International Journal of Engineering & Technology, 7 (4.19) (2018) 349-354

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

The Preprocessing for Predicting of Physical Activity Recognition

Sakchai Muangsrinoon[1*], Poonpong Boonbrahm[2]

[1]School of Informatics, Walailak University, [2] Nakhon Si Thammarat, Thailand
**Corresponding Author Email: [1]msakchai@gmail.com*

Abstract

This experiment examined the preprocessing for predicting of physical activity recognition model to access the relationship between time duration of sensors, the single tri-axial accelerometer, and fitness recognition (sitting, standing, walking, and running). The experimented with sixteen students (62.5% male and 37.5% female, age between eighteen through twenty-three year old) of the Informatics school at Walailak University. The authors had the experimental setup with the split dataset, 80% for training and testing, and 20% for validation, and repeated k-fold Cross-Validation (number=10, repeats=3) for resampling method to evaluate model performance for baseline models. When the authors measured model's performance, the authors found the follows results. First – the raw dataset with 123,156 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor and RF: Random Forest is 100%. Second – the aggregate dataset time duration 1 second with 1,240 samples, the best models performance has accuracy level with RF: Random Forest is 100%.Third – the aggregate dataset time duration 5 seconds with 251 samples, the best models performance has accuracy level with RF: Random Forest is 99.5%. Fourth – the aggregate dataset time duration 10 seconds with 128 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor is 96.82%. Fifth – the aggregate dataset time duration 15 seconds with 86 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor is 96. 21%.Sixth – the aggregate dataset time duration 20 seconds with 66 samples, the best models performance has accuracy level with LDA: Linear Discriminant Analysis is 98%.Seventh – the aggregate dataset time duration 25 seconds with 54 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor is 96.33%. Moreover, finally, Eight – the aggregate dataset time duration 30 seconds with 46 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor is 93.61%. In the future work, the authors planned to get more accuracy model by adding more features from another sensor, heart rate. Mining data collected from sensors provide valuable result in the physical activity recognition area. The improvement in performance is required especially in the healthcare field. The more increasing of using the wearable device, the wider opportunity in the data mining research area can be.

***Keywords**: Accelerometer, Android Wear, Preprocessing, Multiclass classification, Physical activity recognition.*

1. Introduction

According to increasing trend of using the wearable device for health application, there are a lot of set up to research data mining from sensor data such as accelerometer and gyroscope. This experiment examined the physical activity recognition model for the association between the single tri-axial accelerometer with three features on android smart watch and fitness recognition in sixteen students (62.5% male and 37.5%, age between eighteen through twenty-three year old) of the Informatics school at Walailak University. The experiment criteria are, each volunteer has to do physical activity: sitting, standing, walking, and running for five minutes by activity. Thus, the objective of this study is to assess the preprocessing dataset of physical activity recognition model for the association between single tri-axial accelerometer with three features on Android mobile phone and physical activity recognition in sixteen students while having physical activity.

The physical activity recognition application on android wear was developed to support two primary functions on the model of machine learning algorithm, model training and model inference.

For model training, the dataset, physical activity, was trained for sitting, standing, walking, and running with at least five minutes of each activity with the mixed gender of sixteen students. Next, the authors had processed the following steps: Define the problem, Summarize data, Prepare data, Evaluate algorithms, and Finalize model. Predicting and monitoring the student physical activity in daily used with model inference.

The resampling methods: (1) Bootstrap (2) k-fold cross-validation with and without repeats (3) Leave one out cross-validation, were evaluated with Naive Bayes (NB) model. The authors found that The resampling method's performance with Repeated k-fold Cross-Validation (number=10, repeats=3) has the highest accuracy level at 89.84%. So, In this experiment, the authors had the test setup with the split dataset, 80% for training and training, and 20% for validation, and repeated k-fold Cross-Validation (number=10, repeats=3) for resampling method to evaluate model performance for baseline models.

The current multiclass classification algorithms were selected to evaluate. When the authors measured model's performance in the experiment, the authors found the follows results. First – the raw dataset with 123,156 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor and RF: Random Forest is 100%. Second – the aggregate dataset time duration 1 second with 1,240 samples, the best models performance has accuracy level with RF: Random Forest is 100%.Third – the aggregate dataset time duration 5 seconds with 251 samples, the best

models performance has accuracy level with RF: Random Forest is 99.5%. Fourth – the aggregate dataset time duration 10 seconds with 128 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor is 96.82%. Fifth – the aggregate dataset time duration 15 seconds with 86 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor is 96. 21%.Sixth – the aggregate dataset time duration 20 seconds with 66 samples, the best models performance has accuracy level with LDA: Linear Discriminant Analysis is 98%.Seventh – the aggregate dataset time duration 25 seconds with 54 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor is 96.33%. Moreover, finally, Eight – the aggregate dataset time duration 30 seconds with 46 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor is 93.61%.
The authors organize the content of this paper as follows; Section II Related works, Section III Methodology, Section IV Results and Discussion, Section V Conclusion and Future Works.

2. Related Works

The authors had reviewed the previous works that related to the physical activity recognition area and the ensemble algorithms area as follows.

2.1. Physical Activity Recognition

Saez Y. Et al. [26] mentioned that physical activity is the primary indicator of identifying a quality of person's health. Much Physical activity recognition's research had been studying: Ellis, K. et al. [11] presented the model of life living dataset that much variety and noisy data by using triaxial accelerometers and GPS: Global Positioning System. Morales J. &Akopian D. [19] presented the result of reviewed research papers which related to signals, data capture and preprocessing, unknown on-body locations and orientations, selecting features, activity models, and classifiers, metrics for movement execution, and how to evaluate the usability of a system. Bieber G., Kirste T., and Gaede M. [13] proposed Low sampling rate for physical activity recognition to prevent the power consumption of the sensors with high accuracy rate. Natalie Jablonsky N. et al. [24] investigated the performance of C5.0, decision tree algorithm, with tenfold cross-validation and 80/20 training/test resample methods to multi-sensors and body's location dataset, triaxial accelerometers, with exergame hardware. Baldominos A., Saez Y., &Isasi P. [4] proposed the improvement of activity recognition model by using an activity recognition chain (ARC), optimized genetic algorithms with higher accuracy and lower sensors. They evaluated their model by using leave-one-subject-out cross-validation with an average classification accuracy of about 94%.
Obuchi M. et al. [23] studied the answer rate to ESM: Experience Sampling Method communicates to decrease people's mental burden by checking breakpoints during fitness and sending a message at that time. They found that the improvement was very well at a transition to their activity from "walking" to "sedentary."
Ellis, K. et al. [11] presented the multi-level classifier with random forest algorithm and HMM algorithm for physical activity prediction. Reiss A., Hendeby G., and Stricker D. [7] stated that ConfAdaBoost.M1 algorithm mostly improves the classification performance for more massive and more complex classification datasets of the physical activity recognition. Suarez I. et al. [18] presented the result of improving the recognition accuracy of physical activities, public dataset, by using only the accelerometer. They managed by splitting an accelerometer data to a small- and a high-frequency component with a low-pass filter that provided a new set of features. We used it to a complement to the raw acceleration to reduce the number of sensors needed to recognize physical activities. Majethia R. et al. [25] identified a novel model, crowd-sourced sensor data. They tested their Generic algorithm with a large dataset and found that model accuracy was higher than 95%.

3. Methodology

In this experiment, the physical activity dataset that the authors collected was the multiclass dataset with a label for sitting, standing, walking, and running.
The process of data analysis: (1) Define the problem: load packages, load dataset, split-out validation dataset. (2) Summarize data with descriptive statistics and data visualizations. (3) Prepare data with the technique of data cleaning, feature selection, data transforms. (4) Evaluate algorithms both linear algorithms – LDA: Linear Discriminant Analysis and non-linear algorithms – KNN: k-Nearest Neighbors, SVM: Support Vector Machine, CART: Classification and Regression Trees, RF: Random Forest and compare algorithms, and (5) Finalize Model: predictions on validation dataset, create a standalone model on entire training dataset, and save the model for later use.
Tools: R version 3.3, R-Studio, Windows 10 and caret package. Hardware: Desktop computer CPU i7, RAM 16 GB, GPU GTX1060.

Fig. 1: Workflow of Predict Physical Activity Recognition

Fig. 2: User Interface of Android Mobile Application for Collection the Physical Activity Dataset

4. Result and Discussion

The authors were experimental into two phase: First, Data collection phase – developed an Android Wear application for collecting data and collect data from participants, and Second, Data Analysis phase - manipulated and analyzed a dataset, and measurement models.

4.1. Experiment Setup

Participants - the volunteers who comprise of undergraduate and graduate students from the School of Informatics, Walailak University. There are sixteen volunteers with age range from eighteen-year-olds to twenty-three-year-olds. Among the volunteers, ten are males (62.25%), and six are females (37.50%).
Procedure – First, Data Collection: The authors developed an android application for reading sensors' data from the single tri-axial accelerometer in the Android smartwatch to collect physical activity dataset and managed the physical activity: sitting, standing, walking, and running dataset for every five minutes from the participants. Second, Data Analysis: The authors manipulated and analyzed the dataset with R programming and its library with the step as follows: First, the authors summarized data with descriptive statistics and data visualization. Second, the authors prepared data with data cleaning, feature selection, and data transformation. Third, the authors evaluated algorithms with test options and evaluation metric, spotted check algorithms both linear algorithms and non-linear algorithms to be based line algorithms, and compared algorithms.

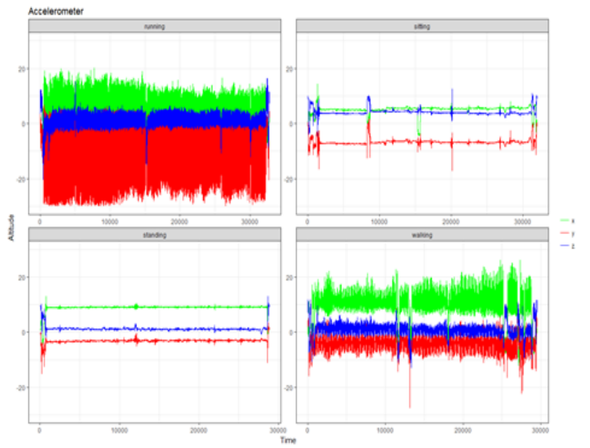

Fig. 3: Graph of Accelerometer (x, y, z) from the Original Dataset

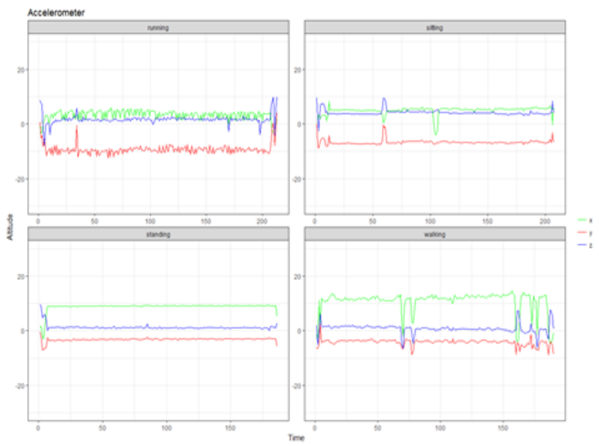

Fig. 4: Graph of Accelerometer (x, y, z) from the Aggregated Dataset with Break One Seconds

4.2. Experiment Results

After, the authors completed the data collection phase. The authors managed and analyzed dataset as follows: prepared data preprocessing and evaluated the algorithms.

1) Preprocessing
First, the authors had got the sensor data loggers (log file) in CSV extension file format from Android wear application. Second, the authors preprocessed the dataset with the following steps: Consolidated all log files into single dataset file with R programming. Cleaned data by removing a feature which had a null value. Filtered only type of sensor: an accelerometer. Aggregated with arithmetic mean of those logs by time and activity as the follows: the completed dataset as an original dataset, the aggregate dataset with break one second, five seconds, ten seconds, fifteen seconds, twenty seconds, twenty-five seconds, and thirty seconds. A visualized that dataset with line graph for each activity. Managed with the proper data format for training and testing the models.

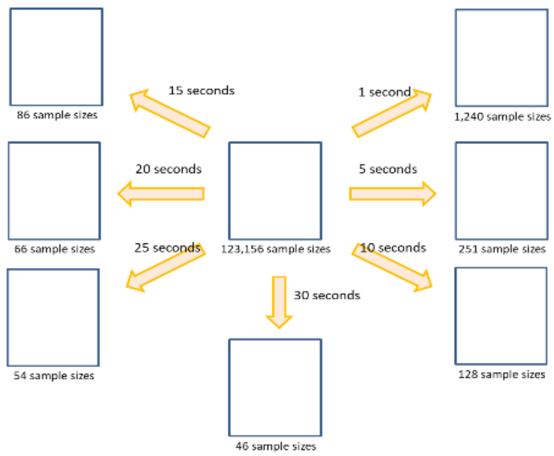

Fig. 5: An Aggregated Dataset with Arithmetic Mean by the Time Duration

2) Split-Out Validation Dataset
In this experiment, we used the split dataset, 80% for training and testing, and 20% for validation, and repeated k-fold Cross-Validation (number=10, repeats=3) for resampling method to evaluate model performance for baseline models

Table 1: The Result of Resampling Method with Naive Bayes

Resampling Methods with Naive Bayes	Accuracy (%)	Kappa (%)
Bootstrap (n=100)	89.05	85.38
k-fold cross-validation without repeats (number=10)	89.67	86.22
Repeated k-fold Cross-Validation (number=10, repeats=3)	89.84	86.43
Leave one out cross-validation	89.86	86.46

From Table 1, The authors found that the result of resampling methods by using Naive Bayes (NB) model performance an accuracy level were Bootstrap (n=100) 89.05%. k-Fold cross-validation without repeats (number=10) 89.67%. Repeated k-fold Cross-Validation (number=10, repeats=3) 89.84%. Leave one out cross-validation 89.86%. In this experiment, the authors selected to use the split dataset, 80% for training and testing, and 20% for validation and repeated k-fold Cross-Validation (number=10, repeats=3) for resampling method to evaluate model performance for baseline algorithms.

3) Evaluate Algorithms
When the authors were developing a predictive model, the authors need to assess the capability of the model on unseen data. Estimating predictive model accuracy need to train with unseen data. In this experiment, the authors presented five approaches for estimating model performance on unseen data. First, split a dataset into train and test subsets. Second, evaluate model accuracy using the bootstrap method. Third, evaluate model accuracy using k-fold cross-validation with and without repeats. Fourth, evaluate model accuracy using leave one out cross-validation.
Model evaluation metric, There is two type of classification models. First, the authors used the metrics: (1) Accuracy and Kappa (2) Area under the curve (AUC), Sensitivity or recall, Specificity or the true negative rate (3) Logarithmic Loss to evaluate the Bi-

nary. Second, the authors measured the metric: (1) Accuracy and Kappa (2) Logarithmic Loss for measuring the performance of the Multiclass class classifications. The root means square error (RMSE) and R2, called as R-squared, measured for Regression models. In this experiment, the collecting dataset was multi-class classification. So we considered model performance with the percentage of Accuracy and Kappa. Compare Algorithms - with baseline algorithms.

Table 2: The comparison of the evaluated algorithms

	Accuracy (%)							
	30	25	20	15	10	5	1	0
LDA	91.94	96.11	98.00	94.92	96.18	97.67	97.68	98.26
CART	67.78	89.67	94.77	93.39	95.86	98.99	87.42	85.15
KNN	93.61	96.33	95.77	96.21	96.82	99.00	99.87	100.00
SVM	91.94	92.83	96.06	95.26	91.61	95.70	97.51	97.59
RF	91.67	93.94	95.89	94.37	96.18	99.50	100.00	100.00
	Kappa (%)							
	30	25	20	15	10	5	1	0
LDA	89.37	94.74	97.27	93.08	94.90	96.89	96.91	97.67
CART	56.75	86.24	89.42	91.05	94.48	98.65	83.08	80.03
KNN	91.59	95.15	92.63	94.88	95.73	98.66	99.82	100.00
SVM	89.37	90.31	94.70	93.55	88.85	94.27	96.68	96.78
RF	89.05	91.77	92.66	92.36	94.92	99.33	100.00	100.00

From table 2, First – the raw dataset with 123,156 samples, the baseline models performance has accuracy level with LDA: Linear Discriminant Analysis is 98.25%, CART: Classification and Regression Tree is 85.15%, KNN: k-Nearest Neighbor is 100%, SVM: Support Vector Machine is 97.58 %, and RF: Random Forest is 100%. Second – the aggregate dataset time duration 1 second with 1,240 samples, the baseline models performance has accuracy level with LDA: Linear Discriminant Analysis is 97.68%, CART: Classification and Regression Tree is 87.41%, KNN: k-Nearest Neighbor is 99.86%, SVM: Support Vector Machine is 97.51%, and RF: Random Forest is 100%.Third – the aggregate dataset time duration 5 second with 251 samples, the baseline models performance has accuracy level with LDA: Linear Discriminant Analysis is 97.67%, CART: Classification and Regression Tree is 98.98%, KNN: k-Nearest Neighbor is 98.99%, SVM: Support Vector Machine is 95.70%, and RF: Random Forest is 99.50%. Fourth – the aggregate dataset time duration 10 second with 128 samples, the baseline models performance has accuracy level with LDA: Linear Discriminant Analysis is 96.18%, CART: Classification and Regression Tree is 95.86%, KNN: k-Nearest Neighbor is 96.82%, SVM: Support Vector Machine is 91.61%, and RF: Random Forest is 96.18%. Fifth – the aggregate dataset time duration 15 second with 86 samples, the baseline models performance has accuracy level with LDA: Linear Discriminant Analysis is 94.92%, CART: Classification and Regression Tree is 93.39%, KNN: k-Nearest Neighbor is 96.21%, SVM: Support Vector Machine is 95.26%, and RF: Random Forest is 94.37%. Sixth – the aggregate dataset time duration 20 second with 66 samples, the baseline models performance has accuracy level with LDA: Linear Discriminant Analysis is 98%, CART: Classification and Regression Tree is 94.77%, KNN: k-Nearest Neighbor is 94.77%, SVM: Support Vector Machine is 96.06%, and RF: Random Forest is 95.89%. Seventh – the aggregate dataset time duration 25 second with 54 samples, the baseline models performance has accuracy level with LDA: Linear Discriminant Analysis is 96.11%, CART: Classification and Regression Tree is 89.67%, KNN: k-Nearest Neighbor is 96.33%, SVM: Support Vector Machine is 92.83%, and RF: Random Forest is 93.94%. Finally, Eight – the aggregate dataset time duration 30 second with 46 samples, the baseline models performance has accuracy level with LDA: Linear Discriminant Analysis is 91.94%, CART: Classification and Regression Tree is 67.78%, KNN: k-Nearest Neighbor is 93.61%, SVM: Support Vector Machine is 91.94%, and RF: Random Forest is 91.67%.

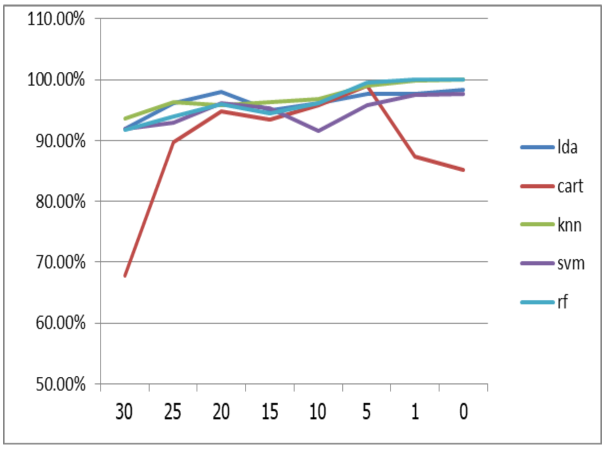

Fig. 6: Graph of the Comparison of the Evaluated Algorithms: Accuracy (%)

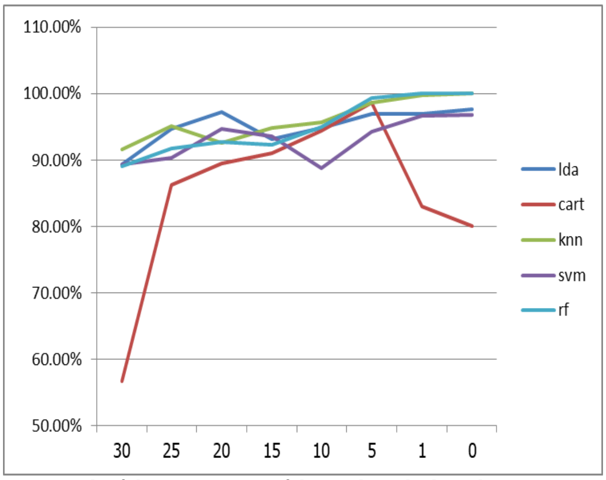

Fig. 7: Graph of the Comparison of the Evaluated Algorithms: Kappa (%)

4) Finalize Model

First, the authors selected Random Forest algorithm to train with 98,526 samples, three predictors, and four classes: 'sitting,' 'standing,' 'walking,' 'running,' resampling: Cross-Validated (10 fold, repeated three times) for the original dataset. The authors found that the final value used for the model was mtry: 2, Accuracy: 100%, and Kappa: 100%.

Second, the authors selected Random Forest algorithm to train with 796 samples, three predictors, and four classes: 'sitting,' 'standing,' 'walking,' 'running,' resampling: Cross-Validated (10 fold, repeated three times) for the original dataset. The authors found that the final value used for the model was mtry: 2, Accuracy: 100%, and Kappa: 100%.

Table 6: Confusion Matrix of Predictions on Validation Dataset with Random Forest (RF) Algorithm for the Original Dataset

Prediction	**Reference**			
	sitting	**standing**	**walking**	**running**
sitting	6396	0	0	0
standing	0	5774	0	0
walking	0	0	5902	0
running	0	0	0	6558

Table 7: Confusion matrix of Predictions on validation dataset with Random Forest (RF) algorithm for the original dataset with break one seconds.

Prediction	Reference			
	sitting	**standing**	**walking**	**running**
sitting	51	0	0	0
standing	0	46	0	0
walking	0	0	47	0
running	0	0	0	53

Then the authors use Random Forest algorithm to predict on validation dataset both the original dataset and the original dataset with break one seconds. The authors had got both the overall statistics: Accuracy: 100% and Kappa: 100%, the confusion matrix as shown in Table 6 and Table 7.

5. Conclusion

When the authors measured model's performance, the authors found the follows results. First – the raw dataset with 123,156 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor and RF: Random Forest is 100%. Second – the aggregate dataset time duration 1 second with 1,240 samples, the best models performance has accuracy level with RF: Random Forest is 100%.Third – the aggregate dataset time duration 5 seconds with 251 samples, the best models performance has accuracy level with RF: Random Forest is 99.5%. Fourth – the aggregate dataset time duration 10 seconds with 128 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor is 96.82%. Fifth – the aggregate dataset time duration 15 seconds with 86 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor is 96. 21%.Sixth – the aggregate dataset time duration 20 seconds with 66 samples, the best models performance has accuracy level with LDA: Linear Discriminant Analysis is 98%.Seventh – the aggregate dataset time duration 25 seconds with 54 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor is 96.33%. Moreover, finally, Eight – the aggregate dataset time duration 30 seconds with 46 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor is 93.61%.

The results indicate that the authors can prepare the dataset with aggregated dataset with arithmetic mean with properly break duration time in seconds with the smaller samples while the model can provide the highest accuracy as the follows: the raw dataset with 123,156 samples, the best models performance has accuracy level with KNN: k-Nearest Neighbor and RF: Random Forest is 100% and the aggregate dataset time duration 1 second with 1,240 samples, the best models performance has accuracy level with RF: Random Forest is 100%.

In the future work, the authors planned to get more accuracy model by adding more features from another sensor, heart rate. Mining data collected from sensors provide valuable result in the physical activity recognition area. The improvement in performance is required especially in the healthcare field. The more increasing of using the wearable device, the broader opportunity in the data mining research area can be.

Acknowledgement

This research was supported by the undergraduate and graduate students from the School of Informatics, Walailak University.

References

[1] AbrielFilios, Sotiris Nikoletseas, and Christina Pavlopoulou. 2015. Efficient Parameterized Methods for Physical Activity Detection using only Smartphone Sensors. In Proceedings of the 13th ACM International Symposium on Mobility Management and Wireless Access (MobiWac '15). ACM, New York, NY, USA, 97-104. DOI=http://dx.doi.org/10.1145/2810362.2810372

[2] Aftab Khan, Sebastian Mellor, Eugen Berlin, Robin Thompson, Roisin McNaney, Patrick Olivier, and Thomas Plötz. 2015. Beyond activity recognition: skill assessment from accelerometer data. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15). ACM, New York, NY, USA, 1155-1166. DOI: http://dx.doi.org/10.1145/2750858.2807534

[3] Alejandro Baldominos, Carmen del Barrio, and YagoSaez. 2016. Exploring the Application of Hybrid Evolutionary Computation Techniques to Physical Activity Recognition. In Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion (GECCO '16 Companion), Tobias Friedrich (Ed.). ACM, New York, NY, USA, 1377-1384. DOI: https://doi.org/10.1145/2908961.2931732

[4] Alejandro Baldominos, YagoSaez, and Pedro Isasi. 2015. Feature Set Optimization for Physical Activity Recognition Using Genetic Algorithms. In Proceedings of the Companion Publication of the 2015 Annual Conference on Genetic and Evolutionary Computation (GECCO Companion '15), Sara Silva (Ed.). ACM, New York, NY, USA, 1311-1318. DOI: http://dx.doi.org/10.1145/2739482.2768506

[5] Allan Stisen, Henrik Blunck, Sourav Bhattacharya, Thor SiigerPrentow, MikkelBaunKjærgaard, AnindDey, Tobias Sonne, and Mads Møller Jensen. 2015. Smart Devices are Different: Assessing and MitigatingMobile Sensing Heterogeneities for Activity Recognition. In Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems (SenSys '15). ACM, New York, NY, USA, 127-140. DOI: http://dx.doi.org/10.1145/2809695.2809718

[6] Andreas Bulling, Ulf Blanke, and Bernt Schiele. 2014. A tutorial on human activity recognition using body-worn inertial sensors. ACM Comput. Surv. 46, 3, Article 33 (January 2014), 33 pages. DOI=http://dx.doi.org/10.1145/2499621

[7] Attila Reiss, GustafHendeby, and Didier Stricker. 2015. A novel confidence-based multiclass boosting algorithm for mobile physical activity monitoring. Personal Ubiquitous Comput. 19, 1 (January 2015), 105-121. DOI: http://dx.doi.org/10.1007/s00779-014-0816-x

[8] Benyue Su, Qingfeng Tang, Jing Jiang, Min Sheng, Ali Abdullah Yahya, and Guangjun Wang. 2016. A novel method for short-time human activity recognition based on improved template matching technique. In Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry - Volume 1 (VRCAI '16), Vol. 1. ACM, New York, NY, USA, 233-242. DOI: https://doi.org/10.1145/3013971.3014004

[9] Dan Morris, T. Scott Saponas, Andrew Guillory, and Ilya Kelner. 2014. RecoFit: using a wearable sensor to find, recognize, and count repetitive exercises. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14). ACM, New York, NY, USA, 3225-3234. DOI: https://doi.org/10.1145/2556288.2557116

[10] Daniel Castro, Steven Hickson, Vinay Bettadapura, Edison Thomaz, Gregory Abowd, Henrik Christensen, and Irfan Essa. 2015. Predicting daily activities from egocentric images using deep learning. In Proceedings of the 2015 ACM International Symposium on Wearable Computers (ISWC '15). ACM, New York, NY, USA, 75-82. DOI: https://doi.org/10.1145/2802083.2808398

[11] Ellis, K., Godbole, S., Kerr, J., &Lanckriet, G. (2014). Multi-sensor physical activity recognition in free-living. Proceedings of the ... ACM International Conference on Ubiquitous Computing .UbiComp (Conference), 2014, 431–440. http://doi.org/10.1145/2638728.2641673

[12] Fabio Aiolli, Matteo Ciman, Michele Donini, and OmbrettaGaggi. 2014. ClimbTheWorld: real-time stairstep counting to increase physical activity. In Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MOBIQUITOUS '14). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, 218-227. DOI: http://dx.doi.org/10.4108/icst.mobiquitous.2014.258013

[13] Gerald Bieber, Thomas Kirste, and Michael Gaede. 2014. Low sampling rate for physical activity recognition. In Proceedings of the 7th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA '14). ACM, New York, NY, USA, Article 15, 8 pages. DOI: https://doi.org/10.1145/2674396.2674446

[14] Giancarlo Fortino, Raffaele Gravina, Wenfeng Li, and Congcong Ma. 2015. Using cloud-assisted body area networks to track people physical activity in mobility. In Proceedings of the 10th EAI

International Conference on Body Area Networks (BodyNets '15). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, Belgium, 85-91. DOI=http://dx.doi.org/10.4108/eai.28-9-2015.2261424

[15] HaodongGuo, Ling Chen, Yanbin Shen, and Gencai Chen. 2014. Activity recognition exploiting classifier level fusion of acceleration and physiological signals. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp '14 Adjunct). ACM, New York, NY, USA, 63-66. DOI=http://dx.doi.org/10.1145/2638728.2638777

[16] Heba Aly and Mohamed A. Ismail. 2015. ubiMonitor: intelligent fusion of body-worn sensors for real-time human activity recognition. In Proceedings of the 30th Annual ACM Symposium on Applied Computing (SAC '15). ACM, New York, NY, USA, 563-568. DOI: http://dx.doi.org/10.1145/2695664.2695912

[17] Henrik Blunck, Sourav Bhattacharya, Allan Stisen, Thor SiigerPrentow, MikkelBaunKjærgaard, AnindDey, Mads Møller Jensen, and Tobias Sonne. 2016. ACTIVITY RECOGNITION ON SMART DEVICES: Dealing with diversity in the wild. GetMobile: Mobile Comp. and Comm. 20, 1 (July 2016), 34-38. DOI: http://dx.doi.org/10.1145/2972413.2972425

[18] Isabel Suarez, Andreas Jahn, Christoph Anderson, and Klaus David. 2015. Improved activity recognition by using enriched acceleration data. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15). ACM, New York, NY, USA, 1011-1015. DOI: http://dx.doi.org/10.1145/2750858.2805844

[19] Jafet Morales, David Akopian, Physical activity recognition by smartphones, a survey, Biocybernetics and Biomedical Engineering, Volume 37, Issue 3, 2017, Pages 388-400, ISSN 0208-5216, https://doi.org/10.1016/j.bbe.2017.04.004

[20] Jeffrey W. Lockhart and Gary M. Weiss. 2014. Limitations with activity recognition methodology & data sets. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp '14 Adjunct). ACM, New York, NY, USA, 747-756. DOI=http://dx.doi.org/10.1145/2638728.2641306

[21] Jie Wan, Michael J. O'grady, and Gregory M. O'hare. 2015. Dynamic sensor event segmentation for real-time activity recognition in a smart home context. Personal Ubiquitous Comput. 19, 2 (February 2015), 287-301. DOI: http://dx.doi.org/10.1007/s00779-014-0824-x

[22] Lijie Xu and Kikuo Fujimura. 2014. Real-Time Driver Activity Recognition with Random Forests. In Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '14). ACM, New York, NY, USA, , Article 9 , 8 pages. DOI=http://dx.doi.org/10.1145/2667317.2667333

[23] Mikio Obuchi, Wataru Sasaki, Tadashi Okoshi, JinNakazawa, and Hideyuki Tokuda. 2016. Investigating incorruptibility at activity breakpoints using smartphone activity recognition API. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16). ACM, New York, NY, USA, 1602-1607. DOI: https://doi.org/10.1145/2968219.2968556

[24] Natalie Jablonsky, Sophie McKenzie, Shaun Bangay, and Tim Wilkin. 2017. Evaluating sensor placement and modality for activity recognition in active games. In Proceedings of the Australasian Computer Science Week Multiconference (ACSW '17). ACM, New York, NY, USA, Article 61, 8 pages. DOI: https://doi.org/10.1145/3014812.3014875

[25] Rahul Majethia, AkshitSinghal, Lakshmi Manasa K, KunchaySahiti, Shubhangi Kishore, and Vijay Nandwani. 2016. AnnoTainted: Automating Physical Activity Ground Truth Collection Using Smartphones. In Proceedings of the 3rd International on Workshop on Physical Analytics (WPA '16). ACM, New York, NY, USA, 13-18. DOI: http://dx.doi.org/10.1145/2935651.2935653

[26] Saez, Y., Baldominos, A., &Isasi, P. (2017). A Comparison Study of Classifier Algorithms for Cross-Person Physical Activity Recognition. Sensors (Basel, Switzerland), 17(1), 66. http://doi.org/10.3390/s17010066

[27] Sara Khalifa. 2017. Energy-efficient human activity recognition for self-powered wearable devices. In Proceedings of the Australasian Computer Science Week Multiconference (ACSW '17). ACM, New York, NY, USA, Article 78, 2 pages. DOI: https://doi.org/10.1145/3014812.3018840

[28] Thomas Phan. 2014. Improving activity recognition via automatic decision tree pruning. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp '14 Adjunct). ACM, New York, NY, USA, 827-832. DOI=http://dx.doi.org/10.1145/2638728.2641310

[29] Vijay Rajanna, Raniero Lara-Garduno, Dev JyotiBehera, KarthicMadanagopal, Daniel Goldberg, and Tracy Hammond. 2014. Step up life: a context aware health assistant. In Proceedings of the Third ACM SIGSPATIAL International Workshop on the Use of GIS in Public Health (HealthGIS '14), Daniel W. Goldberg, Ori Gudes, and YaronKanza (Eds.). ACM, New York, NY, USA, 21-30. DOI: https://doi.org/10.1145/2676629.2676636

[30] Wenchao Jiang and Zhaozheng Yin. 2015. Human Activity Recognition Using Wearable Sensors by Deep Convolutional Neural Networks. In Proceedings of the 23rd ACM international conference on Multimedia (MM '15). ACM, New York, NY, USA, 1307-1310. DOI: https://doi.org/10.1145/2733373.2806333

[31] Xing Su, Hanghang Tong, and Ping Ji. 2014. Accelerometer-based Activity Recognition on Smartphone. In Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management (CIKM '14). ACM, New York, NY, USA, 2021-2023. DOI=http://dx.doi.org/10.1145/2661829.2661836