted features and learning algorithms. Overall, most systems provide similar accuracy levels, but since each system works with a specific dataset and activity set, there is no significant evidence to claim that a system is more accurate than the others.

Recently, one of the most complete activity recognition benchmark datasets has been published, and is now publicly available to the research community [12]. In this paper, we make use of this dataset to not only evaluate different classification techniques for activity recognition, but also to provide an analysis on optimizing the performance and complexity of the final model.

The rest of the paper is organized as follow: In section 2, the proposed feature extraction and classification techniques with segmentation methods are described. Experimental results are presented and discussed in details in section 3. Finally, section 4 concludes the paper.

## 2. METHODS
Activity recognition means identifying the actions of one or more individuals using a series of observations and environmental conditions [13] as formulated as follows:

*Definition:* With $l$ extracted features from the wearable sensors, given a set $W = \{w_1, w_2, \dots, w_m\}$ of labeled and equal-sized time windows, and a set $A = \{a_1, a_2, \dots, a_n\}$ of activity labels, the goal is to find the best classifier model $C$, such that for any $w_k$ which contains a feature set $F_k = \left\{f_{k,1}, f_{k,2}, \dots, f_{k,l}\right\}$, the predicted label $\hat{a}_k = C(F_k)$ is as identical as possible with the actual activity performed during $w_k$.

The recognition procedure starts with collecting data from the motion sensors. In this work, two window-based segmentation methods are applied to investigate the effectiveness of the classifiers with respect to the window length and the percentage of adjacent windows overlap. These techniques provide low complex implementations and reasonable performance which will be evaluated in details in section 3. The first popular method is dividing the data stream into windows of fixed length with no inter-window gaps and no degree of overlap between adjacent windows (Fixed-size Non-overlapping Sliding Window (FNSW)) [11]. The second method is Fixed size Overlapping Sliding Window (FOSW)

MOBIHEALTH 2015, October 14-16, London, Great Britain

DOI 10.4108/eai.14-10-2015.2261719

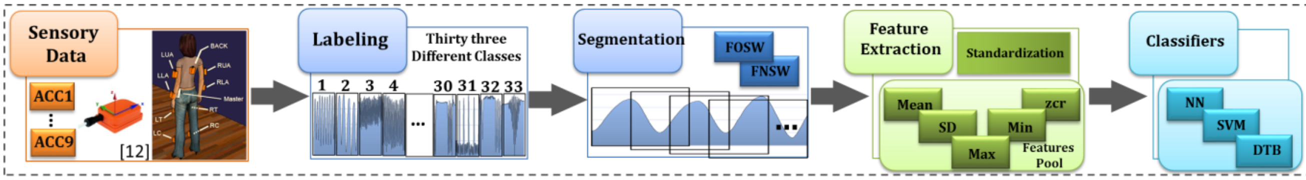


**Figure 1. Activity classification procedure**

which is similar to FNSW except that the windows overlap during segmentation [14].

It is noteworthy that, the tradeoff between overlap value and total number of processing windows should be considered by recognition system designers. Subsequently, a feature extraction process is carried out to extract the statistically important characteristics of the sensor data used to differentiate diverse classes or states. The choice of features with high information content and low complexity for classification purpose is a fundamental phase and a highly problem-dependent task [15]. The time domain features used in this paper are simple, easy to calculate and properly interpret the activity parameters, as proved in [10]. These features include mean, standard deviation, maximum, minimum and zero crossing rate. These are some of the features most widely used in activity recognition such as [16]-[18]. In order to keep the experimental setup and classifiers performance comparable with [10], no feature selection has been applied in this work.

The features are extracted from each separate window of the data and then used as inputs to the classifiers. Likewise, the evaluations are provided on three well-known machine learning techniques including Neural Network (NN), Support Vector Machine (SVM) and Decision Tree Bagging (DTB). These classifiers have been verified to provide significant outcomes while using sensory data [19]-[21].

## 2.1 Classification Algorithms

In this paper, we are dealing with the supervised methods which use the class label when discretizing features. Thus, after data segmentation, the features are extracted from each separate window of the data and then used as inputs to the classifiers. We choose three classifiers to discriminate among thirty three types of activity. Support Vector Machines (SVM) and Neural Networks (NN) have been broadly used in human activity recognition. SVM relies on finding optimal separating decision hyperplanes between classes with the maximum margin between patterns of each class. The key advantage of this classifier is the ability of minimizing both structural and empirical risks [22]. These properties make SVM to be a strong generalization for new data classification even in case of limited training dataset [20]. In this paper, we make use of SVM with Error Correcting Output Codes (ECOC) with one-versus-one (OVO) coding design in our multi-class classification. The idea of using ECOC [23] is to break the multiclass task into several binary classification tasks and then combining the results of these classifiers to obtain the final outcome. Since SVM is very popular in binary classification, we choose to use ECOC for combining $k(k-1)/2$ multiple binary SVMs with linear kernel function [24] where $k$ refers to the number of classes and is thirty three in the proposed classifier.

Artificial neural networks provide a robust tool which help people to analyze, model and make sense of big data across a wide range of applications [25]. NN inspired from simulation of biological nervous system in the human brain which is based on propagating activation signals and encoding knowledge in the network links. The basic neural network architecture has three layers including: input, hidden and output layers. The data is propagated through successive layers, and the final result is available at the output layer. Multilayer perceptron neural network (MLP) uses more than one hidden layer in its structure which might help in solving complex problems where a single hidden layer cannot provide an acceptable result [26]. We have used 20 hidden layers in our experimental results.

Among different methodologies to combine classification models, ensembles of decision trees are described as the most accepted approaches [27]. DTB builds an ensemble of classification trees (in our tests, 50) and uses bagging to combine the predictions [28]. In general, classifier ensembles are clearly more expensive, computationally speaking, as they require several models to be trained and evaluated.

In the next section, we compare the results of the described classifiers based on the mentioned feature extraction for distinguishing among 33 different activities with nine accelerometer sensors. Fig. 1 depicts the whole system flow of the illustrated activity recognition.

# 3. EXPERIMENTAL RESULTS

## 3.1 Data Collection and Setup

3-axis accelerometer is the most widely used sensor to recognize fitness and ambulation activities due to its compact size, low-power requirement and capacity to provide data directly related to the kinematics of subject. Therefore, in this paper, we have used accelerometer sensors data which recently released in one of the most complete activity recognition benchmark datasets [12][29]. The dataset comprises the readings of nine Xsens MTx [30] motion sensors recorded from seven females and ten males, with ages ranging from 22 to 37 years old while performing 33 fitness activities. The sensors are mounted on different parts of the body including: Left Calf (LC), Left Thigh (LT), Right Calf (RC), Right Thigh (RT), back (BACK), Left Lower Arm (LLA), Left Upper Arm (LUA), Right Lower Arm (RLA) and Right Upper Arm (RUA). The sensory nodes are able to measure the motions experienced by each body limb and trunk, thus better capturing the body dynamics. The results demonstrate the possibility of acquiring precise activity recognition from a single type of sensor while giving more credits to accelerometer-based approaches [31] as a simple and cost effective solution for activity detection applications.

## 3.2 Data Segmentation Analysis

In this section, we analyze the general effects of the windowing operation on the activity recognition problem. The evaluation of the activity classification is performed first through a 10-fold random-partitioning cross-validation process. Then, a Leave-One-Subject-Out cross validation is evaluated in which a single subject is iteratively left out from the training dataset and considered in the test set. The procedure is then repeated for all subjects whose data are complete in the dataset. In fact, as summarized in [32] and

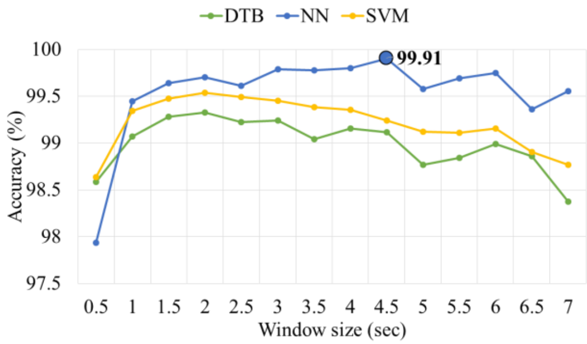


**Figure 2. Accuracy rates for FNSW segmentation with nine sensors**

according to [33], Leave-One-Subject-Out is the best technique for risk estimation, whereas 10-fold is the most accurate approach for model selection.

The accuracy rate is used as the performance measure, which is defined as the proportion of correct classifications with respect to the total classified instances as follows.

$$CM = \begin{bmatrix} M_{11} & \cdots & M_{1l} \\ \vdots & \ddots & \vdots \\ M_{l1} & \cdots & M_{ll} \end{bmatrix} \tag{1}$$

$$Accuracy = \frac{\sum_{i=1}^{l} CM(i,i)}{\sum_{i=1}^{l}\sum_{j=1}^{l} CM(i,j)} \times 100 \tag{2}$$

Each row in confusion matrix $CM$, indicates the instances in a true class, while each column represents the instances in a predicted class. The results for diverse window sizes and each specific methodology with FNSW are depicted in Fig. 2.

Considering all nine sensors data, the NN model shows the best performance with accuracy above 99% in all window sizes swept from 1s to 7s. Increasing the window size more than 2s causes a worsening of the recognition performance for both DTB and SVM. The minimum accuracy values for NN and SVM are obtained in 0.5s, while the accuracies increase when the window is enlarged to 1s. The results demonstrate that we could improve the best accuracy by about 2% for identifying 33 activities compared to the best derived classification model (k-nearest neighbors) in [10].

The results are also evaluated based on FOSW segmentation with five different overlap values including 10%, 25%, 50%, 75% and 90%. The best accuracy values of 99.98%, 99.96% and 99.95% are achieved with NN, SVM and DTB through window sizes 4.5s and 1.5s with 90% overlap. The accuracies decrease in average by 9.45% and 7.73% for FNSW and FOSW (90%) while considering Leave-One-Subject-Out cross validation.

In order to keep the outcomes comparable with [10] and [12], the classifiers accuracies are reported based on 10-fold random-partitioning cross-validation in the next sections.

## 3.3 Sensors Placements Evaluation

Wearing nine IMUs is uncomfortable, invasive, expensive, and not suitable for continuous monitoring of daily activities. Therefore, we will consider each accelerometer data, separately to see how well a single sensor can recognize the activity set. Fig. 3 visualizes the discrimination potential of each sensor for 33 fitness activities, individually. The assessments are computed based on both FNSW and FOSW segmentations in Fig. 3 (a) and (b), respectively. It is observed that, with FNSW, the best accuracy is obtained with DTB in window size 2.5s using the accelerometer on the left tight (LT) of the subjects. NN provides the second best accuracy of 88.96 with sensor on the right tight (RT) and window size 4s. Fig. 3 (b) shows the accuracy values of each sensor with FOSW and overlap 90%. An average improvement of 8% is observed for the case of using fixed size overlapping sliding window. The DTB model is shown to be quite robust considering its moderate performance drop by reducing the number of sensors to one (from 99.95% to 99.60%). All best cases of classification accuracies are derived from the accelerometer on the left tight. Among different parts of the body, the accelerometer on the back overcomes the performance obtained from the sensors on the arms, whereas the lower body (tights and calves) provides the best performance compared to upper body

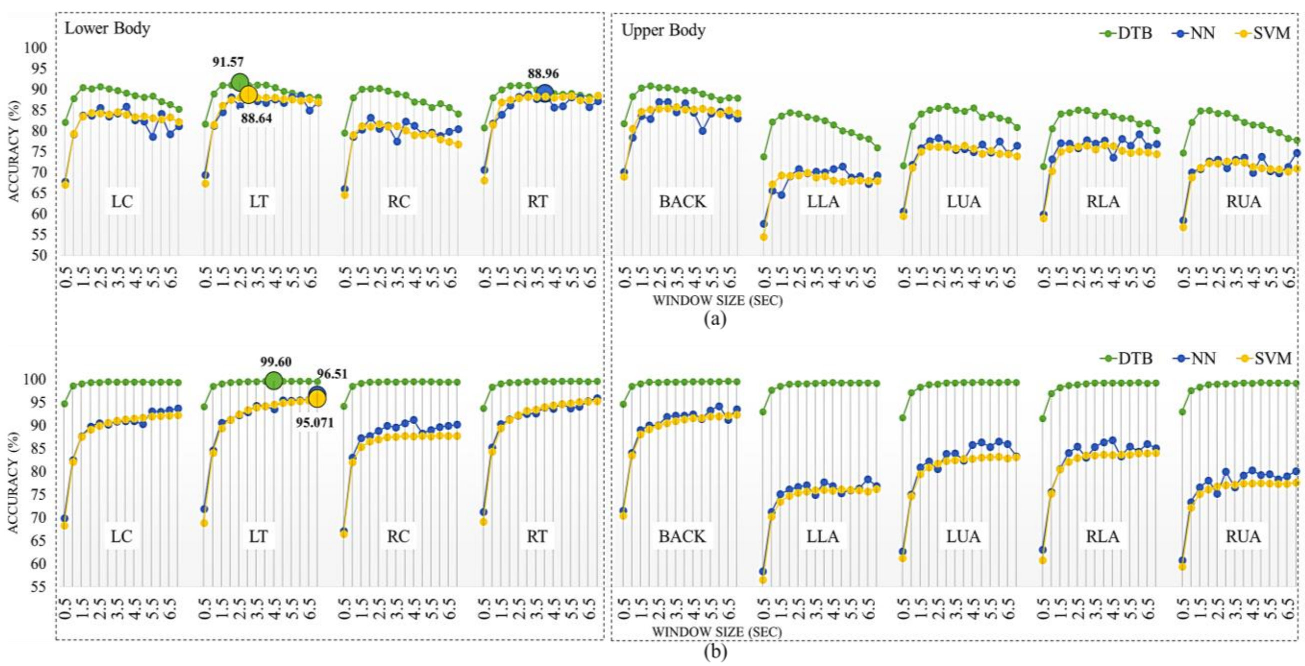


**Figure 3. The classification accuracy based on each sensor on different locations using (a) FNSW, and (b) FOSW (90%)**

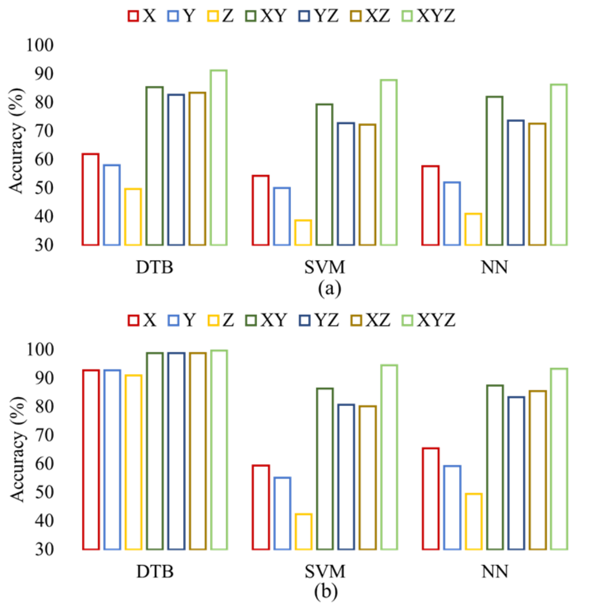


**Figure 4. The impact of accelerometer axes reduction on the classification accuracy using (a) FNSW and (b) FOSW (90%)**

(arms and back). The single sensor classification evaluation across all subjects and sensors shows an improvement of 7.19% with FNSW compared to [10][12] while this number increases to approximately 20% using FOSW technique.

## 3.4 Accelerometer Axes Reduction

In this section, the impact of acceleration of different dimensions (X, Y, and Z) on the best extracted model (single accelerometer on left tight) is investigated. Reducing the number of accelerometer axes translates into a faster detection at the expense of using less data for the features computation. Fig. 4 (a) and (b) show the accuracy values of three classifiers while a single, double and triple axes of accelerometer are considered for FNSW and FOSW (90%), respectively. Decreasing the number of axes can be interpreted as a feature selection technique to reduce computations and simplify the learning models. It is worth noting that, decreasing the number of axes by one represents a reduction of features by more than 33%. Therefore, these figures are devised as a perfect means to visually illustrate the trade-off between performance and the number of axes. In addition, considering Fig. 4 (b), DTB can still provide more than 91% accuracy with a single axis while this number increases to 98% with two axes. Consequently, despite a small drop in accuracy (about 7%), we could reduce the redundancy among features by more than 66% for the single-axis condition.

## 3.5 Sampling Rate Analysis

Another important point of discussion is how to reduce computations, storage, and energy consumption by means of reducing sampling rate while implementing activity recognition. It should be carefully chosen to guarantee a reasonable response time and battery life while keeping the accuracy sufficiently high.

In our experiment, we have swept the sampling rate (by resampling technique) from 1Hz to 50Hz for the best model obtained from a single accelerometer on the LT of the subjects. Fig. 5 displays the behavior of the recognition accuracy as a function of the accelerometer sampling rate in both FNSW and FOSW (90%) segmentations and all three classifiers. Interestingly, we found that no significant gain in accuracy is achieved for sampling rate above 7 Hz. Therefore, it is concluded that low frequency sampling of accelerometer data can lead to classification results competitive with previous results for much higher sampling rates.

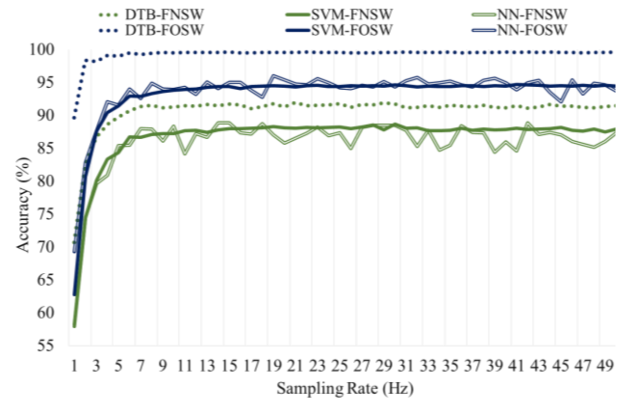


**Figure 5. The impact of sampling rate on the classification accuracy using FNSW and FOSW (90%)**

## 4. Conclusion

With the growth of sensor technology and the analysis methods, recognition systems based on wearable sensors carry the advantages of simple setup, high reliability and accuracy as well as providing useful information for physical training applications. Being able to recognize the state of a person can also provide us valuable information to have a better understanding of numerous medical conditions and treatments. In this paper, we exploited recent advances in wearable sensing, signal processing and machine learning principles to provide an accurate activity recognition system. The results demonstrated the possibility of acquiring a precise model with accelerometer sensors. Comparing with previous work, about 2% and 7.19% improvements were achieved for identifying 33 activities, considering all sensors and single sensor, respectively. We also could achieve a great increase of 20% through investigations on the degree of overlap between windows in the segmentation phase. Such a reliable recognition model could encourage people to do more outdoor activities and also be an assessment of activities of daily living for chronic treatment. The results of reducing the number of axes in accelerometer sensor as well as sampling rate have been also discussed in this paper. For instance, we concluded that, even low sampling rates of accelerometer data collected from the left tight location can still provide promising results.

## 5. REFERENCES

[1] S. Kumar, W. Nilsen, M. Pavel, and M. Srivastava, "Mobile health: Revolutionizing healthcare through transdisciplinary research." IEEE Computer, vol. 46, no. 1, pp. 28-35, 2013.

[2] A. Roshan Fekr, M. Janidarmian, K. Radecka, Z. Zilic, "Movement analysis of the chest compartments and a real-time quality feedback during breathing therapy" Network Modeling Analysis in Health Informatics and Bioinformatics, 4:21, 2015.

[3] A. Roshan Fekr, M. Janidarmian, K. Radecka, Z. Zilic, "Tidal volume variability and respiration rate estimation using a wearable accelerometer sensor," in Wireless Mobile Communication and Healthcare (Mobihealth), pp.1-6, Nov. 2014

[4] R. Khusainov, D. Azzi, I. Achumba, and S. Bersch, "Real-Time Human Ambulation, Activity, and Physiological

Monitoring: Taxonomy of Issues, Techniques, Applications, Challenges and Limitations," Sensors, vol. 13, no. 10, pp. 12852–12902, Sep. 2013.

[5] L. Chen, J. Hoey, C.D. Nugent, D.J. Cook, Z. Yu, "Sensor-Based Activity Recognition," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol.42, no.6, pp.790-808, Nov. 2012.

[6] Z. He, L. Jin, "Activity recognition from acceleration data using AR model representation and SVM," Machine Learning and Cybernetics, 2008 International Conference on, vol.4, no., pp.2245-2250, July 2008.

[7] Y. Nam and J. W. Park, "Physical activity recognition using a single triaxial accelerometer and a barometric sensor for baby and child care in a home environment," Journal of Ambient Intelligence and Smart Environments, vol. 5, no. 4, pp. 381–402, 2013.

[8] A.M. Khan, Y. Lee, S.Y. Lee, T. Kim, "A Triaxial Accelerometer-Based Physical-Activity Recognition via Augmented-Signal Features and a Hierarchical Recognizer," Information Technology in Biomedicine, IEEE Transactions on , vol.14, no.5, pp.1166-1172, Sept. 2010.

[9] O.D. Lara, M.A. Labrador, "A Survey on Human Activity Recognition using Wearable Sensors," Communications Surveys & Tutorials, IEEE, vol.15, no.3, pp.1192-1209, 2013.

[10] O. Banos, J.-M. Galvez, M. Damas, H. Pomares, and I. Rojas, "Window Size Impact in Human Activity Recognition," Sensors, vol. 14, no. 4, pp. 6474–6499, Apr. 2014.

[11] S.J. Preece, J.Y. Goulermas, L.P. Kenney, D. Howard, K. Meijer, R. Crompton, "Activity identification using body-mounted sensors—A review of classification techniques" Physiological Measurement, vol. 30, no. 4, 2009.

[12] O. Banos, M.A. Toth, M. Damas, H. Pomares, I. Rojas, "Dealing with the effects of sensor displacement in wearable activity recognition," Sensors vol. 14, no. 6, pp. 9995-10023 2014.

[13] W. Liu, X. Li, and D. Huang, "A survey on context awareness," in Proceedings of the International Conference on Computer Science and Service System (CSSS '11), pp. 144–147, June 2011.

[14] E. Keogh, S. Chu, D. Hart, M. Pazzani, "An Online Algorithm for Segmenting Time Series," In Proceedings of the International Conference on Data Mining, pp. 289–296, December 2001.

[15] A. Mannini, and A.M. Sabatini,"Machine learning methods for classifying human physical activity from on body accelerometers," Sensors, vol.10, pp.1154-1175, 2010.

[16] N. Kern, B. Schiele, A. Schmidt, "Multi-sensor activity context detection for wearable computing," In Proceedings of the First European Symposium on Ambient Intelligence (EUSAI), Veldhoven, The Netherlands, pp.220–232, November 2003.

[17] N. Ravi, P. Mysore, M.L. Littman, "Activity recognition from accelerometer data," In Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence, Pittsburgh, PA, USA, pp. 1541–1546, July 2005.

[18] J.R. Kwapisz, G.M. Weiss, S.A. Moore, "Activity recognition using cell phone accelerometers.," In Proceedings of the 17th Conference on Knowledge Discovery and Data Mining, San Diego, CA, vol. 12, pp.74–82, USA, August 2011.

[19] M. Janidarmian, K. Radecka, Z. Zilic, "Automated diagnosis of knee pathology using sensory data," Wireless Mobile Communication and Healthcare (Mobihealth), pp.95-98, 3-5 Nov. 2014.

[20] A. Roshan Fekr, M. Janidarmian, K. Radecka, Z. Zilic, "A medical cloud-based platform for respiration rate measurement and hierarchical classification of breath disorders," Sensors 2014, vol.14, pp.11204-11224.

[21] A. Roshan Fekr, M. Janidarmian, K. Radecka, Z. Zilic, "Respiration Disorders Classification with Informative Features for m-health Applications," in Biomedical and Health Informatics, IEEE Journal of , vol.PP, no.99, pp.1-1

[22] O. Chapelle, P. Haffner, and V. N. Vapnik, "Support vector machines for histogram-based classification," IEEE Trans. Neural Netw., vol. 10, no. 5, pp. 1055–1064, Sep. 1999.

[23] T. G. Dietterich and G. Bakiri, "Solving multiclass learning problem via error-correcting output codes," Journal of Artificial Intelligence Research, pp. 263–286, 1995.

[24] Y. Liu, "Using SVM and error-correcting codes for multiclass dialog act classification in meeting corpus," in Proc. Interspeech-ICSLP, pp. 1938-1941, Sep. 2006.

[25] Y. K. Irfan, P. H. Zope, and S. R. Suralkar, "Importance of Artificial Neural Network in Medical Diagnosis disease like acute nephritis disease and heart disease," International Journal of Engineering Science and Innovative Technology (IJESIT), vol 2(2), pp. 210-217, 2013.

[26] Kaushal Kumar & Abhishek, "Artificial Neural Networks for Diagnosis of Kidney stone Disease", I.J. Information technology and computer science, vol.7, pp. 20-25, July 2012.

[27] J. Abellán and A. R. Masegosa, "Bagging decision trees on data sets with classification noise," in Proc. 6th Int. Conf. Found. Inf. Knowl. Syst., Sofia, Bulgaria, pp. 248–265, Feb. 2010.

[28] Z.A. Pardos, N.T. Heffernan, "Using HMMs and bagged decision trees to leverage rich features of user and skill from an intelligent tutoring system dataset," In Proceedings of the KDD cup workshop, pp. 24–35, 2010.

[29] O. Banos, M.A. Toth, M. Damas, H. Pomares, I. Rojas, O. Amft, "A benchmark dataset to evaluate sensor displacement in activity recognition," Proceedings of the 14th International Conference on Ubiquitous Computing, Pittsburgh, USA, September 2012.

[30] Xsens Technologies B.V. XM-B Technical Documentation, 2009. Available online: http://www.xsens.com (accessed on 12 May 2014).

[31] A. Roshan Fekr, K. Radecka, Z. Zilic, "Design and Evaluation of an Intelligent Remote Tidal Volume Variability Monitoring System in e-Health Applications," Biomedical and Health Informatics, IEEE Journal of, vol. 19, no. 5, pp. 1532 - 1548, 2015.

[32] S. Arlot, A. Celisse, "A survey of cross-validation procedures for model selection," Stat. Surv., vol.4, pp.40–79, 2010.

[33] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," In Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI'95), vol.2, pp. 1137-1143, 1995.