15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-NCON2018)

Smartwatch-based Sitting Detection with Human Activity Recognition for Office Workers Syndrome

Sakorn Mekruksavanich*, Narit Hnoohom† and Anuchit Jitpattanakul‡
*Department of Computer Engineering, School of Information and Communication Technology
University of Phayao, Phayao, Thailand
Email: sakorn.me@up.ac.th
†Image, Information and Intelligence Laboratory, Department of Computer Engineering
Faculty of Engineering, Mahidol University, Nakhon Pathom, Thailand
Email: narit.hno@mahidol.ac.th
‡Department of Mathematics, Faculty of Applied Science
King Mongkuts University of Technology North Bangkok, Bangkok, Thailand
Email: anuchit.j@sci.kmutnb.ac.th

***Abstract*—The main cause of Office Workers Syndrome (OWS) is habitual unhealthy behavior. One such example is the habit of siting for long periods at a computer. It has recently become simple to study this kind of human activity by using smartwatch sensors, and as a result, HAR (human activity recognition) has become a field attracting considerable interest. Smartwatch accelerometers have been especially useful in HAR, while gyroscopes are also helpful, especially when combined with the accelerometer. This study therefore suggests the use of HAR to detect sitting in order to identify the threat of OWS through the use of data from the accelerometer and gyroscope in a smartwatch. Two ensemble learning approaches have therefore been examined in order to establish their capability in recognizing periods of sitting. The results reveal that the accelerometer and gyroscope are complementary in their application and using them in combination thus improves the accuracy of recognition. The ensemble learning-based techniques were able to achieve an accuracy level of 93.57% for activity recognition when detecting sitting.**

Keywords*—*smartwatch; activity recognition; office workers syndrome; ensemble learning

I. Introduction

The International Labor Organization (ILO) (2016) recently reported that among countries where people work very long hours (47% exceed 48 hours per week), Thailand is listed in third place. The hours worked by many Thai people mean that the time they spend at work exceeds the time they spend at home [1]. While OWS is not a medical condition with a specific diagnosis, the symptoms are commonly observed in office workers who spend long durations in unhealthy conditions where their muscles are used incorrectly [2]. The most common problem stems for poor posture when working with computers for long periods. Sitting in one position leads to back and shoulder pain, headaches, numbness in the hands and arms, worsening eyesight and often dry eyes. The pain experienced may at first seem trivial, but can become chronic over time, and may lead to long-term problems with the vertebrae of the spine, causing abnormalities which can adversely affect peoples quality of life as they are increasingly unable to move easily and face long-term pain. In some cases, emotional symptoms can be observed, as workers with OWS begin to feel excessively tired or start to suffer depression.

The Internet of Things (IoT) is one possibility for finding relief since it offers numerous applications, including health care. The IoT makes it possible for a patient to be monitored remotely for health status, and can assist with fitness programs [3]. It can work effectively by using activity tracking technology or ingestible biosensors which can provide the data required in order to analyze a patients health on a constant basis. The IoT allows the connection of a number of smart medical devices which provide sensing, diagnosis, or imaging. In recent years, wearable technology has become increasingly prevalent, so biosensors can be incorporated easily to monitor health and fitness, while wireless communication technology allows the data gathered to be transmitted, stored, or analyzed. Smartphones are no considered wearable technology even though they are carried almost permanently by the user, since they are usually kept in a pocket or handbag. In contrast, smartwatches can easily be worn and do not inconvenience the user. They can also perform many of the functions usually associated with a smartphone. The design is typically focused on optimizing the computing capability while maintaining a miniaturized form so that the smartwatch will not have any adverse impact upon the activities of the wearer. Smartwatches are therefore the ideal technology to gather physiological data obtained directly from the bodily activities of the users [4].

HAR is currently a very appealing topic for research projects since it can be applied in numerous different ways [5]. It can be applied commercially, often through gaming and other functions which require interactions between computers and humans, such as health care applications. For the disabled and the elderly in particular, HAR can have valuable applications in monitoring and assisting these people. Various sensors can be used and linked to wearable devices such as smartwatches, or common technology in the shape of the smartphone [6], [7], [8], [9], [10].

This research study proposes a system for detecting sitting which applies HAR with smartwatch technology in order to combat the problem of OWS. Ten participants are involved in the study and wear smartwatches which gather accelerometer and gyroscope data. The data can then be used to construct a model capable of detecting sitting activity through the use of two approaches which apply ensemble learning. These approaches are the stacking method, and majority voting.

The structure of the paper takes the following form. Section II presents an account of the framework used to detect siting activity using a smartwatch to combat OWS. Section III explains the applications of machine learning for HAR. Section IV presents the design of the experiment along with the results, while Section V contains the conclusions.

II. The Framework of Smartwatch-Based Sitting Detection for Office Worker Syndrome

Activity recognition normally comprises four steps: preprocessing, extraction of features, training the model, and classification. The following generic model can be used to employ the basic steps within a system for recognition. The steps in this case begin with data collection, extraction of features, training the model, classification, and finally the detection of sitting, as indicated in Fig. 1.

A summary of some of those final five stages is presented as follows:

A. Data Collection

In this study, ten participants volunteered to provide the raw data. The subjects were aged 24 - 30 and were required to wear a smartwatch on the left wrist while carrying out six different activity types: sitting, standing, lying, walking, walking upstairs, and walking downstairs. An *Apple Watch Series 2* running on *WatchOS 4* version 4.2 was used to gather data, while the smartwatch application employed was SensorLog, which is available through the App Store. The application gathers data via the sensors in the smartwatch which are linked to the accelerometer and gyroscope and use a sample rate of 30 Hz. Ten samples were gathered for each activity of each subject, using each axis such that, for example, acceleration can be measured in the *x*, *y*, *z*-directions while gyroscope activity is also measured in the *x*, *y*, *z*-directions. The resulting dataset thus comprised 600 samples of human activity. Fig. 2 and Fig. 3 show examples of these data.

B. Feature Extraction

It was not possible to use the data from the accelerometer and gyroscope sensors directly for classification, so the step of extracting the features serves as a form of preprocessing which prepares the raw data for use. Machine learning models are more effective when proper feature sets are used.

It can be imagined that S_t represents a particular data fragment within a data window while S_i represents the *i*th data item provided by the accelerometer sensor or gyroscope. Statistical formulas can then be applied in order to calculate the nine feature functions applied in the course of this study.

1) Mean value

$$mean = \frac{1}{N}\sum_{i=1}^{N} S_i$$

2) Standard deviation

$$std = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(S_i - mean)^2}$$

3) Median absolute deviation

$$mad = median(|S_i - mean(S_t)|)$$

4) Maximum value

$$max = MAX(S_t)$$

5) Minimum value

$$min = MIN(S_t)$$

6) Signal magnitude

$$sma_{xyz} = \frac{1}{3}\left\{\sum_{j=1}^{N}|S_{xi}| + \sum_{j=1}^{N}|S_{yi}| + \sum_{j=1}^{N}|S_{zi}|\right\}$$

7) Energy measure

$$energy = \frac{1}{N}\sum_{i=1}^{N} S_i^2$$

8) Interquartile range

$$iqt = Q_3(S_t) - Q_1(S_t)$$

9) Correlation between axis

$$c_{xy} = \frac{\sum_{i=1}^{N}(S_{xi}-mean(S_x))(S_{yi}-mean(S_y))}{\sqrt{\sum_{i=1}^{N}(S_{xi}-mean(S_x))^2)\sum_{i=1}^{N}(S_{yi}-mean(S_y))^2}}$$

C. Classification

HAR is normally addressed as a problem of supervised learning and can involve common supervised learning models such as decision tree classifiers, support vector machines, and the multilayer perceptron. Among these classifiers, some offer greater accuracy when employed in certain contexts: multilayer perceptron models, for instance, showed better performance in more complex activities such as moving up or down stairs [11]. In the case of simple activities such as sanding or sitting, the decision tree approach was shown to be superior in terms of accuracy [12].

III. HAR Using Ensemble Learning

The study involved experimentation being conducted using different machine learning methods. The basic machine learning model was used along with the ensemble machine learning model in order to assess the recognition capabilities when applied to the HAR context. The study employed three basic models for machine learning: decision tree, support vector machine, and multilayer perceptron. These three approaches were then combined with the two ensemble-based approaches of stacking and majority voting.

A. Decision Tree (DT):

Decision trees are very widely used when machine learning is applied in a practical context [13]. They are able to determine underlying patterns where the roots are certain attributes while the extended branches form leaf nodes which can be considered as the concept, which in this case is the intuitive cognition of the subject. Trees are readily comprehensible if the attributes are seen to be rational.

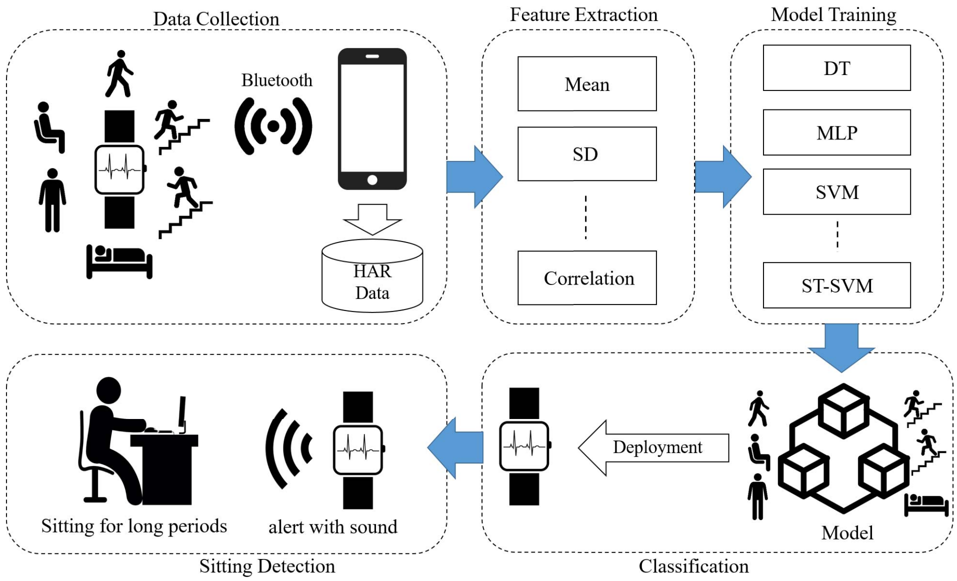

Fig. 1: Flow diagram of the proposed framework of sitting detection

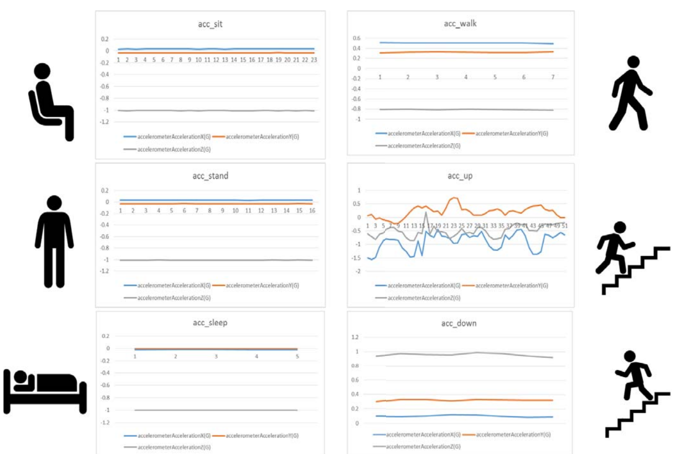

Fig. 2: An example of data from the accelerometer on smartwatch: sitting, standing, laying, walking, walking upstairs and walking downstairs

B. Support Vector Machine (SVM):

The original application of the support vector machine was to address problems of two-class pattern recognition [15]. The underlying principle is to construct a hyperplane capable of separating positive and negative instances while simultaneously maximizing the smallest margins. Support vector machines offer two main strengths: first of all it is not always necessary to use feature selection, and secondly there is often no need for parameter tuning.

C. Multilayer Perceptron (MLP):

Neural networks serve as tools for virtual intelligence tool which can trigger the brain to analyze situations and deliver results [14]. This algorithm for machine learning provides strong results in generalization, non-linear mapping, and self-organization. MLP is capable of providing reasonable outputs when given previously unencountered inputs. Because of this ability to generalize, it becomes possible to train networks when presented with representative input-output pairs, and to generate very good results when the network is trained on all of the feasible input-output pairs.

In order to enhance accuracy, one possible approach is to combine the basic machine learning models. This technique is called ensemble learning, and several previous studies have confirmed that when base models are combined, the accuracy is improved when compared to any of the base models in isolation [16]. For the purposes of this study, the HAR problem

Fig. 3: An example of data from the gyroscope on smartwatch: sitting, standing, laying, walking, walking upstairs and walking downstairs

was addressed using a pair of ensemble learning algorithms: stacking, and majority voting. Each method combined the three basic models comprising DT, SVM, and MLP.

D. Majority Voting (MV):

Majority voting is the most common ensemble method, which does not require any parameter tuning once the base classifiers have been constructed [17]. Assuming *d(y)* is the domain of the class label, *y*, $y_k(x)$ is the class label of HAR, *x* assigned by *k*th base model, and $v(y_k(x),c_i)$ is the indicator function:

$$v(y_k(x), c_i) = \begin{cases} 1 & \text{if } y_k(x) = c_i \\ 0 & \text{if } y_k(x) \neq c_i \end{cases} \tag{1}$$

The formula to compute *c(x)* assigned to an unlabeled HAR *x* is given as below:

$$c(x) = arg \max_{c_i \in d(y)} \left(\sum_k v(y_k(x), c_i) \right) \tag{2}$$

E. Stacking (ST):

Stacking ensemble methods use a meta classifier which is able to utilize the classification findings from numerous base classifiers in order to arrive at the final classification [18]. Stacking implementation requires a frame structure consisting of two layers, in which the various base classifiers are derived from the training dataset in the 0 level layer. The meta-classifier in the 1 level layer can then combine the individual classifiers, so the approach can use a decision tree (known as ST-DT), support vector machine (ST-SVM), or multilayer perceptron (ST-MLP). The two-layer stacking structure used in this study addressing the HAR problem with this ensemble learning technique is shown in Fig. 4.

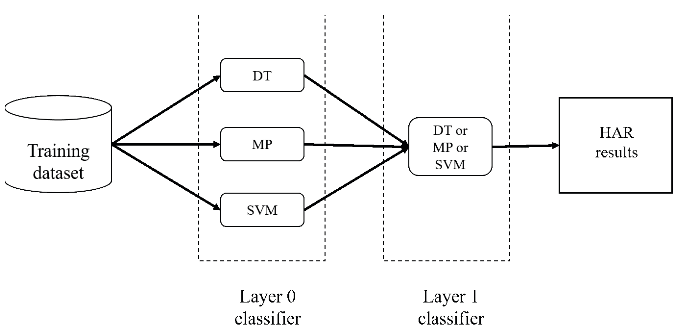

Fig. 4: Two-layer stacking diagram of the HAR problem

IV. Experimental Design and Results

To assess the various ensemble learning methods using the data gathered, a WEKA machine learning tool was employed. Seven classifiers (DT, MLP, SVM, MV, ST-DT, ST-MLP, and ST-SVM) were then used with the dataset representing the activities of the ten participants mentioned in Section II in order to evaluate their performance. A 10-fold cross validation technique was applied in assessing the various classifiers. The cross validation approach requires that each fold or data item holds all classes in similar proportions in order to ensure the fairness of the process. The seven most common classifiers are shown in Table I, along with short notes describing the classifiers to be applied in subsequent sections.

TABLE I: Seven classifiers used in this work

Type of Classifiers	WEKA-version	Notation
Decision trees	J48	DT
Neural networks	MulitlayerPerceptron	MLP
Support vector machines	LibSVM	SVM
Majority voting	Vote	MV
Stacking	Stacking with J48	ST-DT
Stacking	Stacking with MLP	ST-MLP
Stacking	Stacking with SVM	ST-SVM

During the experimental process, the accuracy value measured HAR performance. These values were calculated from the value of true positive (*TP*), true negative (*TN*), false positive (*FP*), and false negative (*FN*) defined by the formula given below:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{3}$$

The study employed data from the accelerometer and gyroscope on a smartwatch for the representation of each kind of activity. The six different activities typically performed by people were then evaluated when combined with the seven classifiers which are stipulated in Table I. The results were found to be reasonable and can be observed in Tables II and III.

TABLE II: Experimental results of HAR recognition performance of base classifiers

Classifiers	Recognition Performance (%Accuracy)		
	Acc. data	Gyro. data	Acc. and Gyro. data
DT	85.9000	80.3000	87.0571
MLP	87.6143	84.7571	88.3429
SVM	85.8571	83.7857	88.9429

Table II summarizes the broad recognition performance for each of the three base classifiers, revealing that the best performance came from SVM in applying data from the accelerometer and gyroscope to achieve an accuracy level of 88.9429 percent.

Table III summarizes the performances for each of the classifiers when used in the two ensemble learning approaches. According to the findings, the results were improved in terms of recognition performance for all seven classifiers. MV obtained the best achievement using data from the accelerometer and gyroscope at 89.6571 percent for accuracy.

In order to develop the required framework, the researchers conducted experiments design to examine the problem of detecting sitting behavior. The data were collected and then divided into two activity classes. Sitting activity was classed

TABLE III: Experimental results of HAR recognition performance of ensemble classifiers

Classifiers	Recognition Performance (%Accuracy)		
	Acc. data	Gyro. data	Acc. and Gyro. data
MV	88.6429	84.3857	89.6571
ST-DT	85.7571	81.2286	86.5286
ST-MLP	87.0000	82.6000	87.7714
ST-MVM	87.9429	83.0429	88.0143

as SITTING, while lying, standing, walking, moving upstairs, and moving downstairs were all classed as OTHER. The results for these experiments are presented in Tables IV and V.

TABLE IV: Experimental results of sitting detection of base classifiers

Classifiers	Recognition Performance (%Accuracy)		
	Acc. data	Gyro. data	Acc. and Gyro. data
DT	91.9429	88.5429	93.2286
MLP	85.5714	85.7000	85.7857
SVM	85.9000	85.7143	86.9143

Table IV shows a summary of how well each base classifier performed with respect to the activity data collected from a smartwatch. Based on the results, the decision tree model performed with the highest accuracy on sitting detection. It yielded the highest accuracy with 93.2286 percent.

TABLE V: Experimental results of sitting detection of ensemble classifiers

Classifiers	Recognition Performance (%Accuracy)		
	Acc. data	Gyro. data	Acc. and Gyro. data
MV	87.0000	85.7143	87.0000
ST-DT	91.3571	88.2571	93.5714
ST-MLP	91.9429	87.9857	92.7143
ST-MVM	91.9429	88.4286	92.5714

Table V shows a summary of how each improved ensemble classifier performed to solve the sitting detection problem. Based on the results, the stacking model ST-DT performed with the highest accuracy at 93.5714 percent.

V. Conclusion

This paper has presented a framework for the purpose of detecting sitting activity using HAR to combat OWS. The study used data gathered from a smartwatch containing an accelerometer and gyroscope in order to construct the framework with classifiers. Furthermore, a pair of ensemble learning approaches were examined in order to address the problem. The experiments revealed that the combination of accelerometer and gyroscope performs better than either in isolation to achieve optimal recognition performance. It was also discovered that the use of stacking for the ensemble learning approach was able to enhance the recognition capability to achieve an accuracy level of 93.5714 percent.

Acknowledgment

This paper was supported in part by the School of Information and Communication Technology, University of Phayao, Thailand.

References

[1] Internayional Labour Organization (ILO), "ILO in Thailand," 2017. [Online]. Available: http://www.ilo.org/asia/countries/thailand/lang–en/index.htm. [Accessed: 30-Sep-2017].

[2] PRLog, "Office workers syndrome: a new condition to focus on with your massage therapist," 2017. [Online]. Available: http://www.prlog.org/12145934-office-workers-syndrome-newcondition-to-focus-on-with-your-massage-therapist.html. [Accessed: 30-Sep-2017].

[3] G. Sun, F. Yu, X. Lei, Y. Wang, and H. Hu, "Research on mobile intelligent medical information system based on the internet of things technology," *2016 8th International Conference on Information Technology in Medicine and Education (ITME)*, pp. 260–266, 2016.

[4] D. Tian, X. Xu, Y. Tao, and X. Wang, "An improved activity recognition method based on smart watch data," in *2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, vol. 1, July 2017, pp. 756–759.

[5] R. Poppe, "A survey on vision-based human action recognition," *Image and Vision Computing*, vol. 28, no. 6, pp. 976 – 990, 2010.

[6] H. Ghasemzadeh and R. Jafari, "Physical movement monitoring using body sensor networks: A phonological approach to construct spatial decision trees," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 1, pp. 66–77, Feb 2011.

[7] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *SIGKDD Explor. Newsl.*, vol. 12, no. 2, pp. 74–82, Mar. 2011.

[8] T. Brezmes, J.-L. Gorricho, and J. Cotrina, "Activity recognition from accelerometer data on a mobile phone," in *Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living*, S. Omatu, M. P. Rocha, J. Bravo, F. Fernández, E. Corchado, A. Bustillo, and J. M. Corchado, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 796–799.

[9] W. Wu, S. Dasgupta, E. E. Ramirez, C. Peterson, and J. G. Norman, "Classification accuracies of physical activities using smartphone motion sensors," *J Med Internet Res*, vol. 14, no. 5, p. e130, Oct 2012.

[10] X. Liu, L. Liu, S. J. Simske, and J. Liu, "Human daily activity recognition for healthcare using wearable and visual sensing data," in *2016 IEEE International Conference on Healthcare Informatics (ICHI)*, Oct. 2016, pp. 24–31.

[11] M. F. A. Abdullah, A. F. P. Negara, S. Sayeed, D. Choi, and K. Anbananthen, "Classification algorithms in human activity recognition using smartphones," vol. 6, pp. 415–432, 01 2012.

[12] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive Computing*, A. Ferscha and F. Mattern, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 1–17.

[13] P. E. Utgoff, "Incremental induction of decision trees," *Machine Learning*, vol. 4, no. 2, pp. 161–186, Nov 1989.

[14] L. Fausett, Ed., *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1994.

[15] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag New York, Inc., 1995.

[16] L. Rokach, "Ensemble-based classifiers," *Artificial Intelligence Review*, vol. 33, no. 1, pp. 1–39, Feb 2010.

[17] G. Brown, J. L. Wyatt, and P. Tiňo, "Managing diversity in regression ensembles," *J. Mach. Learn. Res.*, vol. 6, pp. 1621–1650, Dec. 2005.

[18] E. K. Tang, P. N. Suganthan, and X. Yao, "An analysis of diversity measures," *Machine Learning*, vol. 65, no. 1, pp. 247–271, Oct 2006.