Human activity recognition by wearable sensors in the smart home control problem

I.Y. Nebogatikov and I.P. Soloviev
Saint Petersburg State University, Saint Petersburg, 199034, Russia

E-mail: st049100@student.spbu.ru, i.soloviev@spbu.ru

Abstract. In this work we analyze and compare machine learning methods for recognizing human activity in the context of smart home by using data obtained from an optical heartbeat sensor and an accelerometer embedded in a smart watch, and find a number of activity classes that can be predicted. The conclusion is made that for such type of problems the random forest method with about 10 classes shows the best results.

1. Introduction
The "smart home" research direction includes many different tasks, the range of which is not limited by monitoring and control of household appliances, but also implies the creation of rather complex strategies for supporting various types of inhabitants of this house. The strategies rely on monitoring both the devices and the activity of the inhabitant, in particular on the technology using wearable sensors.

In general, we interpret the task of a smart home as a response to a social demand to create a comfortable and safe environment for humans, possibly with their pets. This presupposes the creation of a whole system of social and technological support for the inhabitants of a smart home. If we proceed from the “average” person and “standard” housing, then the support strategy can be reduced to monitoring the presence of an inhabitant and detecting some type of his activity from a small set of possibilities, and standard technical control over the state of household appliances. At the same time, additional assumptions about the characteristics of an inhabitant of the house or the house itself may require the creation of a flexible control and support strategy, an expanding of the range of household appliances, a system of communication with medical and emergency services, etc. There are various scenarios that require more flexible control. These include preventing accidents when potentially dangerous electrical devices are left on for a long time, saving energy and water, monitoring the condition of a person who may need medical attention, and others. Further, we clarify exactly which tasks we solve in this work.

Only some problems, such as smoke or gas monitoring, can be solved using information from the sensors. However, for more complex scenarios we need to know what an inhabitant is doing now. The monitoring parameters from devices make it possible to recognize human activity type in cases when patterns of activities are previously defined. Supporting strategies rely on static device management and wearable devices. The term static devices means not only sensors in household appliance, but also camera systems with video and sound recording. Vision-based algorithms show a good recognition ability [1] but have some drawbacks. For example, one drawback is the blind spots of cameras; it could
cost a lot to reduce the number of blind spots with additional cameras. Wearable devices may form a complex system of electrocardiographs, pulse oximeters and other medical devices. Smart watches and fitness trackers with accelerometer and heart rate monitoring sensors are the most commonly used wearable devices. In some cases using data from more than one device may lead to incorrectly detected activity classes when sensors are inconsistent [2]. For example, light measurement from Bluetooth beacons can change because of the weather but not human activity.

Data about inhabitants’ activities can be used for programming user scenarios that depend on current activity types, for example “if inhabitants are sleeping, then all the lights and taps should be checked and turned off”. Another example of application is related to inhabitants’ health monitoring. Based on activity data we may detect circadian rhythms for using in Chrono therapeutic healthcare [3]. Long periods of inactivity or very high activity could be a cause of insomnia [4], and we could detect these situations and warn users about the influence on their health. These problems became more actual in recent years, because the number of remote employees increases, and the largest part of their activities is related to actions in the home.

There are two main approaches to recognizing human activity: classification and recognition of predefined types of activities; and clustering with automatic pattern extraction.

In this paper, we compare machine learning methods of human activity recognition in the scope of the smart home control problem. The processed data are in open access. We focus on the classification problem, because classified activities can be used in the automatization of the smart home automatic scenarios programming. For simplicity, we consider the model with one person in the home.

2. Related works
Activity monitoring and recognition with wearable devices are the subject of interest for many researches.

In [5], the authors used a wristband with heart rate and accelerator sensors to detect home-specific activities. They selected 5 base types of activities: standing, sitting, householding and cycling with 2 intensity levels. They applied Random Forest (RF) and Support Vector Machines (SVM) classifiers to detect the type of activity. The average recognition accuracy with leave-one-subject-out (LOSO) cross-validation were 89.2% and 85.6% respectively, F1-score values for RF and SVM were 0.89 and 0.86 respectively.

In [6], the authors used a modification of the SVM algorithm with hierarchy to detect whether a person sits, stands, walks or runs. Hierarchical classifier contains two steps, on the first one the basic type of activity (keeping still or moving) is recognized. Then the same sensor values are used in the second classifier for more accurate definition. With a high sampling rate, more than 5 Hz, the accuracy of activity recognition is above 99%.

In [7], the authors used combinations of Naive Bayes, random tree and k-nearest neighbors (KNN) classifiers to recognize 13 activity types with wristband and smartphone data. F-measure of the classifier is near to 1. The authors state that complex activities can be recognized in a reliable way by using hierarchical classification and adaptive segmentation window size (2–30 s).

3. Data description
In the presented work we used open access datasets on human activity data.

3.1. Heart rate and breathing data [8]. The dataset contains 5 activity types: running, walking and inactivity, climbing and descending stairs. There are 87235 rows for basic activities, and 41459 more measurements for climbing and descending. The data do not have timestamps, and the measurements could not be ordered and filtered. In the considered smart home model we do not collect breathing data, and use heart rate data only.

3.2. An Open Dataset for Human Activity Analysis [9] contains records about 10 types of activity collected during 3 days: cooking, eating, working at computer, running, sitting, sleeping, walking,
playing video games, watching television, and training. The number of records is 4742. The dataset contains information from optical heart rate sensor and accelerometer. Triaxial accelerometer output is three numbers for each axis. Although this dataset contains the information about the measurement time, it is difficult, as in previous case, to apply filters because the intervals between some measurements are more than several hours.

4. The description of machine learning methods
This paper discusses classification algorithms and metrics used for solving related problems, mentioned in [2, 5, 6, 7], AdaBoost and a multilayer neural network.
A more detailed description and the parameters are presented in the following subsections.

4.1. The description of algorithms
We use Python and Scikit learn library to process data. The machine learning algorithms and their parameters are:
- Multi-layer Perceptron classifier with 100 layers, ReLU activation function and Adam solver [10];
- K nearest neighbors with K=10 [11];
- Random forest with 100 decision trees [12];
- Naïve Bayes classifier [13];
- AdaBoost with Decision Tree Classifier as a base algorithm [14]; this classifier creates copies of base classifiers with adjusted weights that focus more on examples that were incorrectly labeled by previous classifiers;
- Support vector machines [15].

4.2. The description of metrics
- Accuracy: the ratio of correctly labeled examples to the total set;
- Precision: the ratio of correctly labeled examples to the total number of labeled examples;
- ROC AUC: area under “error” (receiver operating characteristic) curve.

5. Analysis
The selected datasets were firstly divided into training and test sets with a ratio of 85-15.

Table 1. The values of metrics with heartbeat sensor for 5 activity types

| Activity type     | Random forest | Support vector machines |
|-------------------|---------------|-------------------------|
|                   | Precision     | AUC                     | Precision     | AUC           |
| Climbing stairs   | 0.61919       | 0.564738                | 0.160682      | 0.609572      |
| Inactivity        | 0.487542      | 0.926043                | 0.268873      | 0.930411      |
| Run               | 0.599251      | 0.771685                | 0.639908      | 0.789139      |
| Descending stairs | 0.243863      | 0.595384                | 0.265470      | 0.614261      |
| Walk              | 0.370785      | 0.624040                | 0.392319      | 0.636381      |

As mentioned above, random forest and support vector machines methods have better quality than other classifiers when detecting types of activity. These methods were therefore used to process heartbeat data. The accuracy values are 0.438487 and 0.386946 for random forest and SVM respectively. Since accuracy indicates the quality of the classifier as a whole, to study how the classifier performs on
each activity class, precision and ROC AUC values should be calculated for each activity class. Precision and ROC AUC metric values are shown in Table 1.

Precision for inactivity is significantly lower for the SVM classifier than for the random forest algorithm, because heart rate for walking and inactivity states have similar value ranges and data for these classes are not separated enough to be linearly separable. Other activity types are more separable and SVM shows better precision results. The ROC AUC values for SVM are higher than the corresponding values for random forest because these values are calculated on training sets that contain more data and outliers have less impact to values calculation.

The climbing and descending stairs activities data has been excluded to increase data separability and cover only basic activities. In this case, the accuracy value is 0.666590 for the random forest and 0.614473 for the support vector machines. As in the case of 5 activity classes, the accuracy of the random forest classifier exceeds the accuracy of SVM classifier. A more detailed assessment of the quality of classifiers, calculated for each activity class separately, using precision and ROC AUC values is shown in Table 2.

Table 2. The values of metrics with heartbeat sensor for 3 activity types

| Activity type | Random forest | Support vector machines |
|---------------|---------------|-------------------------|
|               | Precision | AUC | Precision | AUC |
| Inactivity    | 0.616322 | 0.923518 | 0.386822 | 0.931720 |
| Run           | 0.754862 | 0.802214 | 0.805437 | 0.816909 |
| Walk          | 0.612167 | 0.724526 | 0.633569 | 0.730616 |

According to Table 2, inactivity recognition accuracy is lower for support vector machine classifier, the same is true for 5 classes processed. The accuracy of recognizing inactivity is lower for the support vector machine classifier, the same could be observed for the case with 5 classes. Random forest algorithm is more sensible for features difference and shows greater precision values. The detection quality of running is significantly increased for 3 activity types, because the case with 5 activities had two similar activities: climbing and descending stairs.

The second dataset was processed by all the classification algorithms mentioned above. Accuracy values are shown in Table 3.

Table 3. Recognition accuracy values for An Open Dataset for Human Activity Analysis dataset processed with different sensor combinations

| Algorithm              | Accuracy (heartbeat sensor) | Accuracy (accelerometer) | Accuracy (both sensors) |
|------------------------|-----------------------------|--------------------------|-------------------------|
| Neuron net             | 0.286517                    | 0.351124                 | 0.369382                |
| K nearest neighbors    | 0.237360                    | 0.419944                 | 0.433989                |
| Random forest          | 0.147472                    | 0.428371                 | 0.501404                |
| Naive Bayes            | 0.290730                    | 0.271067                 | 0.300562                |
| AdaBoost               | 0.308989                    | 0.275281                 | 0.289326                |
| Support vector machines| 0.116573                    | 0.137640                 | 0.143258                |
Table 4. The values of metrics for An Open Dataset for Human Activity Analysis dataset processed with random forest

| Activity type  | Heart beat | | Accelerometer | | |
|---------------|-----------|---|---------------|---|
|               | Precision | AUC | Precision | AUC |
| Cooking       | 0.171429  | 0.750000 | 0.250000 | 0.792779 |
| Eating        | 0.189189  | 0.683665 | 0.393939 | 0.780057 |
| On computer   | 0.250000  | 0.565559 | 0.457364 | 0.770287 |
| Running       | 0.594595  | 0.879178 | 0.357143 | 0.815145 |
| Sitting       | 0.290323  | 0.593627 | 0.414062 | 0.720457 |
| Sleeping      | 0.265823  | 0.693415 | 0.822222 | 0.911407 |
| Training      | 0.064935  | 0.639364 | 0.380952 | 0.799205 |
| Video games   | 0.265823  | 0.607804 | 0.384615 | 0.771417 |
| Walking       | 0.800000  | 0.542663 | 0.374172 | 0.678598 |
| TV            | 0.081081  | 0.609222 | 0.642857 | 0.869282 |

Table 5. The values of metrics An Open Dataset for Human Activity Analysis dataset processed with random forest

| Activity type   | Both sensors | | |
|-----------------|--------------|---|
|                 | Precision    | AUC |
| Cooking         | 0.400000     | 0.893726 |
| Eating          | 0.370370     | 0.839258 |
| On computer     | 0.544715     | 0.818275 |
| Running         | 0.727273     | 0.435128 |
| Sitting         | 0.485714     | 0.515893 |
| Sleeping        | 0.897436     | 0.897072 |
| Training        | 0.500000     | 0.636019 |
| Video games     | 0.400000     | 0.576913 |
| Walking         | 0.452962     | 0.520933 |
| TV              | 0.611111     | 0.475904 |
It can be seen from the tables above that 3 variants of sensors were considered: the heartbeat sensor, the accelerometer, and both of them.

This dataset has a significantly smaller size than the previous one and is more affected by outliers in the training set; hence the K nearest neighbors, AdaBoost, Naive Bayes and neural network classifiers show higher accuracy when using only the heart rate sensor. Only random forest shows an increase in accuracy when using both sensors. Support vector machine classifier accuracy slightly increases because of complication data structure. The accuracy of the support vector machine classifier does not increase significantly due to the complexity of the data structure.

Precision and ROC AUC values for random forest classifier are shown in tables 4 and 5. Precision values for basic activities when using only heart beat sensor are higher than for cooking or eating; similar results were obtained when processing the first dataset, where the precision of the recognition of climbing stairs is significantly less than walking. Some types of inactivity, like sleeping or watching television, can be detected using only the accelerometer. However, using both sensors increases these results and reduces the differences between precision for the same type of activity with different sensors.

For the random forest when using both sensors, features importance was calculated. For this algorithm, importance is calculated as the ratio of the number of times a feature is used for splitting to the number of all splits in all trees [16]:

- 0.479728 for heartbeat sensor;
- 0.520272 for accelerometer.

These values are close to 0.5 and depend on dataset separation, so the values can be assumed to be equal.

To summarize, accuracy values for random forest and SVM classifiers for both datasets are shown in Table 6. We see that only basic types of activities can be recognized with one sensor. To improve accuracy, we need to reduce the number of classes, remove derived or related types, or add more sensors to the model.

Table 6. Accuracy values for both datasets with random forest and SVM classifiers

| Classes            | Random Forest | Support Vector Machines |
|--------------------|---------------|-------------------------|
| 5 classes (heart rate) | 0.438487      | 0.386946                |
| 3 classes (heart rate) | 0.666590      | 0.614473                |
| 10 classes (heart rate) | 0.310393      | -                       |
| 10 classes (accelerometer) | 0.428370      | -                       |
| 10 classes (both sensors) | 0.512640      | -                       |
6. The description of the proposed data collection method

As shown in [2, 7], sensors for activity recognition should be carefully selected to reduce the influence of unrelated features on the classification result. A high quality of classification can be reached using wearable devices such as smartphones and wristbands. We can get accelerometer data with steps count activity, because most vendors calculate activity on the wristband and do not send raw data with Bluetooth.

We can get the accelerometer value in the form of the number of steps, because most manufacturers program the devices to calculate the activity level on the wristband and do not send raw data via Bluetooth.

Using a smartphone, we can get the location of the inhabitant in the home. It can be calculated using the signal quality indicators (received signal strength indication — RSSI) of Wi-Fi or Bluetooth connection with smartwatches [17].

If our goal is health monitoring and, for example, quality of sleep, a wristband can calculate sleep phases and their durations.

7. Conclusions

In this work, machine learning algorithms for human activity recognition in the smart home problem were compared. Activity data for a single inhabitant inside the home were classified. The comparison of results shows that using only one sensor leads to recognition of a few classes of activity. To detect about ten classes both sensors should be used.

Built models could not be used in a real smart home environment because of the quality of open datasets.

In future work we propose to collect a sufficient amount of data to be able to filter them, and to increase the number of activity types. Random forest classifier is proposed to process data, as in many scenarios it shows robust accuracy which exceeds the classification accuracy of other algorithms for different task parameters.

References

[1] Al-Faris M, Chiverton J, Ndzi D, Ahmed A 2020 A Review on Computer Vision-Based Methods for Human Action Recognition Journal of Imaging 6 46

[2] Kirienko A and Soloviev I 2017 Human Behaviour Analysis in Context of Smart Environment Automation Computer tools in education 1 15–29

[3] Huang Q, Cohen D, Komarzynski S, Li X, Innominato P, Lévi F and Finkenstädt B 2018 Hidden Markov models for monitoring circadian rhythmicity in telemetric activity data J. R. Soc. Interface 15 20170885

[4] Hartescu I and Morgan K 2018 Regular physical activity and insomnia: An international perspective Journal of Sleep Research 28 30117220

[5] Mehrang S, Pietila J, Tolonen J, Helander E, Jimison H, Pavel M and Korhonen I 2017 Human Activity Recognition Using A Single Optical Heart Rate Monitoring Wristband Equipped with Triaxial Accelerometer EMBEC 65 587–90

[6] Tang T, Zheng L, Weng S, Peng A and Zheng H 2018 Human Activity Recognition with Smart Watch based on H-SVM International Conference on Frontier Computing 422 179–86

[7] Shoaib M, Bosch S, Incel O, Scholten H and Havinga P 2016 Complex Human Activity Recognition Using Smartphone and Wrist-Worn Motion Sensors Sensors 16 426

[8] Bouarada O. heart rate and breathing data [Electronic resource]. Access mode: https://www.kaggle.com/onnnsbrd/heart-rate-and-breathing-data?select=final_ecg_data.csv

[9] Jafarnejad S, An Open Dataset for Human Activity Analysis. [Electronic resource]. Access mode: https://www.kaggle.com/sasanj/human-activity-smart-devices

[10] Diederik P and Jimmy B 2015 Adam: A Method for Stochastic Optimization Preprint arXiv:1412.6980

[11] Cunningham P and Delany S J 2020 k-Nearest neighbour classifiers Preprint arXiv:2004.04523
[12] Oshiro T M, Perez P S and Baranauskas J A 2012 How Many Trees in a Random Forest? Lecture Notes in Computer Science 7376 154–68
[13] Rish I 2001 An empirical study of the naive Bayes classifier IJCAI 2001 workshop on empirical methods in artificial intelligence 3 41–6
[14] Freund Y and Schapire R E 1999 A Short Introduction to Boosting Journal of Japanese Society for Artificial Intelligence 2 771–80
[15] Evgeniou T and Pontil M 2001 Support Vector Machines: Theory and Applications Machine Learning and Its Applications 2049 249–57
[16] Menze B H, Kelm B M, Masuch R, Himmelreich U, Bachert P, Petrich W and Hamprecht F A 2009 A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data BMC Bioinformatics 10 213
[17] Correa J, Katz E, Collins P, Griss M 2018 Room-Level Wi-Fi Location Tracking Carnegie Mellon University