Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 163 (2019) 54–61

Procedia Computer Science

www.elsevier.com/locate/procedia

16th International Learning & Technology Conference 2019

Smartphone-Based Human Activity Recognition Using Bagging and Boosting

Abdulhamit Subasi[a], Asalah Fllatah[b], Kholoud Alzobidi[b], Tayeb Brahimi[c], Akila Sarirete[b]

[a]*Department of Information System, College of Engineering, Effat University, P.O. Box 34689, Jeddah 21478, Saudi Arabia*
[b]*Department of Computer Science, Effat University, P.O. Box 34689, Jeddah 21478, Saudi Arabia*
[c]*Department of Electrical and Computer Engineering, Effat University, P.O. Box 34689, Jeddah 21478, Saudi Arabia*

Abstract In today's healthcare applications, the use of mobile technologies brings together physicians and patients for intelligent and automatic monitoring of daily clinical activities, remote life assistants, and preventive care, especially for the elderly and those under medical control. As smartphones become an important part of our everyday life, they are ever more employed in human activities recognition (HAR) including the monitoring of personal health care and wellbeing. However, HAR is complex and it is important to use the best technology and learn about human activity using machine learning. The purpose of this paper is to develop a HAR system based on the smartphone sensors' data using Bagging and Adaboost ensemble classifiers. The experimental results for the HAR data have been evaluated after performing different data mining techniques. For each subject, the total classification accuracy, the F-measure, and the ROC area were calculated. Adaboost ensemble classifiers algorithm improved significantly the performance of smartphone-based HAR, combined with SVM, it reached 97.44% accuracy compared to the rest of the classifiers. The proposed algorithm of Adaboost SVM can lead to an accurate HAR for elderly and disabled patients who need continuous care as well as it is a tool that supports the decisions of all medical practitioners.

Keywords: Smartphone; Human Activity Recognition (HAR); BAgging; Adaboost, Ensemble Classifier

1877-0509 © 2019 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)
Peer-review under responsibility of the scientific committee of the 16th International Learning & Technology Conference 2019.
10.1016/j.procs.2019.12.086

1. Introduction

Today many healthcare applications would benefit from the advance in information and communication technologies, wireless network technology, wearable devices, and smartphone technology to analyze patient big data, medical images, videos, and images, based on data mining techniques that are also capable of recognizing activities of daily living (ADLs) and Human Activity Recognition (HAR) [**Error! Reference source not found.Error! Reference source not found.**]. Monitoring patient's activities have provided effective results especially in cases of elderly and disabled patients who need continuous care. Due to its major effects on a human in general and healthcare in particular, HAR recognition has become an essential field in computing. Moreover, the integration of smartphone in healthcare has led to initiate smart applications such as mobile healthcare and intelligent healthcare monitoring systems [3, 4, 5, 6]. With the increase of the population age of our society and the increase in their health risk, HAR has proven to help detect cases of different health problems such as Parkinson's disease, dyskinesia, tremor, dystonia, or bradykinesia among others [7,8]. There are several ways to deliver activity recognition that are classified by different "taxonomies" which differentiates the types of recognition algorithms based on machine learning and logical modeling and reasoning. According to Anguita et al. [9], the main focus is the HAR dataset and energy efficiency. Energy is believed to be saved by using simpler calculation methods which are hardware friendly. This research presents an effective way of recognition of daily activities using smartphones. Different data mining techniques have been used to examine human behavior based on smartphone technology [2, 9]. The experiments were carried out on 30 volunteers wearing Samsung Galaxy S II smartphone and performing three static posture activities namely standing, sitting, and lying and three dynamic postures activities namely walking, walking downstairs, and walking upstairs [10]. The experiment also includes postural transitions stand-to-sit, sit-to-stand, sit-to-lie, lie-to-sit, stand-to-lie, and lie-to-stand. A user-independent data mining approach for off-line human activity classification is developed based on smartphone sensors' data using Bagging and Adaboost ensemble classifiers. The experimental results for the HAR data are evaluated after performing different data mining techniques. The rest of the paper is organized as follows. Section 2 gives a review of related work in the area of human activity recognition and the usage of smart phones. Sections 3 presents the dataset specification along the data mining techniques used in HAR. Section 4 shows and discusses the experimental results of the data mining techniques. Section 5 summarizes the work of this study, draws a conclusion, and provides some guidelines for future work.

2. Related Work

The spread of mobile phones across the general public has made it an essential everyday element, they can monitor our daily activities to learn from them and help them make decisions [11]. Smartphones are equipped with sensors that are able to recall daily activities, detect falls, evaluate wellness, and much more. These sensors are able to assist in healthcare, by detecting a person's daily activities, evaluating a person's health and predicting when a change in their health occurs which eventually helps in to identify possible diseases. Human activity recognition is a challenge since different people may act differently while performing the same activity. Moreover, for the same activity, the same person may also act differently at different times. Different type of sensors are being used in HAR, they are mainly divided into two categories i) vision based which monitor patient activity using cameras and ii) sensor based technique where physical sensors are attached to the patient's body such as accelerometers and gyroscopes [12, 13]. The vision-based approach is not practical, even though it provides great recognition rate, due to its use of camera and video in indoor environments, this raises privacy and other technical issues. The wearable based approach is a lot more convenient due to its size and flexibility as it allows to carry out devices embedded with sensors [14]. Activity recognition systems on a smartphone can be trained to classify data in real time using the geometric template while the classifier uses a support vector machine. Activity recognition requires a training phase and a recognition phase where the data collected is processed based on the activity models. As for sensors connected to these smart devices, a benefit is that the data collected through these sensors for human activity recognition can be analyzed offline using machine learning tools. Even though that now it can be also implemented in an online pattern, through the more powerful resources that smartphones have now, including processors, memory, and battery [2].

One of the obstacles that HAR recognition is the correct information due to the huge amount of processed data. The supervised learning uses correct information that raises the privacy concerns due to the need for daily living activity and true information [15]. However, In order to solve the privacy issue researcher came up with smartphones sensor that works in background mode, which is achieving a good result in the low-cost matter [16]. Furthermore, the researchers look for new methods and techniques that help to improve the accuracy and integrity level of smartphones sensors. They found out a new method that used the smartphone accelerometer pair with dedicated chest sensor to recognize the human activity by times17]. Lu et al. [18] came up with a new technique which is compressed sensing method to recognize human motion based on compressed sensing theory and the outcomes of their paper reach 86% accuracy. In order to improve the activity recognition field, the authors studied the cases and scenarios of accuracy level when the results can be accepted or rejected. The quality of metrics has been a dialectic issue between researchers who seek to improve the method and tools used to measure the performance. This creates a new method used by model parameter optimization in application to enhance the quality of recognition. However, the activity recognition is not only on information technology environment. It used by the medical researcher in order to enhance the health services [8]. Moreover, Ronao & Cho [19] argued how deep learning network helps Human activity recognition at smartphones. They used this technique to increase human activity recognition efficient by searching the ingrained feature of activities. In the healthcare domain, HAR is used by the medical researcher in order to enhance the health services [8]. Moreover, Ronao & Cho [19] argued how deep learning network helps Human activity recognition at smartphones. They used this technique to increase human activity recognition efficient by searching the ingrained feature of activities. The work of Reiss et al. [20] is a good reference in explaining one of the biggest physical activity monitoring challenges related to the difficulties with complex classification problems that are beyond the potential of actual classifiers. The authors proposed ConfAdaBoost.M1 algorithm that combines some of the benefits of boosting methods. Ronao et al [19] proposed to establish a deep convolutional neural network to analyze HAR operations using smartphones. Results showed that the method has relevant and more complex characteristics. Arif et al. [21] used a smartphone with an accelerometer sensor for monitoring physical activities of a person including walking, sitting, walking, standing, and walking upstairs or downstairs. A 98% classification accuracy has been obtained for the six types of physical activities. Martín et al. [22] focused on exploration Use smartphones for HAR without affecting the user's lifestyle. Guiry et al. [23] described a method to detect human activity using a smartphone and concluded that it is possible to deduce HAR using only five features of two accelerometers, the analysis of the data showed an accuracy of 98%.

3. Dataset and Methodology

3.1. Dataset

The dataset used to bee implemented in this research is based on the UCI repository [10]. Thirty volunteers with age between 19 and 48 wearing Samsung Galaxy S II smartphone on their waist performed six basic activities namely, standing, sitting, lying, walking, ascending stairs, and descending stairs. The analysis also includes postural transitions between the three static postures and the three dynamic postures. Acceleration data was recorded with a frequency of 50Hz while the data was labeled manually using a video recording process. The dataset was divided randomly into training and testing data by 70% and 30% respectively. Noise filters have been applied and then sampled in fixed-width sliding windows of 2.56 sec and 50% overlap. From each window, there is a vector of 561 features that can be obtained by calculating variables from the time and frequency domain [9, 24].

3.2. Data Mining Techniques

The main objective of data mining techniques is to mine knowledge, identify, and extrapolating patterns and new knowledge from a collection of a large number of data sets. These techniques involve an algorithmic and interactive knowledge discovery of unidentified patterns based on finding relationships within a large set of data. Among the techniques used in data mining are tracking patterns, classification, outlier detection, neural networks, decision trees, clustering, association, regression, and prediction. Classification algorithms are applied first on training data to create a model, and then the model is tested on predefined test data to measure its accuracy. Different machine learning classification techniques have been used to perform HAR. There are several classifiers available in WEKA, in this work seven standard prediction algorithms have been used for comparative purposes along with bagging and AdaBoost techniques [25]. These algorithms include k-Nearest neighbors, neural networks, Naïve Bayesian networks, and decision trees. The k-nearest neighbor algorithm is a method broadly applied for pattern recognition comparing a given test tuple with similar training tuples, however, it requires a lot of work for the training [26, 27]. The Support Vector Machine (SVM) is an algorithm that classifies linear and nonlinear data. Although SVMs are very accurate and can model complex nonlinear decision boundaries, yet they could use extremely slow training time [26, 27]. The power of these algorithms decreases with the existence of dependencies between attributes. Another concern about the Naïve Bayes is the assumption that numeric attributes are normally distributed. Yet, standard estimation procedures can be applied to a particular attribute that is likely to follow some other distribution 26, 27]. Decision trees algorithms include a leaf, a branch, and root nodes where each internal node represents a test on an attribute, the branch represents the outcome of the test, and the topmost node in the tree represents the root node. C4.5 algorithm creates decision trees classifiers as well as classifiers in more comprehensible ruleset form. It applies two heuristic measures to evaluate potential tests: (1) information gain that reduces the overall entropy of the subsets, and (2) the default gain ratio that splits information gained by the information offered by the test outcomes [28]. REPTree forms a decision or regression tree using information gain/variance reduction and prunes it employing reduced-error pruning. For optimal speed, REPTree algorithm categories values for numeric attributes only once. Missing values are managed through splitting instances into parts, similar to C4.5 approach [29]. Decision tree forests are a group of tree classifiers grown relevant to random vectors. LADTree is a decision tree algorithm that can handle multiclass issues based on the LogitBoost algorithm [29]. Random Forests are a combination of tree predictors, [30] the generalization error of tree classifiers of forests relies on the power of each tree in the forest and the interrelationship between them.

3.2.1. Bagging and Boosting

Bagging and Boosting are two techniques used to improve prediction rules obtained by a classification algorithm [31, 25]. Both techniques are referred to as “perturb and combine” (P&C) method [32] methods where a classifier is applied to different perturbations of the original data set then, the output is combined with a single classifier method. Bagging implements bootstrap to produce L training sets that trains L base-learners by an unstable learning method, and then, during testing, take an average [33]. Bagging is useful for classification and regression problems. Using the median rather than the average in joining predications creates a more robust regression. Averaging lower variance in one case only and that if the positive correlation is small; an algorithm is stable if distinctive runs of the same algorithm on resampled copies of the identical dataset generates learners with a high positive correlation [34]. Like the bagging, boosting employs voting or averaging to integrate the output of individual models. Bagging constructs individual models separately while boosting is iterative where every constructed model is affected by the performance of those created past. Boosting promotes new models to become experts for instances addressed not right by earlier ones by giving greater weight to those instances. Also, there is the difference between them that is boosting weights a model’s impact by its performance instead of offering weight to all models. The boosting algorithm starts by giving the same weight to all instances in the training data then calls the learning algorithm to compose a classifier for this data and reweights each instance conferring to the classifier's output.

3.2.2. AdaBoost Ensemble Classifier

Boosting can be best defended as an improvement of bagging that emphasis on including base model diversity by shifting the focus during base model formation to instances that turn out the most “predictively difficult”. Adaptive Boosting ensemble classifier (AdaBoost) developed by Freund and Schapire [35] is considered as the most recognized technique of appropriate boosting of two-class classification problems. AdaBoost adds to the generic boosting procedures fundamental defined features as well as model weighting schemes. The model weight relays on the misclassification error of training set data set estimated based on the definition. This increases the weights of the misclassified instances while reducing the weights of appropriately classified instances. The degree of increase and decrease depends on the weight of the model. According to [27] highly accurate models (higher weighted) provides broader instance weight.

4. Result and Discussion

4.1. Performance Evaluation Metrics

The performance measures used in this study is based on the total classification accuracy, the F-measure, and the receiver operating characteristics (ROC). Accuracy is considered as the most important parameter showing the performance of a classifier. It is a measure of the correctly classified instances from different classification techniques such as k-Nearest neighbors, neural networks, Naïve Bayesian networks, and decision trees. For a true a positive TP (a hit), a true negative TN (correct rejection), a false positive FP where the final result is incorrectly expected as sure, and false negative FN, where the final result was incorrectly expected to be poor but it turns out to be truly positive [29], the success rate is defined as the ratio of the number of correctly classified instances to the total number of instances

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} \tag{1}$$

The precision parameter is defined as the ratio of the correctly classified instances to the total number of instances classified

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

The recall parameter is defined as the ratio of the correctly classified instances to the total number of instances belonging to this class

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

Finally, the F-measure is the weighted average of *Precision* and *Recall* together.

$$F_{measure} = \frac{2\, Precision \times Recall}{Precision+Recall} \tag{4}$$

4.2. Validation of the Results

The experiments conducted in this study consider 30 users wearing Samsung Galaxy S II smartphone and performing static and dynamic posture activities: standing, sitting, and lying, walking, walking downstairs, and walking upstairs including postural transitions stand-to-sit, sit-to-stand, sit-to-lie, lie-to-sit, stand-to-lie, and lie-to-stand. The experimental results for the HAR data have been evaluated after performing seven data mining techniques as shown in Table 1 along with Single, Bagging, and AdaBoost classifiers. While most methods show good accuracy, it is clear from Table 1 and Figure 1 that Adaboost ensemble classifiers algorithm improved significantly the performance of smartphone-based HAR when combined with SVM where the accuracy reached 97.44%. Naïve Bayes showed low performance with 74.6% and 74.54% with Single and Bagging classifiers, however, the accuracy is increased when using AdaBoost and reaches 86.72%. However, the accuracy of Random Tree classifier is not improved by using Adaboost ensemble classifier as in the case of Bagging where it reached 93.52% compared to 83.49% with Adaboost. Other than this case, Adaboost performed well in all other cases, with SVM it can certainly be used in human activity recognition in to support the decisions of all medical practitioners.

Table 1. Classification results of Single, Bagging, and Adaboost Ensemble Classifiers.

	Single			Bagging			Adaboost		
	ROC Area	F – measure	Accuracy (%)	ROC Area	F - measure	Accuracy (%)	ROC Area	F - measure	Accuracy (%)
k-NN	0.974	0.953	95.35%	0.994	0.952	95.29%	0.974	0.953	95.35%
SVM	0.97	0.973	97.31%	0.998	0.974	97.38%	0.996	0.974	97.44%
NB	0.957	0.73	74.60%	0.964	0.729	74.54%	0.977	0.868	86.72%
RF	0.999	0.955	95.60%	0.979	0.959	95.95%	0.999	0.96	96.05%
C4.5	0.962	0.918	91.78%	0.998	0.944	94.37%	0.999	0.967	96.74%
REP Tree	0.983	0.899	89.88%	0.998	0.942	94.31%	0.999	0.962	96.24%
Random Tree	0.906	0.837	83.71%	0.995	0.934	93.52%	0.904	0.835	83.49%

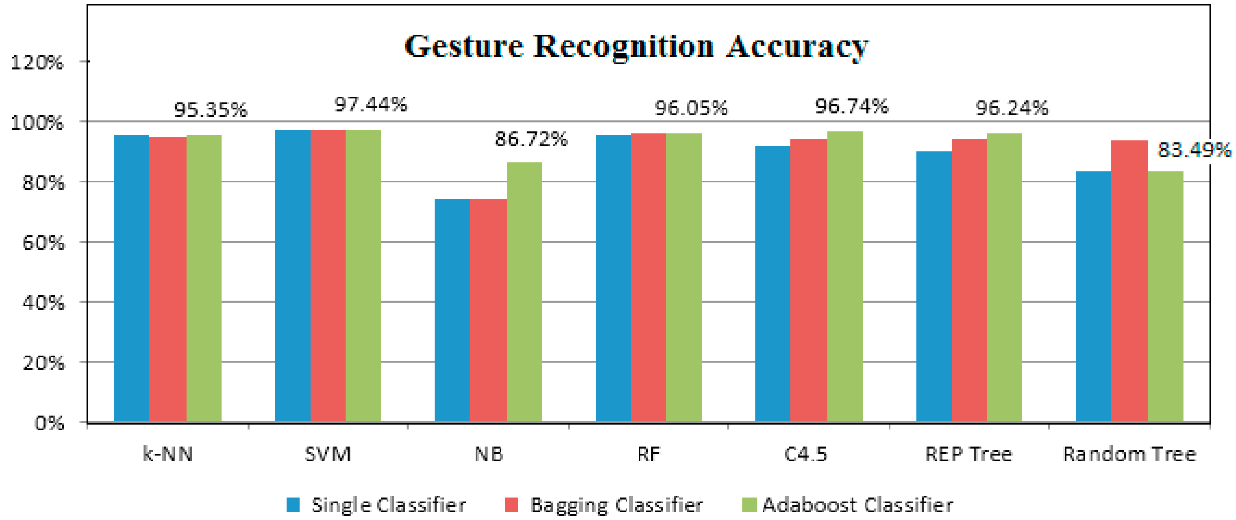

Fig. 1. Comparison of gesture recognition accuracy for Single, Bagging, and Adaboost Classifiers.

5. Conclusion

With the help of smartphones, this study investigated human activity recognition (HAR) using data mining techniques. Using the best technology and benefit from the development of machine learning it was possible to provide accurate information about human activity especially in cases of elderly and disabled patients who need continuous care. In the present study, it has been shown that using Adaboost ensemble classifier in HAR achieved good performance when combined with k-Nearest neighbors, neural networks, Naïve Bayesian networks, and decision trees. Moreover, results from the present study showed that Adaboost ensemble classifiers outperformed among the other data mining techniques selected in this research. Adaboost ensemble classifiers algorithm improved the accuracy of smartphone-based HAR by 97.44% when combined with SVM. The authors believe that the proposed algorithm may lead to better predictions in smartphone-based Human Activity Recognition which is needed not only in the development of accurate and efficient healthcare monitoring systems but also in a smart city environment. Nevertheless, there is always space and opportunities for improvements and provide more accurate and efficient results. This study requires additional research and methodologies to be implemented in the future and include different aspects of the field.

Acknowledgments

The author would like to acknowledge the support of the College of Engineering at Effat University, Jeddah, Saudi Arabia.

References

[1] Xi Liu-Lei Liu-Steven Simske-Jerry Liu. (2016). Human Daily Activity Recognition for Healthcare Using Wearable and Visual Sensing Data. IEEE International Conference on Healthcare Informatics (ICHI).

[2] Shoaib, M., Bosch, S., Incel, O., Scholten, H., & Havinga, P. (2015). A Survey of Online Activity Recognition Using Mobile Phones. Sensors,15(1), 2059-2085. doi:10.3390/s150102059

[3] Guan Y., Zhaohui W., Fanrong M., Qiuyan Y., Shixiong X., (2018). An overview of human activity recognition based on smartphone. Sensor Review, https://doi.org/10.1108/SR-11-2017-0245 Permanent link to this document: https://doi.org/10.1108/SR-11-2017-0245

[4] A Survey of Online Activity Recognition Using Mobile Phones. Sensors,15(1), 2059-2085. doi:10.3390/s150102059

[5] Fareed, U. (2015). Smartphone Sensor Fusion based Activity Recognition System for Elderly Healthcare. Proceedings of the 2015 Workshop on *Pervasive Wireless Healthcare* - MobileHealth 15. doi:10.1145/2757290.2757297

[6] Boulos, M. N. K., Wheeler, S., Tavares, C., & Jones, R. (2011). How smartphones are changing the face of mobile and participatory healthcare: an overview, with an example from eCAALYX. Biomedical Engineering Online, 10(1), 24.

[7] Gjoreski, M., Gjoreski, H., Luštrek, M., & Gams, M. (2016). How Accurately Can Your Wrist Device Recognize Daily Activities and Detect Falls? *Sensors,16*(6), 800. doi:10.3390/s16060800

[8] Woznowski, P., Kaleshi, D., Oikonomou, G., & Craddock, I. (2016). Classification and suitability of sensing technologies for activity recognition. Computer Communications, 89–90, 34–50. https://doi.org/10.1016/j.comcom.2016.03.006

[9] Anguita, D., Ghio, A., Oneto, L., Llanas Parra, F. X., & Reyes Ortiz, J. L. (2013). Energy efficient smartphone-based activity recognition using fixed-point arithmetic. Journal of Universal Computer Science, 19(9), 1295–1314.

[10] "UCI Machine Learning Repository: REALDISP Activity Recognition Dataset Data Set." [Online]. Available: https://archive.ics.uci.edu/ml/datasets/REALDISP+Activity+Recognition+Dataset#. [Accessed: 24-Jun-2018].

[11] Chetty, G., White, M., & Akther, F. (2015). Smartphone-based data mining for human activity recognition. Procedia Computer Science, 46, 1181–1187.

[12] Ye Liu, Liqiang Nie, Li Liu, David S. Rosenblum. (2016). From action to activity: Sensor-based activity recognition. Neurocomputing, 181 (2016) 108–115.

[13] Yu, H., Cang, S., & Wang, Y. (2016). A review of sensor selection, sensor devices and sensor deployment for wearable sensor-based human activity recognition systems. 2016 10th *International Conference on Software, Knowledge, Information Management & Applications* (SKIMA). doi:10.1109/skima.2016.7916228

[14] Wang, A., Chen, G., Yang, J., Zhao, S., & Chang, C.-Y. (2016). A comparative study on human activity recognition using inertial sensors in a smartphone. IEEE Sensors Journal, 16(11), 4566–4578.

[15] Hossain, H. S., Khan, M. A. A. H., & Roy, N. (2016). Active learning enabled activity recognition. In Pervasive and Mobile Computing.

[16] Martín, H., Bernardos, A. M., Iglesias, J., & Casar, J. R. (2013). Activity logging using lightweight classification techniques in mobile devices. Personal and Ubiquitous Computing, 17(4), 675–695. https://doi.org/10.1007/s00779-012-0515-4

[17] Guiry, J. J., Ven, P. van de, Nelson, J., Warmerdam, L., & Riper, H. (2014). Activity recognition with smartphone support. Medical Engineering and Physics, 36(6), 670–675. https://doi.org/10.1016/j.medengphy.2014.02.009

[18] Lu, Y.; Wei, Y.; Liu, L.; Zhong, J.; Sun, L.; Liu, Y. Towards unsupervised physical activity recognition using smartphone accelerometers. *Multimed. Tools Appl.***2017**, *76*, 10701–10719

[19] Ronao, C. A., & Cho, S.-B. (2016). Human activity recognition with smartphone sensors using deep learning neural networks. Expert Systems with Applications, 59, 235–244. https://doi.org/10.1016/j.eswa.2016.04.032

[20] Reiss, A., Hendeby, G., & Stricker, D. (2015). A novel confidence-based multiclass boosting algorithm for mobile physical activity monitoring. Personal and Ubiquitous Computing, 19(1), 105–121. https://doi.org/10.1007/s00779-014-0816-x

[21] Arif, M., Bilal, M., Kattan, A., & Ahamed, S. I. (2014). Better Physical Activity Classification using Smartphone Acceleration Sensor. Journal of Medical Systems, 38(9), 95. https://doi.org/10.1007/s10916-014-0095-0

[22] Martín, H., Bernardos, A. M., Iglesias, J., & Casar, J. R. (2013). Activity logging using lightweight classification techniques in mobile devices. Personal and Ubiquitous Computing, 17(4), 675–695. https://doi.org/10.1007/s00779-012-0515-4

[23] Guiry, J. J., Ven, P. van de, Nelson, J., Warmerdam, L., & Riper, H. (2014). Activity recognition with smartphone support. Medical Engineering and Physics, 36(6), 670–675. https://doi.org/10.1016/j.medengphy.2014.02.009

[24] Reyes-Ortiz, Jorge-L, Oneto, L., Sama, A., Parra, X., & Anguita, D. (2016). Transition-aware human activity recognition using smartphones. Neurocomputing, 171, 754–767.

[25] Sutton, Clifton D. (2004). Classification and Regression Trees, Bagging and Boosting. Handbook of Statistics, 24, DOI 10.1016/S0169-7161(04)24011-1

[26] Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques*. Amsterdam: Elsevier, Morgan Kaufmann.

[27] Cichosz, P. (2014). Data Mining Algorithms: Explained Using R. John Wiley & Sons.

[28] Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., … Philip, S. Y. (2008). Top 10 algorithms in data mining. Knowledge and Information Systems, 14(1), 1–37.

[29] Hall, M., Witten, I., & Frank, E. (2011). Data mining: Practical machine learning tools and techniques. Kaufmann, Burlington.

[30] Breiman, L. (2001). Random forests. Machine Learning, 5–32.

[31] Breiman, Leo. (1996). Bagging predictors. Machine Learning, 24(2), 123–140.

[32] Breiman, Leo. (1998). Arcing classifiers (with discussion). *Ann. Statist* 26, 801-849

[33] Holmes, B. Pfahringer, R. Kirkby, E. Frank, and M. Hall. (2002). Multiclass alternating decision trees. presented at the European Conference on Machine Learning, pp. 161–17

[34] Alpaydin, E. (1997). Voting over multiple condensed nearest neighbors. In Lazy learning (pp. 115–132). Springer.

[35] Freund, Y., Schapire, R. (1996). Experiments with a new boosting algorithm. In: Saitta L. (Ed), *Machine Learning*: Proceedings of the 13th International Conference. Morgan Kaufman, San Francisco, Ca.