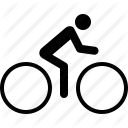
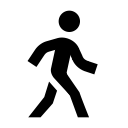
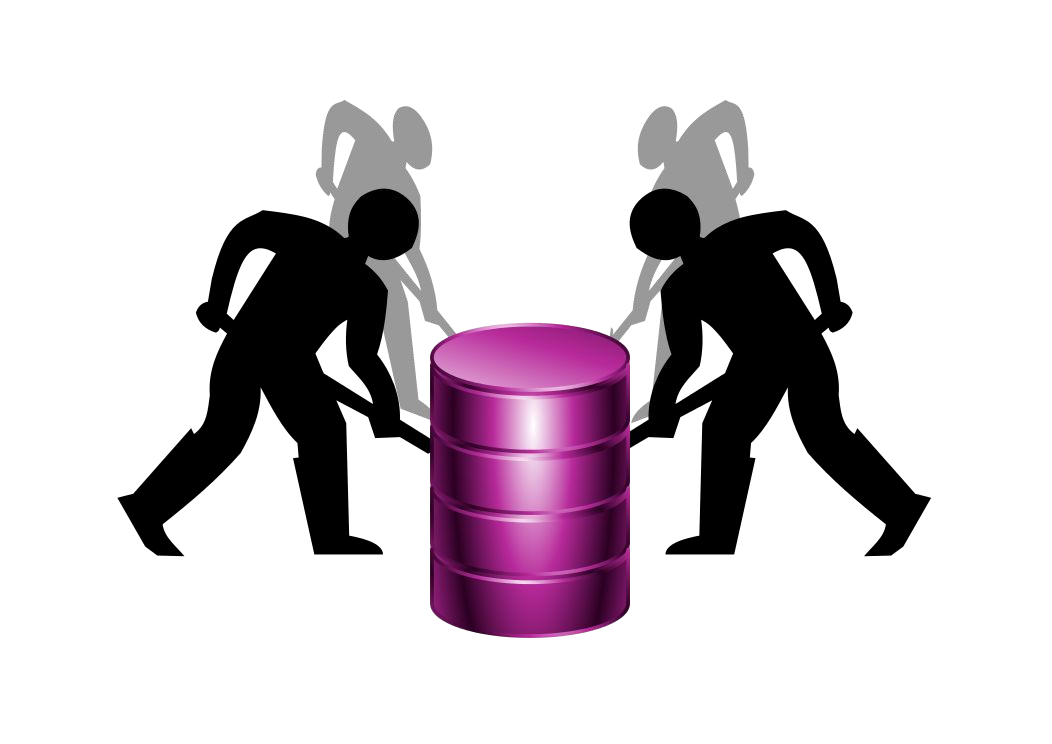
Human activity recognition using

accelerometer data from wearable sensors

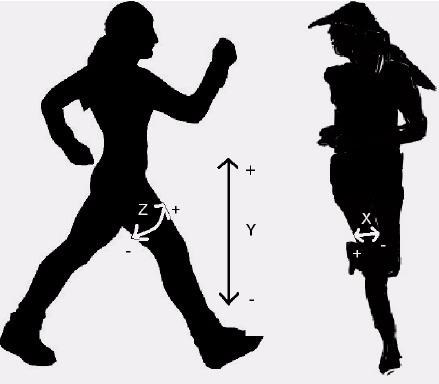
CS 405 Project Presentation By:

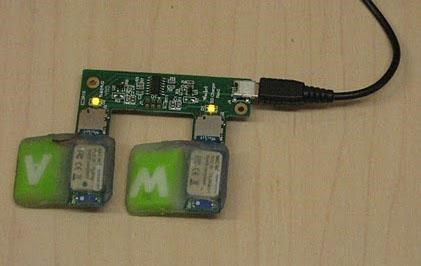


Beknazar Bazarbek uulu COM17

Outline

* Introduction
* Motivation & Goals ● Problem Definition
* Data Mining Steps
* Results and Evaluation
* Conclusion
* What is human activity recognition?
  1. Recognizing multiple sets of daily human activities under real-world condition.
* What devices are being used to collect data for human activities recognition?
  1. Smartphones ○ Wearable devices
* Each of these devices have built in accelerometer

(biaxial/triaxial) that keeps track of human body movement in x,y,z axes.

* Device we are using?
  1. *Wocket accelerometer (+- 4g, sampling rate=90Hz)*
* Wocket accelerometer contains a triaxial accelerometer, a microprocessor, a Bluetooth transmitter and a rechargeable battery.
* These are sufficiently small and can be comfortably worn on all body locations at the same time.
* Raw accelerometer data is acquired and sent using the bluetooth to a smartphone. 

Wocket connected to a charger

Wocket ready to be placed on body

|  |  |  |
| --- | --- | --- |
| ● | ○ | Analyze and understand the process of commercial wearable devices  Commercially available physical activity recognition system like Fitbit, Nike+ FuelBand etc. are widely used but their algorithm has not been validated i.e. it’s still a black box system. |
|  | ○ | In this project, we report our efforts to recognize human activities by working on similar raw accelerometer data. |
|  | ○ | By doing so, we gain in-depth understanding of the activity classification system, and provide recommendation based on our findings. |

●

Applications of these devices in industries such as:

[

Lockhart et al. 2012]

○

Health: Fitness Tracking, Health monitoring, fall detection

○

Social: Share your fitness activities on social networking

sites like Facebook etc.

○

Lifestyle : Context-aware behaviours

○

Targeted Advertising : Advertisement based on user activities.

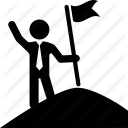
○

Corporate Management and Accounting.



Goals:

* Classifying user’s daily activities by analyzing and processing raw data from wocket accelerometer.
* Suggest best possible position for sensor placement based on the accuracy.
* Suggest best combination of sensor placement sites to classify activities.



* We are working on this project to recognize few of the most important everyday human activities:



* 1. Walking

○ Cycling

○ Lying on back ○ Sitting

* Accelerometer are placed at five body locations at the same time.
  1. Dominant Upper-Arm

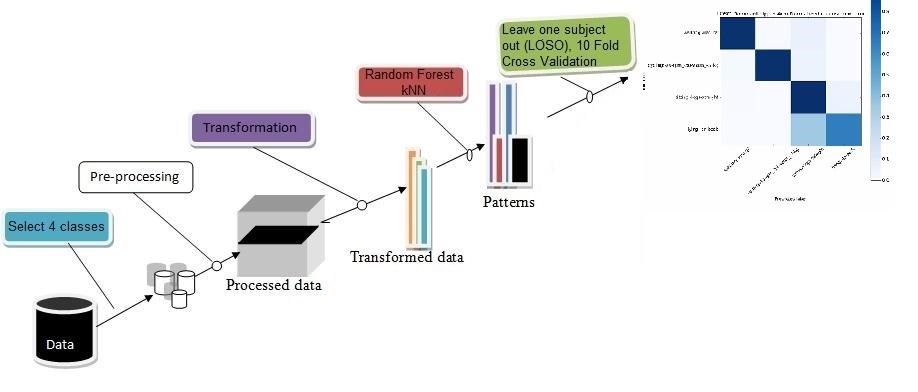
○ Dominant Wrist

○ Dominant Hip

○ Dominant Thigh ○ Dominant Ankle

* These placement sites were selected because of their relevance in

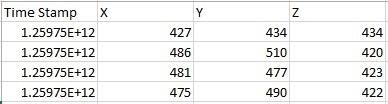
exercise monitoring research.[Mannini et al. 2013]



* Dataset : We have raw data set from 33 participants. For each participant we have following files :
  1. Annotations.csv

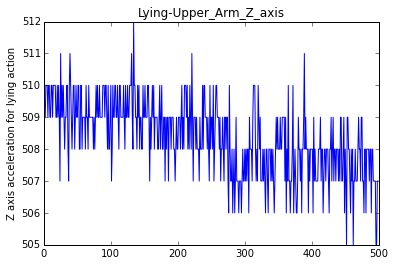
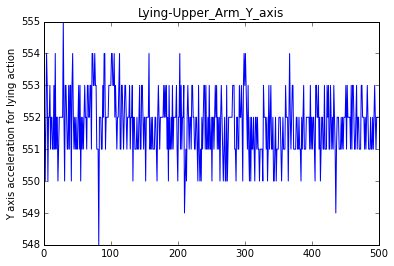
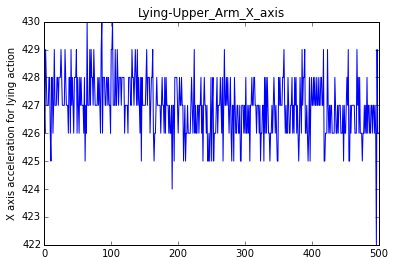
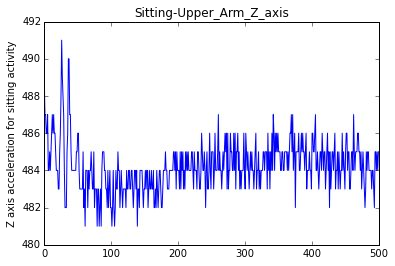
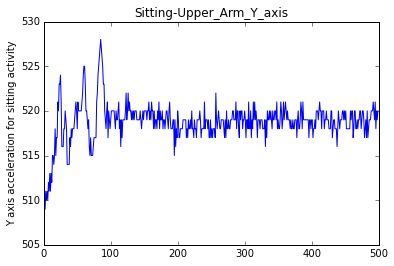
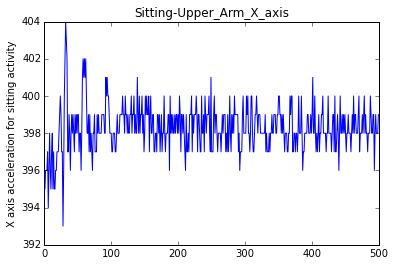
○ Wocket.csv ( A total of 5 files for each sensor location)

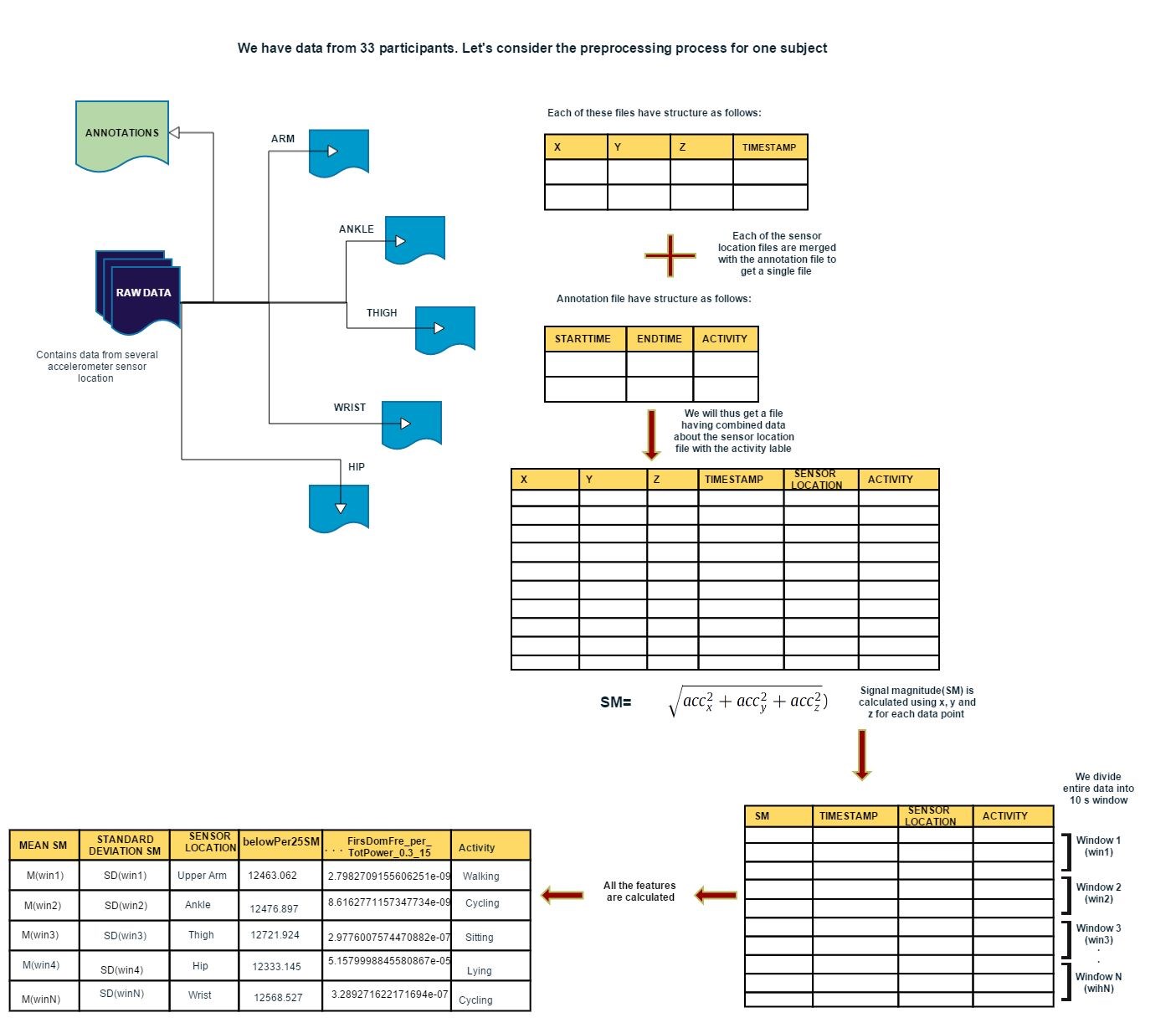
* Demographics: 33 participants
  1. 11 Male, 22 Female, age :18-75, height: 168.5 +/- 9.3cm, weight: 70.0 +/- 15.6 kg

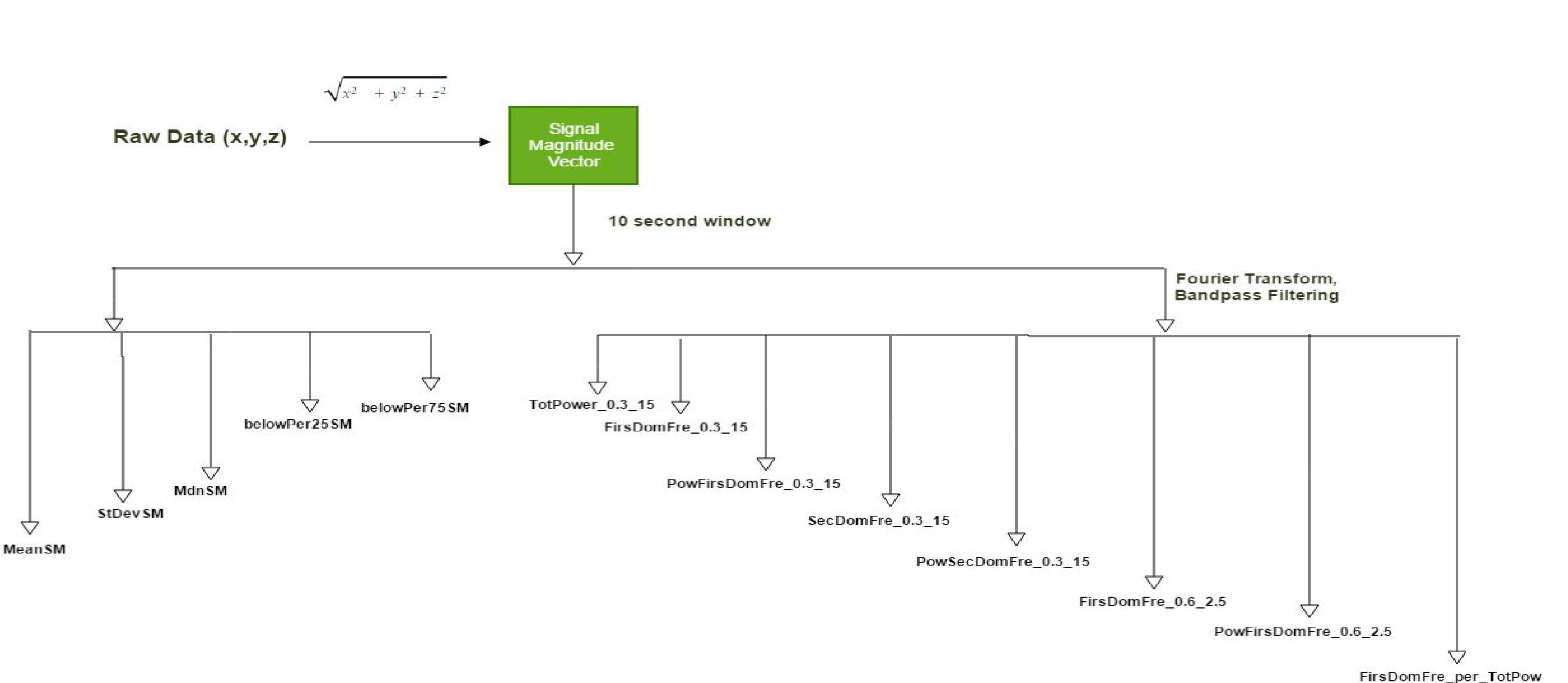


Sample Annotations.csv Sample Wocket.csv

Plotting x,y,z acceleration values for sitting and lying activities for sensor position at upper-arm







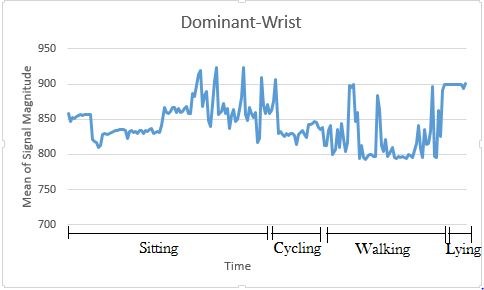
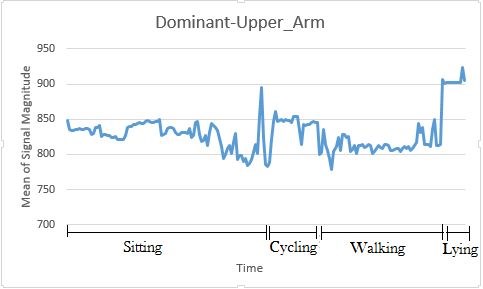
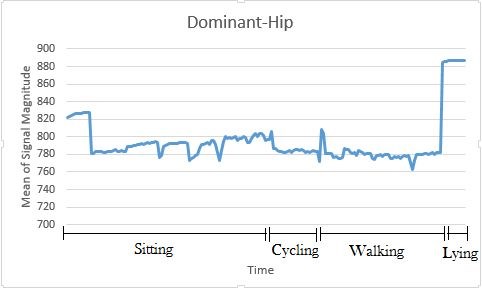
[

Mannini et al.

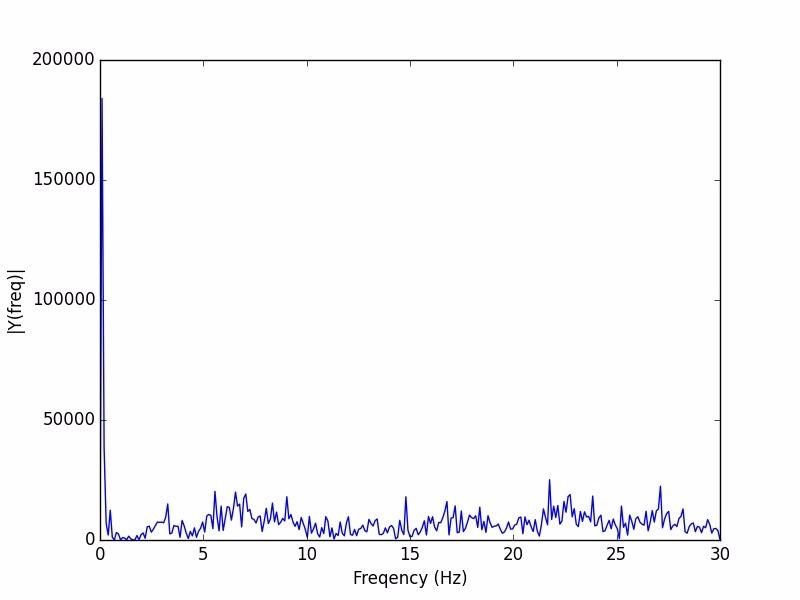
2013]

Features Extraction Methodology

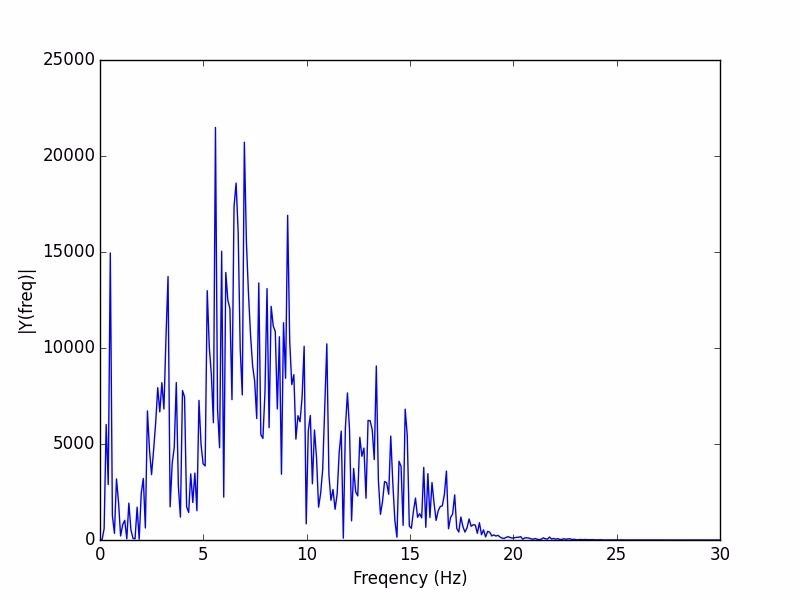
Mean Signal Magnitude for each sensor site



[Mannini et al. 2013]



Frequency before filtering

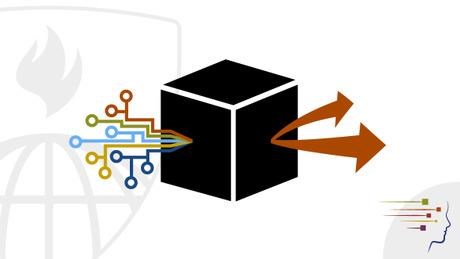


Frequency after filtering

* As the last step of processing the data, we formed following datasets to meet our goal:
  1. Grouped the entire dataset of 33 participants as per the sensor positions.

○ A separate dataset having 33 files, where each file corresponds to a participant.

* Univariate feature selection based on ANOVA : removed features related to SMV percentile



Input Dataset

Result

* Algorithms used:
  1. Random Forest [Ho. et al. 1995]

■ Random Forests are ensemble learning algorithms for classification that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes.

○ k-Nearest Neighbors (k-NN) [ Keller,J.M et al. 1985]

■ The k-NN algorithm is among the simplest of all machine learning algorithms. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its *k* nearest neighbors (*k* is a positive integer, typically small).

* We employed “grid-search” to identify the best parameters suited for above mentioned algorithms.
  1. For Random Forest: Number of Trees ranges from numOfTrees(50-200), InfoGain

(entropy), all features used

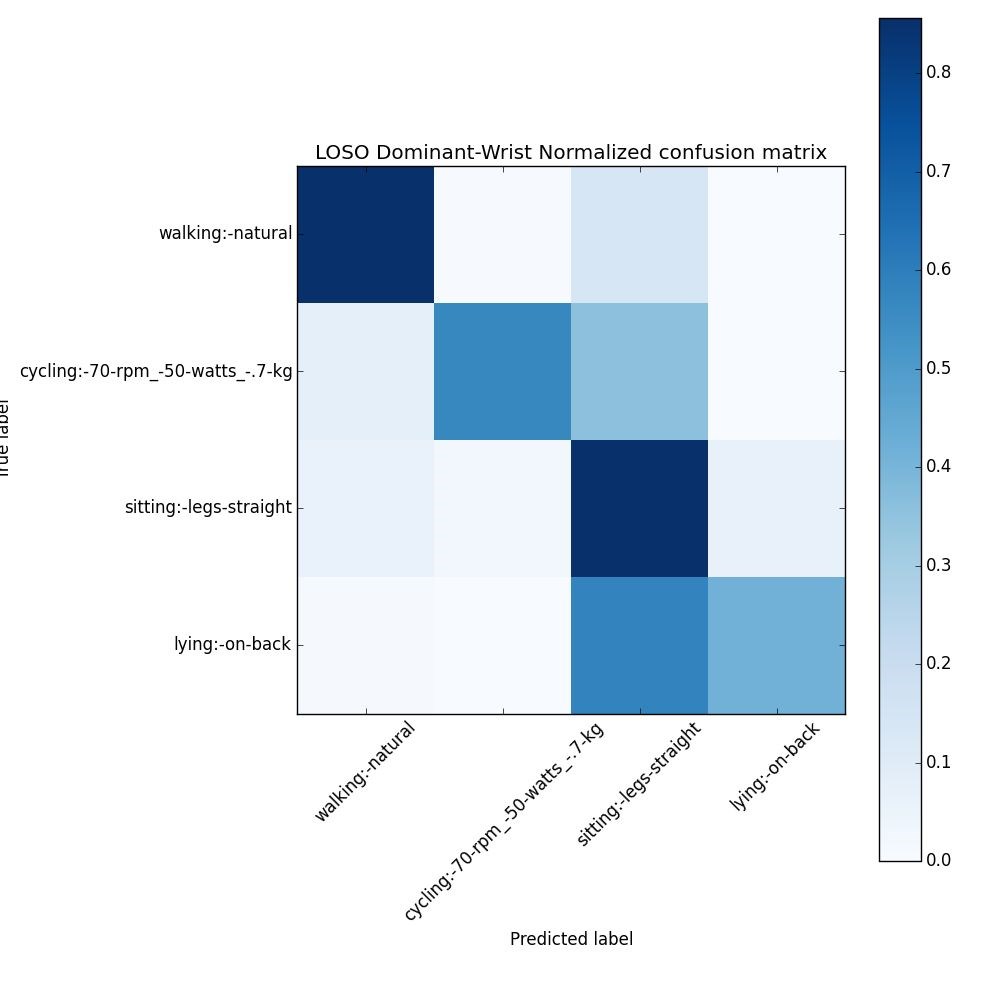
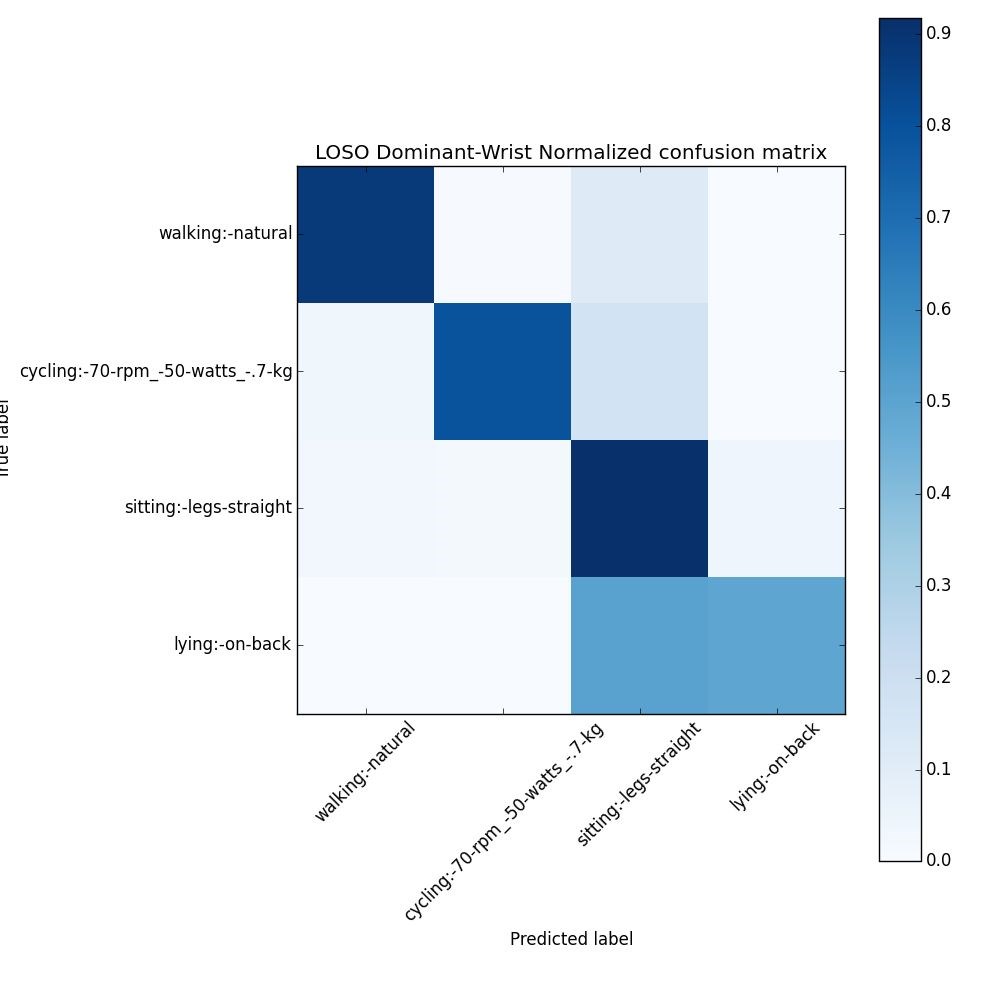
○ For k-NearestNeighbors: Value of k varied from 9 - 11, uniform weight, euclidean distance

Evaluation:

* We used the following methods to evaluate our model
  1. 10-Fold Cross Validation

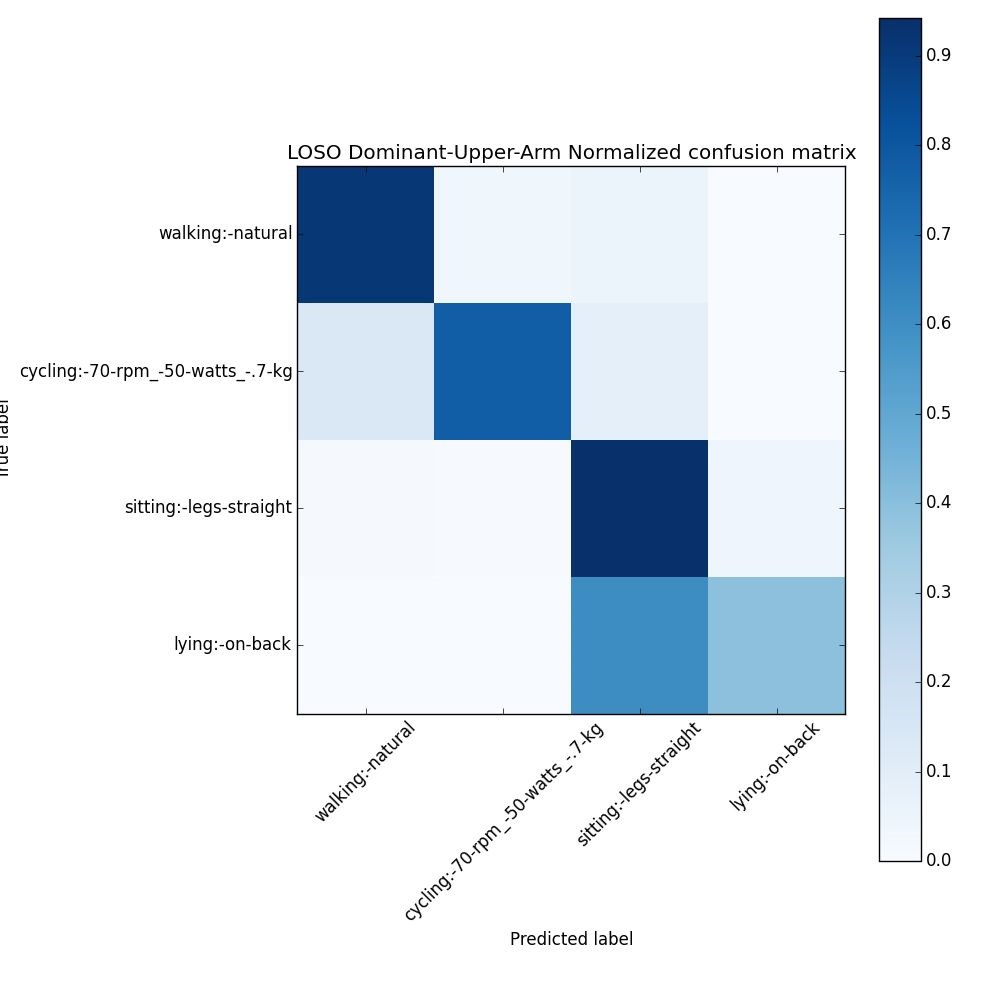
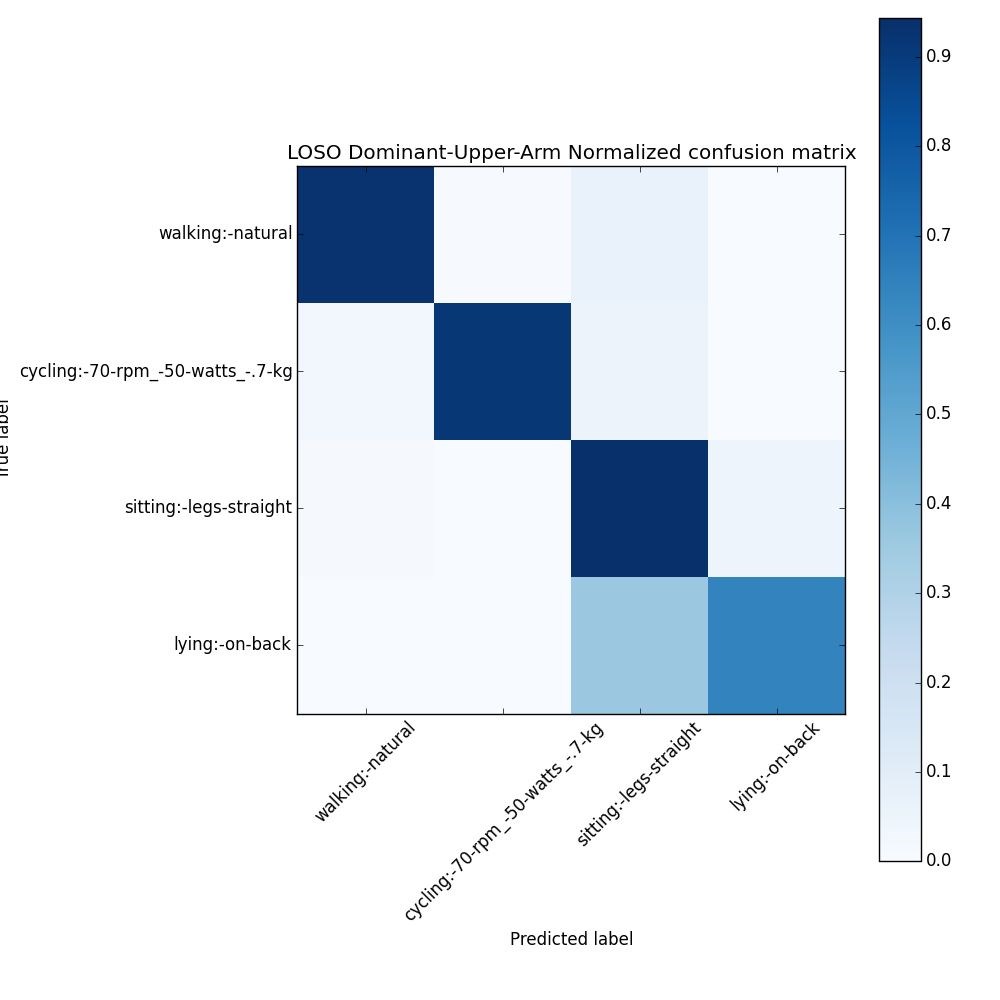
○ Leave-One-Subject-Out (LOSO) : simulating a real-life situation

* Outcome measures included Accuracy, Precision, Recall, and F1-score. ● Confusion matrix after LOSO CV based on Dominant Wrist data:

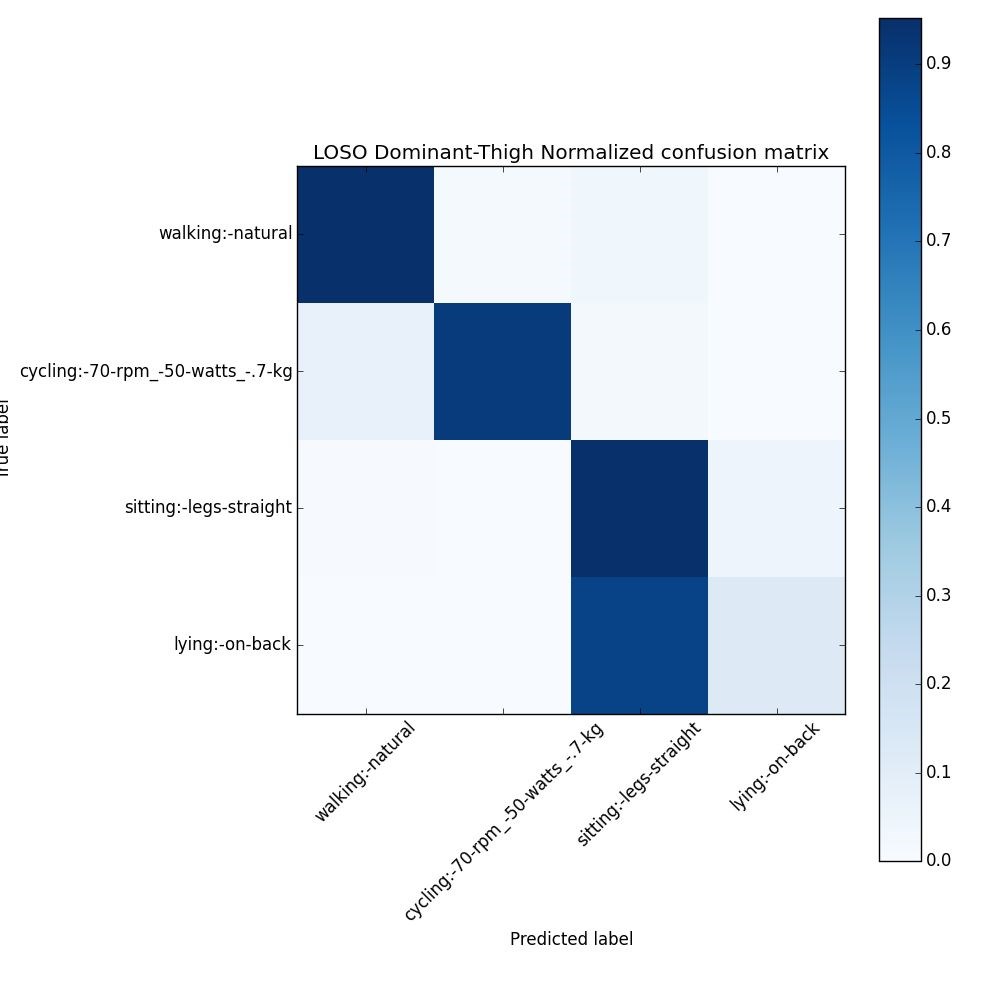
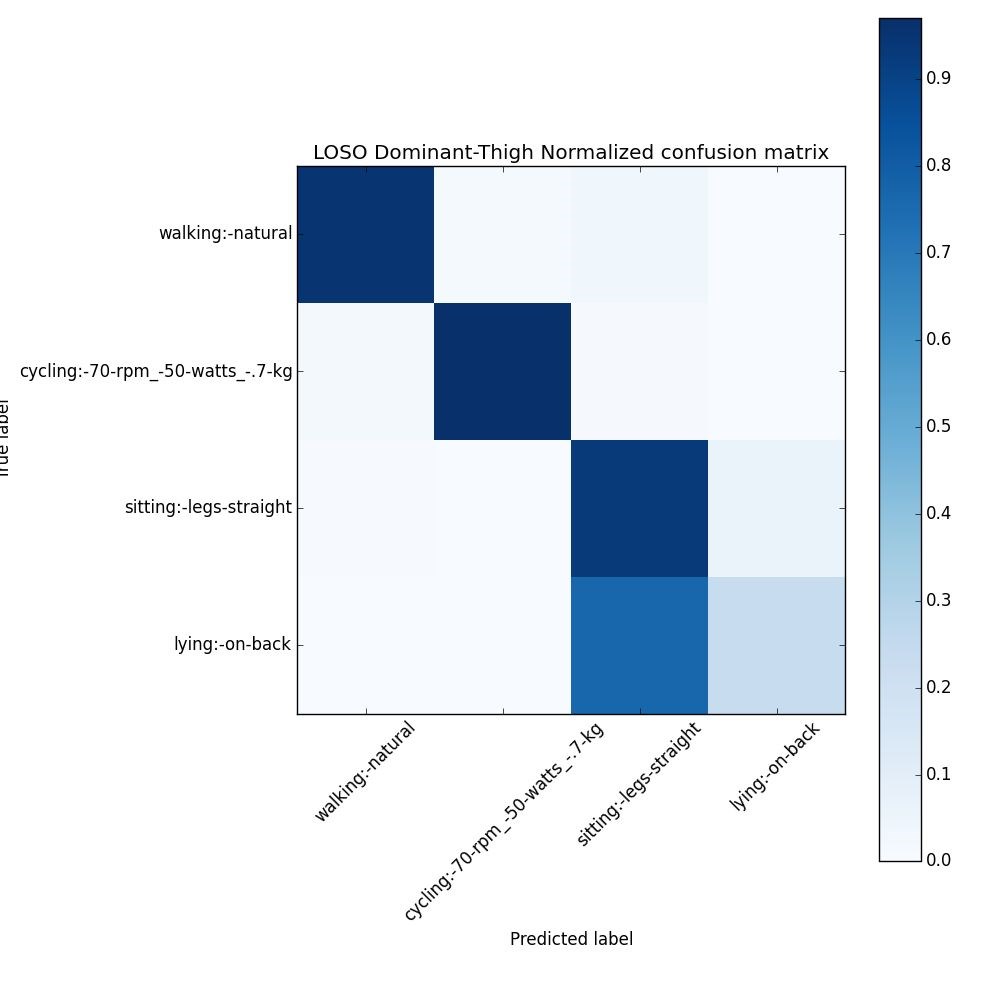


k-NN

* Confusion matrix after LOSO CV based on Dominant Upper arm data:

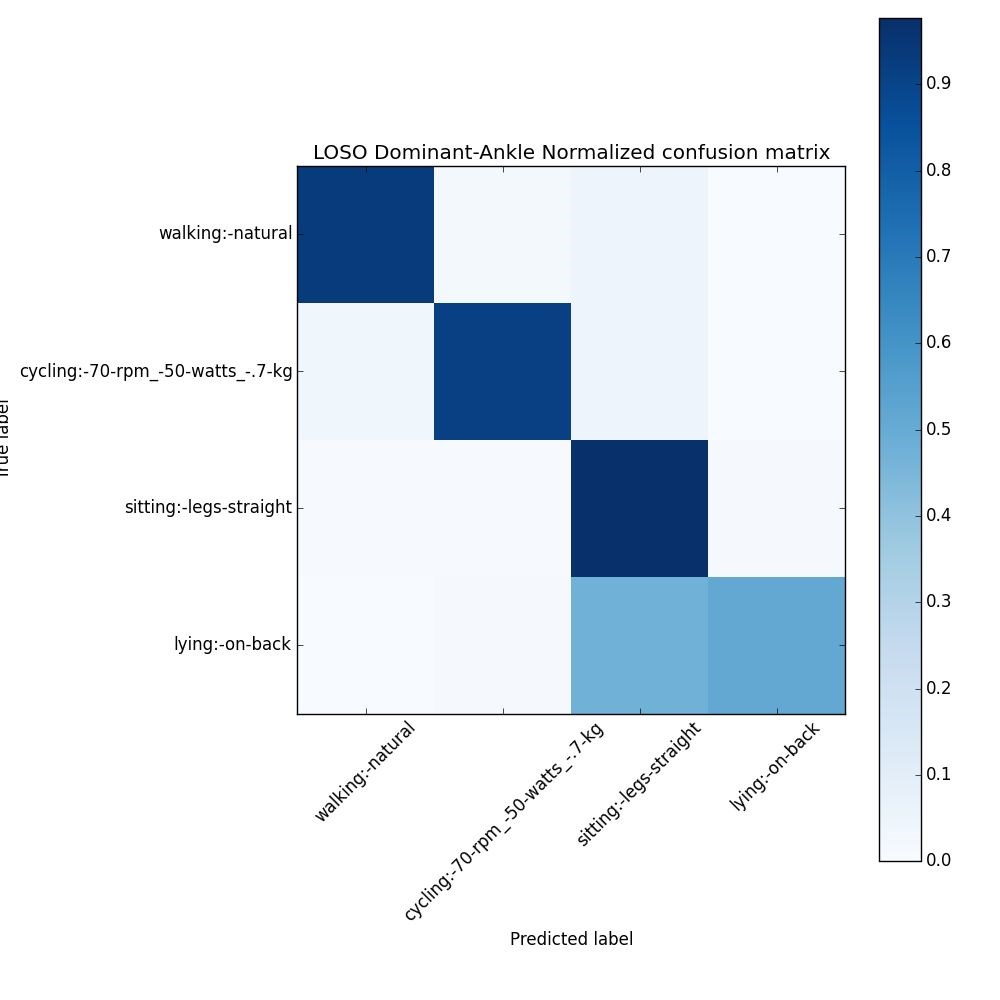
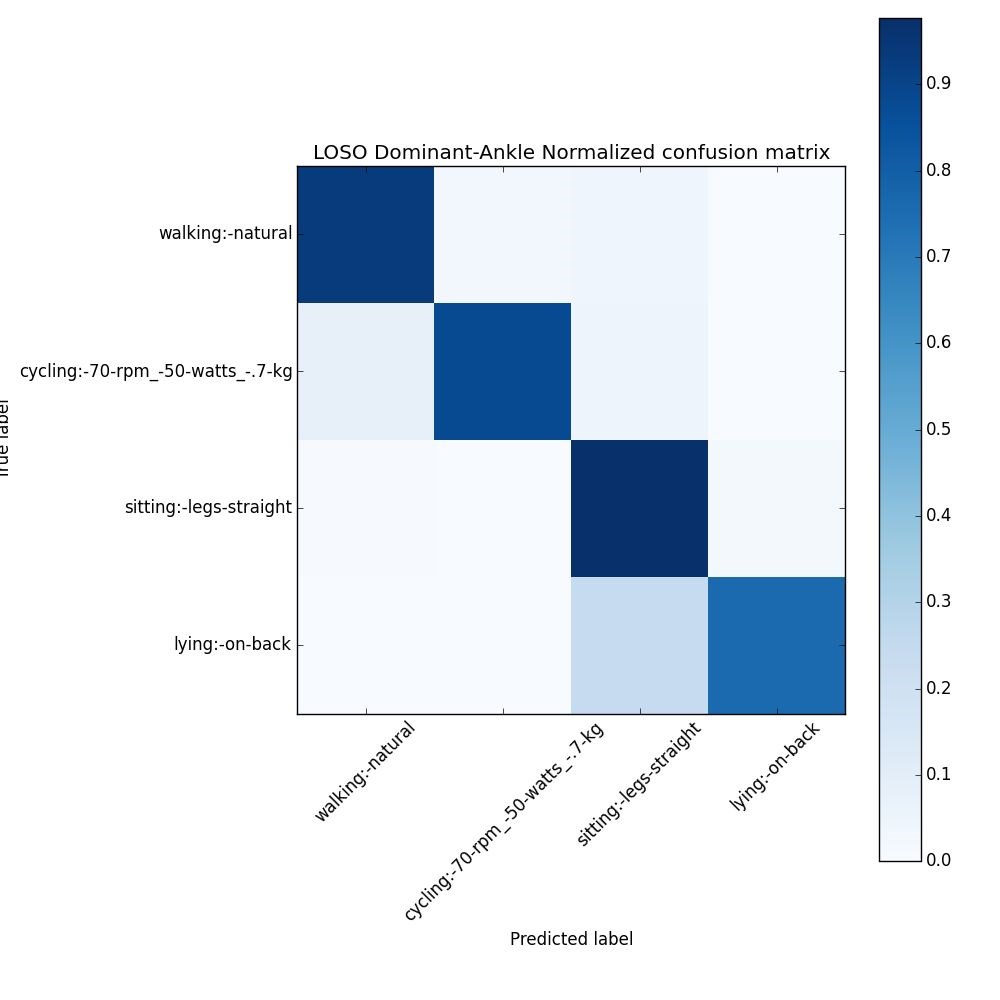
  k-NN

* Confusion matrix after LOSO CV based on Dominant Thigh data:



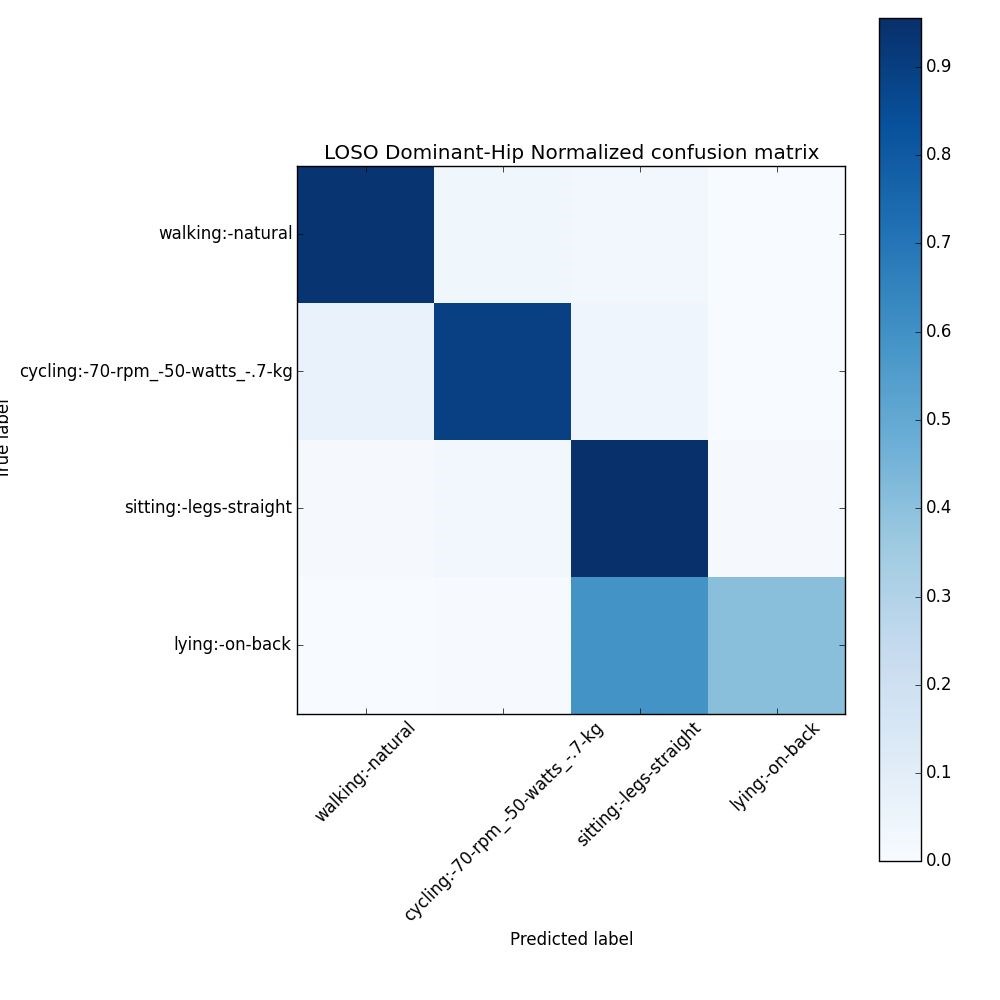
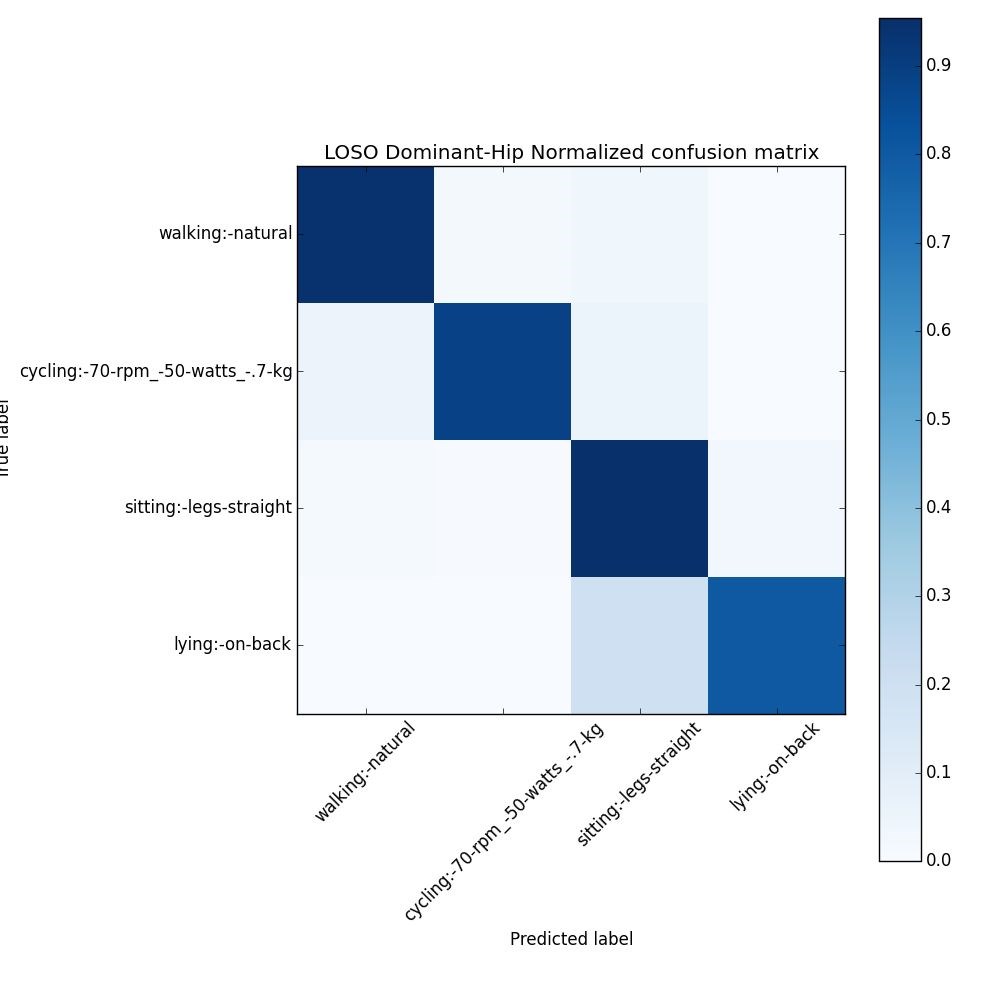
k-NN

* Confusion matrix after LOSO CV based on Dominant Ankle data:



k-NN

* Confusion matrix after LOSO CV based on Dominant Hip data:

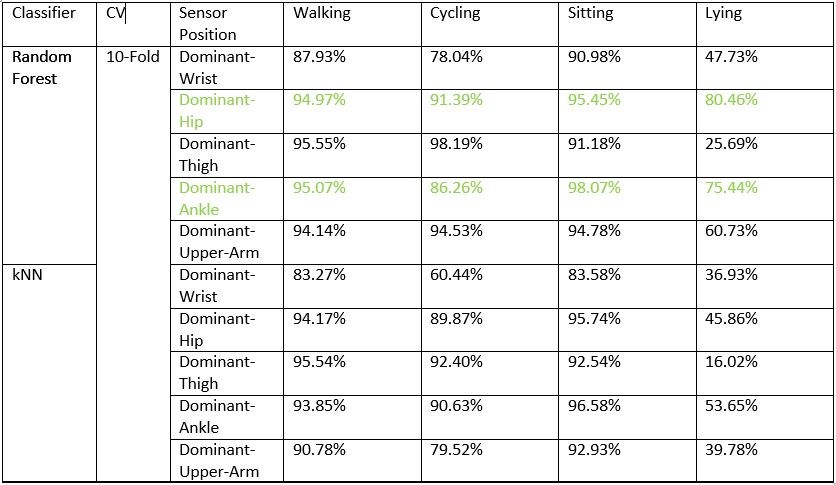


k-NN

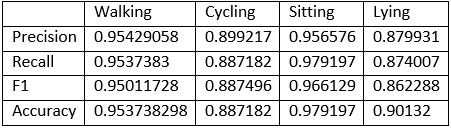
Evaluation Results:



Evaluation Results:



* Combining data from the top two placement sites, Dominant-Hip and Dominant-Ankle



* Thus, we can see in the above table that combining data from DominantHip and Dominant-Ankle improved overall outcome measures.
* Predictive features : Mean, St. deviation, Median, along with Frequency domain features extracted from signal magnitude vector.

* For both LOSO and 10-fold CV, RF performed better than k-NN.
* Data from dominant-hip was the most discriminative in classification, followed by dominant-ankle.
* We recommend combination of both sites for improved classification.

References cited:

* Lockhart, J.W.; Pulickal, T.; Weiss, G.M. Applications of mobile activity recognition. Proceedings of the 14th ACM International Conference on Ubiquitous Computing, Pittsburgh, PA, USA, 5–8 September 2012; pp. 1054–1058
* Mannini A, Intille S S, Rosenberger M, Sabatini A M and Haskell W. Activity recognition using a single accelerometer placed at the wrist or ankle Med. Sci. Sports Exerc. 2013 45 2193–203
* Ho, T.K. Random *Decision Forests*. Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–282.
* Keller,J.M, Gray,M.R, Givens,J.A k-NearestNeighbor. IEEE Transactions on Systems Man and Cybernetics 07/1985; SMC-15(4). DOI: 10.1109/TSMC.1985.6313426

Thank You