**Contemporary Engineering Sciences, Vol. 10, 2017, no. 7, 345 - 352**
**HIKARI Ltd, www.m-hikari.com**
**https://doi.org/10.12988/ces.2017.7218**

# User Behavior Classification Based on Smart Watch and Machine Learning Algorithm

**Min-Cheol Kwon**

Dept. of Secured Smart Electric Vehicle
Kookmin University, Seoul, Korea 136-702

**Sunwoong Choi***

Dept. of Secured Smart Electric Vehicle
Kookmin University, Seoul, Korea 136-702
*Corresponding author



**Abstract**

Recently, many wearable devices have been developed as IoT technology grows. Among them, smart watch is the friendliest wearable device in daily lives. Many companies are trying to improve the device or system to provide personal service as user's behavior. This paper proposes an user behavior classification system using smart watch and machine learning algorithm to provide personal service with wearable devices. Sensing data from accelerometer of smart watch is collected, and then classification is implemented by Two-Class Support Vector Machine, Multi-Class Logistic Regression, Multi-Class Decision Forest. We show that the prediction accuracy is more than 90%.

**Keywords**: Wearable, Internet of Things(IoT), Machine Learning

## 1 Introduction

Recently, as technology of IoT(Internet of Things) has been growing enormously [1], [2], various products of IoT or wearable devices are developed and used in daily lives. The friendliest wearable device is the smart watch, which has integrated IoT technology into the watch. The smart watch is widely used.

Machine learning has been developed [3], [4] for more than decades. Machine learning is a field of artificial intelligence to improve algorithms for computers to learn.

There are many researches to classify human activities using machine learning. Most of them classify a few distinct and active human activities such as walking, running, going upstairs, and going downstairs, for health care services [5], [6], [7]. [8] studies on recognition of a patient's specific behavior for therapeutic purposes or detecting diseases.

In this paper, we focus on inactive human behaviors as compared with the previous studies which focused on active human behaviors. Inactive behavior means that behaviors on desk such as office work(writing documents, or coding), studying(reading a book, or writing with a pencil), and playing computer game.

Accelerometer sensing data is periodically collected from smart watch, and then classification is implemented by Two-Class Support Vector Machine, Multi-Class Logistic Regression, Multi-Class Decision Forest.

The rest of this paper organized as follows. The system is introduced in section 2, process of classifying user behavior is presented in section 3. Performance evaluation is given in section 4. The section 5 is the conclusion.

## 2 System Overview

The proposed system is divided into 4 sections, smart watch, smart phone, server, and machine learning.

### 2.1 System Structure

The System Structure is Fig 1. The smart watch collects the sensor data of accelerometer and information from user's activity, and then sends the data to a smart phone. The smart phone sorts out the data by format and sends it to server. The server performs preprocessing received data. The machine learning model learns to categorize user behavior from preprocessed data.

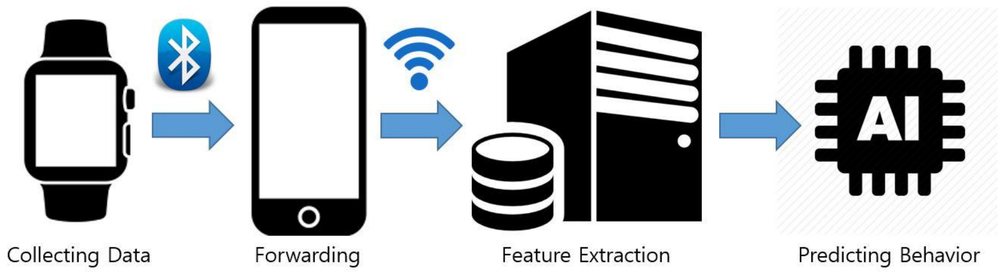


Fig 1. System Structure

2.2 Equipment

In the experiment, apple watch and iPhone is used as experimental devices. The operating system of the server is window 10. The ML studio[9] is utilized as a machine learning tool.

2.3 System Scenario

The user wears a smart watch on user’s wrist. Left-handers wear this to the left, right-handers to the right. When the user wears the watch, it measures the change of accelerometer and transmits the change to the server. The server classifies received data as user behavior.

## 3 User Behavior Classifier

3.1 Data Collection

The user can collect the data from the smart watch. It works in three steps for collecting data by wearing a smart watch. The first step is to get the information of user’s activity. User can specify current activity by just pressing a button. Then the sensor data is labelled to the specified activity in real time. With this labels, the supervised learning is enabled. The second step is to get accelerometers sensor data periodically. After user pressing the button, smart watch starts to get the data from accelerometers. The smart watch gets X-axis, Y-axis Z-axis data of accelerometer ten times in every second. Until the user presses its finish button, the smart watch continuously collects the data. The third step is to send the data to the smart phone. The data set is defined as X-axis, Y-axis, Z-axis data of accelerometer and information of user activity.

The smart phone receives the data from smart watch through Bluetooth and forward the data to server through WiFi.

3.2 Feature Extraction

The server extracts the feature for machine learning from the received dataset. The extracting process is shown as follows. The duration of each epoch in this study is set to 10 seconds. The server calculates its mean, standard deviation, and variance of values of the X-axis, Y-axis and Z-axis of the accelerometer during 10 seconds and extracts a total of 9 features for machine learning. Finally, server gives the preprocessed data to the machine learning model.

3.3 Classification

In this paper, three kinds of machine learning techniques are applied to obtain the results. They are Multi-class Logistic Regression model, Two-class Support Vector Machine model and Multi-class Decision Forest model. We use ML studio of Microsoft as machine learning tool. ML studio is a powerful and simple tool. The ML studio performs machine learning based on data processed by server.

## 4 Performance Evaluation

In this section, we evaluate the proposed system. Table 1 shows the number of data set used for machine learning. The number of dataset with office work features is 9392. The number of dataset with study features is 8666. The number of dataset with game features is 6168. Therefore, the total number of feature dataset is 24226. This data set is randomly divided into a training set and a validation set as 8:2 ratio.

Table 1 Dataset Overview

| **Activity Type** | Office Work | Study | Game | Total |
|---|---|---|---|---|
| **Number of Instances** | 9392 | 8666 | 6168 | 24226 |

4.1 Two-Class Support Vector Machine

Fig 2 shows the confusion matrix of Two-Class Support Vector Machine. The Y-axis of confusion matrix is an actual stage and the X-axis of confusion matrix is the stage predicted by the machine learning model. The ratio that model predicts the dataset of an actual office work stage as office work stage is 78.3%. The model confuses the office work with the game. The other ratios are almost accurate, as each of them is 97.4% and 91.0%. The overall accuracy of this model is 90.07% in table 2.

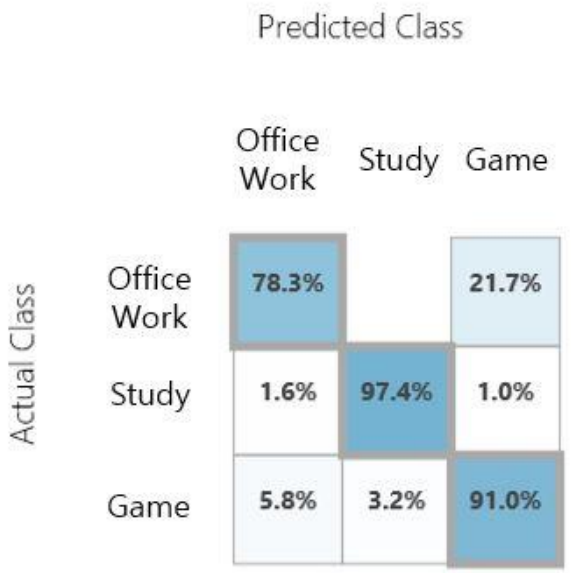


Fig 2 Two-Class SVM Confusion Matrix

Table 2 Overall Accuracy

| Machine Learning Model | Overall Accuracy | Precision | Recall |
|---|---|---|---|
| Two-Class Support Vector Machine | 90.07% | 89.96% | 88.92% |
| MultiClass Logistic Regression | 91.55% | 92.15% | 89.96% |
| MultiClass Decision Forest | 99.93% | 99.93% | 99.93% |

4.2 Multi-Class Logistic Regression

Fig 3 shows the confusion matrix of Multi-Class Logistic Regression. It is slightly better than Two-class Support Vector Machine. Similarly, SVM mainly confuses office work with the game. The ratio of matching the office work stage is only 76.2%. The model usually predicts as the game stage, if the model does not fully match with work dataset. The overall accuracy is 91.55% in table 2.

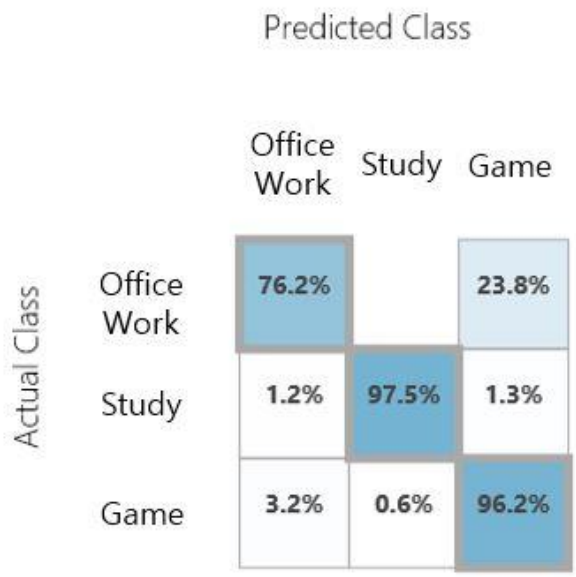


Fig 3 Multi-Class Logistic Regression Confusion Matrix

4.3 Multi-Class Decision Forest

Fig 4 shows the confusion matrix of Multi-Class Decision Forest model. The ratio that model predicts accurate dataset, actual office work stage as office work stage, an actual study stage as study stage, an actual game stage as game stage, are 99.9%. The overall accuracy is 99.93% in table 2. This model has the best accuracy among three models.

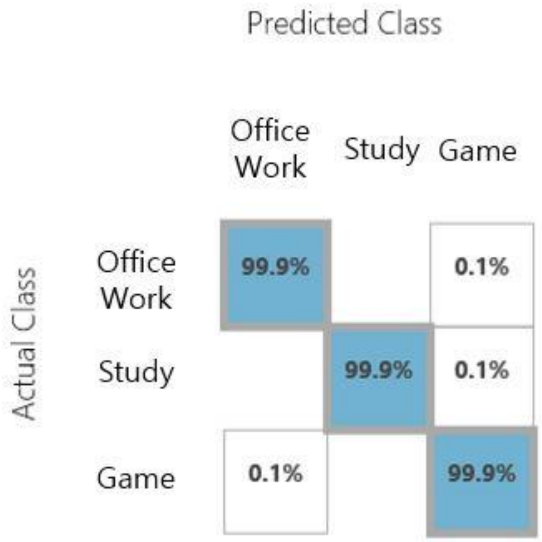


Fig 4 Multi-Class Decision Forest Confusion Matrix

4.4 Discussion

According to results, human behavior classification is possible even with the data sampling in 10 seconds. It is also possible to classify not only active behaviors, but also inactive behaviors, such as office work, studying, or playing computer game. Among the various machine learning models, Multi-Class Decision Forest model is the best to categorize the inactive behavior. Still, the other models are good as well.

## 5 Conclusion

In this paper, user behavior is categorized by smart watch sensor and machine learning. Although previous work classifies the activities in the area of health care, this work presents the possibility that various human activities in daily lives can be classified by machine learning. Being able to distinguish inactive behaviors, as well as active behaviors, it is possible to create an environment where the device of IoT can provide wider services. Moreover, applying more sensors, other than accelerometers, will enable other wearable devices or other IoT devices to be more accurate and classify more human activities.

**Acknowledgements.** This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No. B0717-16-0065, Germanium on Silicon based Sensor which covers Visible spectrum to Short Wavelength Infra-Red range (400-1600nm) for Materials Identification and Application System Development)

## References

[1] Qi-bo Sun et al., Internet of Things: Summarize on Concepts, Architecture and Key Technology Problem, *Journal of Beijing University of Posts and Telecommunications*, **3** (2010), no. 3, 1-9.

[2] Sungmee Park, Sundaresan Jayaraman, Enhancing the quality of life through wearable technology, *IEEE Engineering in Medicine and Biology Magazine*, **22** (2003), no. 3, 41-48. https://doi.org/10.1109/memb.2003.1213625

[3] Yann Lecun, Yoshua Bengio, Geoffrey Hinton, Deep learning, *Nature*, **521** (2015), no. 7553, 436-444. https://doi.org/10.1038/nature14539

[4] Jürgen Schmidhuber, Deep learning in neural networks: An overview, *Neural Networks*, **61** (2015), 85-117. https://doi.org/10.1016/j.neunet.2014.09.003

[5] Xizhe Yin et al., Human activity detection based on multiple smart phone sensors and machine learning algorithms, *2015 IEEE 19th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, IEEE, (2015), 582-587. https://doi.org/10.1109/cscwd.2015.7231023

[6] Yongjin Kwon, Kyuchang Kang, Changseok Bae, Unsupervised learning for human activity recognition using smartphone sensors, *Expert Systems with Applications*, **41** (2014), no. 14, 6067-6074. https://doi.org/10.1016/j.eswa.2014.04.037

[7] Xizhe Yin, Weiming Shen, Xianbin Wang, Incremental clustering for human activity detection based on phone sensor data, *2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD),* IEEE, (2016), 35-40. https://doi.org/10.1109/cscwd.2016.7565959

[8] F. Foerster, M. Smeja, J. Fahrenberg, Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring, *Computers in Human Behavior*, **15** (1999), no. 5, 571-583. https://doi.org/10.1016/s0747-5632(99)00037-0

[9] https://studio.azureml.net/

**Received: February 23, 2017; Published: March 16, 2017**