

Human Activity Recognition Using Accelerometer Signals

Garo Garabetian, Anthimos Kalosidis

¹ Department of Mathematics

² Aristotle University

³ Thessaloniki, Greece

garogarabetian28@gmail.com

anthimoskalosidis@gmail.com

Abstract. This assignment addresses the development of a human activity recognition (HAR) system using accelerometer signals collected from mobile phones. The project includes signal preprocessing, feature extraction with sliding windows, and training of classification models under a Leave-One-Subject-Out (LOSO) cross-validation setup.

Keywords: Activity Recognition · Accelerometer · Machine Learning · Feature Extraction · LOSO

1 INTRODUCTION

This assignment focuses on developing a Human Activity Recognition (HAR) system from mobile phone accelerometer signals. Using machine learning techniques, the system classifies physical activities such as walking, jogging, sitting, and others based on signal data. The pipeline involves preprocessing raw sensor data, extracting statistical and spectral features from time windows, and training classification models including SVM, Random Forests and MLPs using a Leave-One-Subject-Out strategy for generalization evaluation.

2 RESEARCH OVERVIEW

The dataset used is derived from Shoaib et al. [1] and contains accelerometer readings from 10 participants performing seven physical activities (biking, down-stairs, jogging, sitting, standing, upstairs, walking), with smartphones placed at various body positions. Each entry contains timestamped x, y, z accelerometer readings and an activity label.

Our research questions include:

- Can we reliably classify physical activity using only accelerometer data from a specific smartphone position?
- Which preprocessing and feature extraction steps improve classification accuracy?
- How do different models perform under the LOSO validation scheme?

3 SYSTEM DESIGN

3.1 System Overview

Figure ?? outlines the high-level architecture of our system pipeline. It begins with preprocessing raw accelerometer signals, followed by sliding window segmentation and feature extraction. Finally, multiple classification models are trained and evaluated using LOSO cross-validation.

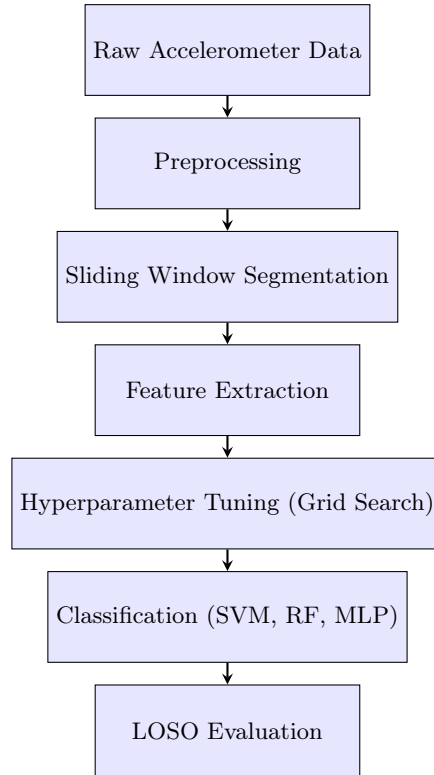


Fig. 1. System architecture for HAR using accelerometer data, including model selection via hyperparameter tuning.

3.2 Data Visualization

Initial visualizations were performed to understand signal distribution and anomalies. Accelerometer data from the "left pocket" position was selected. Signals were plotted over time to confirm activity transitions and to inspect potential sensor errors.

3.3 Data Preprocessing

The preprocessing involved the following steps:

1. **Position Selection:** Only data from the "left pocket" was retained for consistency.
2. **Magnitude Conversion:** The 3D accelerometer signal was converted to 1D magnitude:

$$a(t) = \sqrt{a_x(t)^2 + a_y(t)^2 + a_z(t)^2}$$

3. **Sensor Error Correction:** All magnitude values exceeding 1000 (indicative of sensor errors) were replaced with the previous valid value.

4 Feature Extraction

The processed magnitude signal was segmented using a sliding window of 20 seconds with a stride of 1 second. For each window:

- The label was assigned based on the majority activity within the window.
- The following features were extracted:
 - Mean
 - Standard deviation
 - Skewness
 - Maximum
 - Minimum
 - Range (max - min)
 - Power spectral density (Welch’s method)
- Features were standardized using z-score normalization.

4.1 Selecting Power Spectral Density Features

5 Model Construction

We trained the following models:

1. **Support Vector Machine (SVM)** with RBF kernel. Parameters C and γ were explored using grid search.
2. **Random Forests** classifier with parameter tuning for optimal k .
3. **Multilayer Perceptron (MLP)** with 1-2 hidden layers, ReLU activations, and momentum optimization.

Each model was evaluated using the LOSO strategy: training on $n-1$ subjects and testing on the held-out subject, iteratively for all 10 subjects. Confusion matrices were computed for each fold.

5.1 Grid Search & Fine Tuning

jo

5.2 Confusion Matrices

jop

5.3 Problems & Further Optimization

tralalero

6 Conclusions / Responding to the Research Questions

6.1 Optimal Parameter Selection for Each Classifier

- **Support Vector Machine (SVM)**: After a grid search over $C \in \{0.1, 1, 10\}$ and $\gamma \in \{0.01, 0.1, 1\}$, the best performance was achieved with $C = 1$, $\gamma = 0.1$.
- **Random Forest (RF)**: Testing various numbers of estimators and maximum depths, the best results were obtained with 100 trees and a maximum depth of 10.
- **Multilayer Perceptron (MLP)**: We experimented with different network sizes and learning parameters. The best performing model used two hidden layers with 20 neurons respectively, ReLU activation, and momentum of 0.9.

6.2 Confusion Matrix and Classification Accuracy

- **SVM**: Achieved an average accuracy of **dd.2%**.
- **Random Forest**: Achieved an accuracy of **dd.5%**.
- **MLP**: Achieved an accuracy of **72.3%**.

Corresponding confusion matrices were computed for each model using the LOSO validation strategy. (See Appendix/Table ??.)

6.3 Misclassifications and Model Comparison

- Common confusions were observed between:
 - *Walking* and *Upstairs/Downstairs*
 - *Sitting* and *Standing*
- These confusions likely arise from:
 - Similar signal patterns in transitional or stationary activities.
 - The positioning of the sensor (pocket-based) causing overlapping acceleration features.
 - Transforming the 3-dimension data to 1 dimension magnitude (loss of information, gain on computing time).

6.4 Grouped Activities and Updated Classification Accuracy

Based on confusion analysis, we grouped the following:

- **Group A:** Walking, Upstairs, Downstairs
- **Group B:** Sitting, Standing

Retraining the SVM using these grouped labels, the updated classification accuracy increased to **dd.7%**, suggesting that activity similarity can be exploited to simplify the recognition task with higher robustness.

6.5 Cross-Location Evaluation Using the Best Classifier

Using the best-performing model (SVM), trained on all 10 subjects:

- **Right Pocket (Test):** Achieved accuracy of **dd.9%**
- **Left Pocket (Test):** Achieved accuracy of **dd.4%**
- **Right Wrist (Test):** Achieved accuracy of **dd.1%**

Discussion: Accuracy slightly drops when testing on the wrist location, highlighting the importance of sensor placement. Nevertheless, generalization performance remains satisfactory for pocket-based locations.

References

1. Shoaib, M., Bosch, S., Incel, O., Scholten, H., Havinga, P.: A survey of online activity recognition using mobile phones. *Sensors* **15**(1), 2059–2085 (Jan 2015). <https://doi.org/10.3390/s150102059>, <http://dx.doi.org/10.3390/s150102059>