NOVA SCHOOL OF SCIENCE & TECHNOLOGY

DEPARTMENT OF PHYSICS

SARA LEONOR MARTINS DOS SANTOS

BSc in Biomedical Engineering

UNSUPERVISED HUMAN ACTIVITY RECOGNITION MODELS FOR CHARACTERIZING OFFICE WORK TASKS

MASTER IN BIOMEDICAL ENGINEERING

NOVA University Lisbon
September, 2024

NOVA SCHOOL OF
SCIENCE & TECHNOLOGY

DEPARTMENT OF
PHYSICS

UNSUPERVISED HUMAN ACTIVITY RECOGNITION MODELS FOR CHARACTERIZING OFFICE WORK TASKS

SARA LEONOR MARTINS DOS SANTOS

BSc in Biomedical Engineering

Adviser: Hugo Filipe Silveira Gamboa
Full Professor, NOVA University Lisbon

Examination Committee

Chair: Ricardo Nuno Pereira Verga e Afonso Vigário
Associate Professor, NOVA University Lisbon

Rapporteur: Marília da Silveira Gouveia Barandas
Senior Researcher, Fraunhofer AICOS Portugal

Adviser: Hugo Filipe Silveira Gamboa
Full Professor, NOVA University Lisbon

MASTER IN BIOMEDICAL ENGINEERING

NOVA University Lisbon
September, 2024

Unsupervised Human Activity Recognition Models for Characterizing Office Work Tasks

This document was created with the (pdf/Xe/Lua)LaTeX processor and the NOVAthesis template (v7.1.18) [1].

Acknowledgements

I would like to start by thanking Professor Hugo Gamboa for giving the chance to work on this project and for believing in my capabilities throughout this process. I also want to express my utmost gratitude to Phillip Probst, for helping me climb this tough and scary mountain all the way from base camp to the top. Thank you Phillip, for the guidance, motivation, and patience you showed me all of these months, and for teaching me not only what science is and how it is done, but also for all the life advice and shared laughs. I could not have chosen a better environment than LibPhys, where I always felt welcomed from day one and never hesitated to ask for help. For that, thank you Dr. Luís, Inês, Dânia, Mariana, and Leonor.

To my family who supported me throughout these five years that are now ending. Mom and Dad, I am forever grateful to both of you for always making sure I had everything I needed. Thank you, Mom, for being my biggest supporter, even during my lowest moments, and for always giving me the strength to push forward. Thank you, Dad, for giving me rides home late at night when I was tired after school, even when you were exhausted too. To my bro, thanks for keeping me in check and for always reminding me that life is much more than just studying. Aos meus avós que me viram crescer aos fins-de-semana e a quem sempre me custou dizer adeus. Obrigada Diana, por me aturares nos meus dias piores, ao ouvires os meus desabafos e ao cozinhares para mim. Aos meus gatinhos, as minhas fontes de alegria diárias, que fazem tudo muito mais leve. Ao meu trio Bigodinhos, Tita e Nini, ao incompreendido Sid e à minha primogénita e companheira Nïn.

Finally, to my friends, without whom none of this would have been possible. You were the best thing this University gave me. Márcia, I always miss the words to describe our friendship. I can only thank you for being the best friend I could have ever asked for. I was never alone these five years, if we struggled or laughed or cried, it was always together. Thank you. Jéssica my ratty, thank you for being my partner in arms, not only during these lengthy thesis months, but since the first day I stepped into this place. My memer Alexandra, thank you for making me laugh and for all the support. You can count on me to hold you up!

Abstract

Office workers usually spend most of their time sitting, often with rigid postures, for prolonged periods of time. This has been recognized by the European Union as a risk factor for work-related musculoskeletal disorders (WRMDs). Promoting healthier work environments by encouraging more active work practices is essential for workers occupational health. To study work activities and their distribution over time, Human Activity Recognition (HAR) techniques need to be implemented. These techniques are typically used in broader contexts, relying on supervised methods that require costly, time-consuming label annotation.

This work presents a new dataset comprised of nine office activities from 19 subjects, collected using a smartphone and a smartwatch, to train unsupervised learning models for HAR in office environments. Three experiments were conducted to evaluate the performance of traditional clustering models: feature selection, cluster stability over time, and impact of data imbalance. The results show that, considering only basic activities, such as sitting, standing, and walking, the models accurately cluster each subject using both subject-specific and subject-independent feature sets. However, subjects' data is not similar enough to develop a subject-independent model. With all nine activities, the feature sets lose effectiveness. Smartwatch sensors produce worse results than smartphone sensors or a combination of these two. Clustering performance is impacted by subjects displaying different behaviours over time and by class imbalance. When testing on data from two subjects performing their actual work tasks for a whole workday, the models are able to separate active and non-active activities into two distinct clusters, but struggle to identify the exact activity. This information can be useful for identifying risk factors for WRMDs.

Keywords: Unsupervised learning, Human Activity Recognition, Smart devices, Work-related musculoskeletal disorders, Occupational health

Resumo

Trabalhadores de escritório passam grande parte do seu tempo sentados e com posturas rígidas por longos períodos de tempo. Isto é considerado pela União Europeia como um fator de risco para lesões musculoesqueléticas relacionadas com o trabalho (LMERT). Promover ambientes de trabalho saudáveis, incentivando práticas mais ativas, é essencial para a saúde dos trabalhadores. Para estudar atividades e correspondente duração recorre-se a técnicas de reconhecimento de atividades humanas (RAH). Estas técnicas, geralmente, são aplicadas em contextos mais gerais e recorrendo a métodos de aprendizagem supervisionada, que requerem a tarefa dispendiosa e demorada de anotação de classes.

Este estudo apresenta um novo conjunto de dados, que contém nove atividades de escritório realizadas por 19 sujeitos, adquirido através de um smartphone e um smartwatch, para treino de modelos de aprendizagem não supervisionada para RAH em escritórios. Foram realizadas três experiências para avaliar a performance de modelos tradicionais de *clustering*: seleção de características, estabilidade dos *clusters* ao longo do tempo e impacto de dados desbalanceados. Os resultados mostraram que, considerando apenas atividades básicas como sentar, estar de pé e caminhar, os modelos conseguem agrupar corretamente cada sujeito, tanto com um conjunto de características otimizado como um comum. No entanto, não é possível agrupar os dados dos sujeitos em conjunto. Considerando as nove atividades, os conjuntos de características perdem eficácia. Os sensors do *smartwatch* apresentam piores resultados do que os do *smartphone*. Os modelos de *clustering* são afetados pelos diferentes comportamentos dos sujeitos ao longo do tempo e pelo desbalanceamento de classes. Ao testar os modelos em dados de dois sujeitos a trabalhar normalmente durante um dia, estes conseguem separar atividades ativas e não ativas em dois grupos distintos, mas não qual a exata atividade. Isto pode ser útil para a identificação de fatores de risco para LMERT.

Palavras-chave: Aprendizagem não supervisionada, Reconhecimento de atividades humanas, Dispositivos inteligentes, Lesões musculoesqueléticas relacionadas com o trabalho, Saúde ocupacional

Contents

List of Figures

List of Tables

List of Algorithms

List of Abbreviations

1GM One-stage general model *(pp. 17, 18, 20, 23, 24, 26, 28, 34, 36)*
2GM Two-stage general model *(pp. 18–21, 24, 25, 27, 28, 30, 35–37)*

ACC Accelerometer *(pp. 3, 6–8, 11–16, 33, 38)*
AGG Agglomerative Clustering *(pp. 4, 7, 8, 17, 20, 23–27, 29, 32, 34–37)*
ARI Adjusted Rand Index *(pp. 5, 7, 8, 18, 19, 21–32, 34–38)*
AT Autoridade Tributária e Aduaneira *(pp. 1, 9)*

EMG Electromygraphy *(pp. 1, 11, 13)*
EU European Union *(p. 1)*

GMM Gaussian Mixture Model *(pp. 4, 5, 7, 17, 20, 23–27, 29, 34–38)*
GYR Gyroscope *(pp. 3, 6–8, 11, 13, 14, 16, 31, 33, 38)*

HAR Human Activity Recognition *(pp. 2, 3, 6, 8–11, 13, 17, 18, 21, 22, 32, 35, 37)*
HTA Hierarchical Task Analysis *(pp. 9, 11, 17)*

IMU Inertial Measurement Unit *(pp. 7, 35)*

KDE Kernel density estimation *(pp. 21, 31, 38)*

M_{P+W} Modality smartphone and smartwatch *(pp. 18, 22–27, 32, 34)*
M_P Modality smartphone *(pp. 18, 21–28, 30, 32, 33, 35, 37)*
M_W Modality smartwatch *(pp. 18, 22–27, 32–34, 36)*
MAG Magnetometer *(pp. 3, 7, 11, 13, 14, 16, 22, 31–35, 38)*
ML Machine Learning *(pp. 2, 4, 6, 17)*
MVC Maximum Voluntary Contraction *(pp. 11, 13)*

RF Random Forest *(pp. 8, 20, 28, 36)*

RI Rand Index *(pp. 5, 8)*

SD Standard deviation *(pp. 23, 25, 26, 29, 31, 34–36)*
SMMs Subject specific models *(pp. 17–20, 22–28, 32, 34–36)*

WRMDs Work-related musculoskeletal disorders *(pp. 1, 2, 38)*

1 Introduction

1.1 Context and motivation

Since the early 2000s, the number of people employed in computer-based office work has been steadily increasing across the European Union (EU). Specifically, from 2000 to 2015, the percentage of workers who spent at least a quarter of their workday doing computer work increased from 47 % to 58 % [2]. Office work predominantly consists of low variance activities such as sitting for long periods of time, often with rigid postures [3–5]. This has been linked to musculoskeletal pain, particularly in the lower back, neck, shoulders, and knees [6]. Additionally, office workers are often confronted with high job demands, while being limited by low job resources [7]. A combination of these factors contribute to the development of Work-related musculoskeletal disorders (WRMDs), stress, depression, and anxiety related problems, which are a significant health concern for 7.4 % of European workers [4]. WRMDs are associated with loss of productivity and increased absenteeism, resulting in medical burden and increased economic costs for organizations [2, 8]. This problem has been recognized by the EU, which is actively funding initiatives under the 2021-2024 Horizon Europe program, particularly within Cluster 1, to promote healthier living and working environments [9]. Implementing more active work practices that reduce sitting time and encourage more walking and standing, is crucial for occupational health and has shown positive health outcomes [6, 10].

To address some of the above-mentioned issues, the PrevOccupAI Project (Prevention of Occupational Disorders in Public Administration based on Artificial Intelligence) was carried out with the objective of evaluating occupational risk factors for WRMDs in office workers [11]. Biosignals were acquired with the purpose of studying workers' postures during their workday. Data was collected from 40 workers across four Autoridade Tributária e Aduaneira (AT) offices during five consecutive days. Acquisitions were scheduled so that a smartphone acquired data during the whole workday, and a smartwatch and two Electromygraphy (EMG) sensors acquired four times a day for 20 minutes. As subjects were not observed during work, the resulting dataset is unlabeled.

While PrevOccupAI focused more on postural information, the dataset offers the

opportunity to also study work activities and how they are distributed over time. To achieve this, Human Activity Recognition (HAR) techniques need to be implemented. HAR is an active research area, driven by advancements in sensor technology, particularly in smart devices. These devices, which are widely used by the population, have built-in sensors, eliminating the need for additional hardware [12]. In HAR, Machine Learning (ML) models are used to infer activities from sensor data. The most common approach is supervised learning, specifically classification methods, which yield the most accurate outcomes [13]. However, supervised techniques require annotating labels, which is costly and time-consuming [13], and can raise privacy concerns if done using cameras [14]. Furthermore, HAR is often done in broader contexts with the purpose of distinguishing simple activities [12, 13]. Therefore, to do HAR in a scenario similar to PrevOccupAI, datasets comprised specifically of office activities are needed, and unsupervised learning schemes should be implemented in order to handle unlabeled data. Leveraging a dataset of office activities to train unsupervised learning models that can effectively distinguish between activities, could be a powerful tool for identifying risk factors for WRMDs. Providing companies with reports on employees' activities and their corresponding durations could help identify risk factors, leading to future recommendations to improve workers' occupational health.

1.2 Objectives

In line with the context and motivation provided in the previous section, and as an extension of the PrevOccupAI project, this thesis has the following objectives:

- **(O1):** Use smart devices, in particular smartphones and smartwatches, to collect a diverse dataset that includes various office work activities for developing unsupervised HAR systems.
- **(O2):** Employ traditional clustering models for HAR on the collected dataset.
- **(O3):** Find consistent feature sets that can be used for HAR in the context of office work.
- **(O4):** Study if there is enough similarity within the data to develop a subject-independent clustering model.
- **(O5):** Evaluate the stability of clusters over time, as well as the impact of imbalanced data on the clusters.
- **(O6):** Test the proposed models on data from two subjects carrying out their actual work tasks for an entire workday.

The research work described in this dissertation was carried out in accordance with the norms established in the ethics code of Universidade Nova de Lisboa. The work described and the material presented in this dissertation, with the exceptions clearly indicated, constitute original work carried out by the author.

2

Theoretical Concepts

2.1 Human Activity Recognition

Human Activity Recognition (HAR) is the process of automatically identifying human activities using data from various sources, primarily inertial sensors, by means of artificial intelligence techniques [15]. The sensors can be embedded in electronic devices, such as smartphones and smartwatches, and body-worn [15, 16]. HAR has diverse applications, including health monitoring, surveillance, sports analysis, and more [17].

2.2 Inertial Sensors

Inertial sensors are sensors that measure linear acceleration, angular velocity, and magnetic field, typically across three orthogonal planes [18]. These sensors include the Accelerometer (ACC), Gyroscope (GYR), and Magnetometer (MAG) [16]. Advances in micro-electromechanical systems have significantly reduced the size and cost of these devices, making them suitable for integration into smart devices [19]. These smart devices, such as smartphones and smartwatches, include analog-to-digital converters within their processors [20], eliminating the need for manual conversion.

The ACC measures the rate of change of velocity (acceleration) of an object in meters per second squared (m/s^2) [21]. Specifically, it measures the force exerted on a mass inside the device [18]. The GYR measures angular velocity, which corresponds to the rate of change in the rotational position of a body around an axis, typically measured in radians per second (rad/s) [22, 23]. The MAG measures the strength and direction of magnetic fields near an object, usually reported in microteslas (μT) [24]. The MAG is commonly used to detect the orientation and direction of an object by sensing the Earth's magnetic field [25].

2.3 Unsupervised Learning

Unsupervised learning is a type of Machine Learning (ML) in which an algorithm identifies underlying patterns or groupings within a dataset without prior knowledge of groups or classes [26, 27]. One of the most common unsupervised learning models that allows for the detection of these patterns is clustering. A cluster is a group of data points with similar characteristics, and different clusters exhibit distinct features [28]. This thesis employs the following clustering models: KMeans, Agglomerative Clustering (AGG), and Gaussian Mixture Model (GMM).

2.3.1 KMeans

KMeans is one of the most widely used clustering algorithms and operates as a partitioning method [29]. This method divides n unlabeled data points into k clusters, where k is specified by the user [29, 30]. In KMeans, each cluster is represented by a cluster center, which corresponds to the mean position of all data points within that cluster [28, 29]. The KMeans algorithm begins by randomly assigning initial cluster centers. It then proceeds in two steps: first, data points are assigned to the nearest cluster center based on a distance metric, typically the euclidean distance. Next, the cluster centers are updated by computing the mean of the data points within each cluster, shifting the centers to the new average position [28]. With KMeans, points can be assigned to preformed clusters by computing the mean to the closest cluster center [31].

2.3.2 Agglomerative Clustering

Hierarchical clustering is a cluster analysis technique that organizes clusters into a hierarchy, which can be either agglomerative or divisive. The AGG algorithm starts by treating each data point as its own cluster [29] and with each iteration, pairs of close clusters are merged [30]. AGG can be implemented in various ways, depending on the criteria used for merging clusters, known as the inter-cluster dissimilarity measure or linkage. Some of the most common linkage methods include single, complete, average, and Ward linkage [32].

Single linkage merges pairs of clusters that have the minimum distance between them, while complete linkage merges pairs with the smallest maximum distance. For average linkage, the distance between two clusters is defined as the average of all distances between the data points in the two clusters. During each iteration, the pairs of clusters with the smallest average distance are merged. These three linkage methods can be computed using different distance metrics, whereas Ward linkage only allows for euclidean distance. Ward linkage minimizes the increase in total within-cluster variance when merging two clusters by calculating the total sum of squared deviations from the mean within each cluster and then merging the pair that results in the smallest increase in this value [32]. This thesis used the Ward linkage method. In AGG, the user can define the number of

clusters to be identified. If this value is unknown, a distance threshold must be set, beyond which clusters will not be merged [28].

2.3.3 Gaussian Mixture Model

Model-based clustering algorithms assume that each cluster follows a specific probabilistic model [29]. In the case of GMM, the algorithm assumes that all data points are generated by a mixture of a finite number of Gaussian distributions, each with its own mean and covariance matrix [33]. The number of Gaussian distributions corresponds to the number of clusters the user expects to find. This model iteratively refines the cluster assignments and the parameters of the Gaussian distributions using the expectation-maximization algorithm. The implementation details of GMM are thoroughly described in [33]. Similar to KMeans, GMM can also assign new data points to the Gaussian distribution that best represents them.

2.3.4 Clustering Metrics

To evaluate clustering performance, external or internal metrics can be employed [34]. External metrics are used when ground truth labels are available. One such metric is the Rand Index (RI), which is analogous to accuracy in supervised settings. The RI evaluates the similarity between two clusterings, the predicted clusters obtained from clustering algorithms, and the true clusters [34, 35]. It measures how consistently pairs of data points are either grouped together or separated in both clusterings. However, the RI does not account for the likelihood of random correct cluster assignments. Therefore, the Adjusted Rand Index (ARI) is often preferred, as it adjusts for this chance. From [36] and [37], the RI and ARI are defined as follows:

$$\text{RI} = \frac{\text{number of agreeing pairs}}{\text{number of pairs}} \quad (2.1) \qquad \text{ARI} = \frac{\text{RI - Expected RI}}{\text{Max RI - Expected RI}} \quad (2.2)$$

The ARI returns values close to 0 for random labelings and values close to 1 when the predicted and true clusters are identical. This metric can also yield negative values for particularly poor clusterings.

When labels are not available, internal validation metrics such as the silhouette coefficient can be used. This metric combines both cohesion, which measures how closely related the points within the same cluster are, and separation, which measures how distinct different clusters are from each other [34]. According to [38], the silhouette coefficient is calculated as follows:

$$\text{Silhouette Coefficient} = \frac{(b - a)}{\max(a, b)} \quad (2.3)$$

Where a is the mean intra-cluster distance, and b is the mean distance between a sample and the nearest cluster to which the sample does not belong to. Scores close to 1 indicate well-separated clusters and 0 indicate overlapping clusters.

3

Literature Review

Advances in sensor technology and increased computational power have significantly contributed to the rapid growth of HAR, making this a prominent research area [39]. Numerous review and survey papers have been published over the years [12, 13, 40–42]. These surveys cover key aspects of HAR research, such as the types of sensors used, the recognized activities, and employed ML models and the utilized approaches (i.e., supervised or unsupervised). This section focuses on HAR datasets and the commonly applied unsupervised learning models.

3.1 Activity Recognition Datasets

HAR can be complex, influenced by factors such as the variety and number of activities being studied, the duration of these activities, the variability between subjects, as different individuals can often perform the same activity in different ways, and variability over time within the same subject. These factors can impact HAR frameworks, often limiting their generalizability [43]. In the literature, there is often a distinction between basic and complex activities. Basic activities typically involve low-variability movements that can consist of clear, repetitive patterns, making them easier to distinguish. In contrast, complex activities usually consist of compound actions that may include combinations of basic activities, interactions with objects, or sequences of movements, resulting in much higher variability [12, 13, 17, 41, 44].

The studies by Machado et al. [45] and Kwon et al. [46] are examples of research focused on lab-collected basic activities such as standing, sitting, walking, running, and lying down. In [45], data was collected from eight subjects using a waist-mounted tri-axial ACC, with each activity lasting one minute. In [46], tri-axial ACC and GYR data were collected from a single subject using a smartphone placed in the pants pocket. Regarding publicly available datasets, UCI-HAR [47], HHAR [48], and MHEALTH [49] primarily feature basic activities. The UCI-HAR dataset includes data from 30 subjects performing six activities—standing, sitting, lying down, walking, and going up and down stairs—recorded via a smartphone placed at the waist, capturing ACC and GYR data,

each lasting roughly 30 seconds. The HHAR dataset includes the same activities as UCI-HAR, with the addition of biking and running, each lasting five minutes, performed by nine subjects using the ACC and GYR from eight smartphones worn at the waist and four smartwatches on the wrists. The MHEALTH dataset also comprises the same basic activities as UCI-HAR, along with additional movements such as waist bending forward, crouching, and front arm raises, performed by ten subjects. Data was collected using three wearable devices equipped with ACC, GYR, and MAG sensors placed on the chest, right wrist, and left ankle, with each activity lasting one minute [49].

Other datasets focus on more complex activities, such as SKODA [50], PAMAP2 [51], and REALDISP [52]. SKODA involves ten activities performed by eight subjects on a car assembly line, with an Inertial Measurement Unit (IMU) placed on the back of their hands. The PAMAP2 dataset includes 18 different activities, ranging from sports (e.g., playing football, rope jumping) to household tasks (e.g., computer work, house cleaning), recorded from nine subjects over nearly ten hours. Each subject wore three IMUs: one on the chest, one on the dominant wrist, and one on the dominant ankle. REALDISP features 33 fitness exercises performed by 17 subjects, with data collected using nine IMUs placed on both arms, both legs, and the back.

3.2 Unsupervised Learning Schemes for HAR

Unsupervised machine learning models are primarily used for exploratory data analysis to identify patterns within the data [13]. In the study by Machado et al. [45], data was acquired as described in the previous section. Researchers extracted features from a tri-axial ACC, resulting in a 180-dimensional feature vector which was subsequently reduced using feature selection techniques. KMeans clustering was then applied to these vectors with various window sizes. A window size of five seconds resulted in the highest ARI of 99.3 % in a subject-specific approach, and 88.6 % in a person-independent approach. The researchers acknowledged the need for longer data acquisitions, more subjects, and activities of varying intensity. Similarly, Kwon et al. [46] acquired data as described earlier. They extracted the mean and standard deviation from the time and frequency domains of the sensor data, then clustered it under two conditions: one where the number of clusters (k) is known and matches the number of activities, and another where k is unknown. In the first scenario, KMeans, GMM, and AGG with average linkage were used, achieving ARIs of 72.0 %, 100 %, and 80.0 %, respectively. With k unknown, the number of clusters obtained was higher than the number of activites. KMeans and AGG achieved close to 80.0 % ARI, while GMM maintained 100 % ARI. DBSCAN was also employed in this scenario, reaching a 90.0 % ARI. In this work, more subjects are needed to further validate the obtained results.

Similar experimental setups to those in [45] and [46] were used in [53] and [54]. In [53], an additional smartwatch and pedometer readings were incorporated. The clustering models used included KMeans, Spectral Clustering, AGG (average and Ward linkage),

DBSCAN, and Mean Shift. KMeans was the best-performing model, followed by AGG with Ward's method and Spectral Clustering. DBSCAN struggled to cluster effectively, discarding many samples as noise, while AGG with average linkage and Mean Shift also performed poorly. No clustering metrics were provided in this study. In [54], an adapted Hidden Markov Model achieved an ARI of 91.4 % for clustering basic activities, which was slightly lower than the accuracy of a random forest and *k*-nearest neighbors.

In recent years, deep learning models have increasingly been used for HAR, either through a combination of supervised and unsupervised methods or using a fully unsupervised approach. For instance, the work in [55] introduced a variational autoencoder trained on unlabeled ACC and GYR data from a wristband. The model, tested on the HHAR dataset, used the embedded vectors as input features for a Random Forest (RF) and k-nearest neighbors, achieving F1 scores of 91.5 % and 90.2 %, respectively, outperforming the 90 % achieved with handcrafted features. In a fully unsupervised approach, Abedin et al. [56] proposed a multi-task autoencoder and evaluated it using the UCI HAR, SKODA, and MHEALTH datasets. They compared traditional clustering models (KMeans, AGG with Ward, complete, and average linkage) on raw signal windows, clustering with embedded vectors, and an end-to-end approach optimizing these vectors for clustering. The optimized embedded vectors yielded the highest test accuracy: 75.4 % on UCI-HAR, 53.48 % on SKODA, and 56.9 % on MHEALTH. Similarly, [57] extracted statistical features from sensor data, using them as input to an autoencoder, where the resulting embedded vectors were used for KMeans clustering. This achieved a RI of 92.1 % on PAMAP2, 65.1 % on REALDISP, and 67.2 % on UCI-HAR, though it still performed worse than the averaging 99 % from supervised methods. In [58], the same approach was implemented with the addition of the cluster labels being used as pseudo-labels to train a deep neural network, which outperformed the precision of fully supervised deep learning models by 2 %.

The studies mentioned above were conducted in general contexts, not focusing specifically on office work activities. Traditional clustering models achieve high scores in scenarios where the activities being studied are basic. However, models' performances tend to decrease when considering complex activities and longer recordings. This can be explained by the lack of clear or repetitive patterns, which allows individuals to perform these activities in varied ways. This variability further increases when considering multiple subjects with different behaviors. Furthermore, existing studies are usually limited to controlled and balanced datasets for model training and testing, not validating their results against unrestricted, real-world settings. To address these issues, this thesis presents a new HAR dataset comprising nine office work activities collected from 19 subjects for training clustering models with varying the complexity of the input data. These are tested over three different experiments: feature selection, cluster stability over time, and impact of cluster imbalance. Lastly, the proposed framework is tested on data from two subjects performing their actual work tasks for an entire workday.

4

Materials and Methods

This chapter presents the proposed framework for HAR in office environments. It begins with a description of the two acquired datasets: a labeled HAR dataset consisting of office tasks performed by 19 subjects, used for model training, and a second dataset used for evaluating the models in a real-world office scenario. The latter comprises data from two subjects who carried out their actual work tasks for over six hours using the PrevOccupAI setup. Next, the pre-processing and feature extraction techniques applied to this data are detailed. The labeled HAR dataset was used to train multiple clustering models of varying complexity across three experiments. Finally, these models were tested on the real-world dataset.

The presented work was implemented in Python [59] using NumPy [60], pandas [61], SciPy [62], scikit-learn [63], TSFEL [64], and biosignalsnotebooks [65, 66] and is available on GitHub[1].

4.1 Definition of Human Activities and Partitioning into Main and Sub-activities

The clustering approach in this study was based on the concept that human activities can be categorized into main activities and sub-activities. This methodology was adapted from Hierarchical Task Analysis (HTA) [67], an ergonomic technique used to break down complex tasks into a hierarchy of main tasks and associated sub-tasks, facilitating a detailed examination of each component.

In an office setting, workers typically engage in tasks that fall into three main categories: sitting, standing, and walking. The goal of this study is to distinguish between these activities. Each main activity can be divided into several sub-activities. These were chosen based on common office practices and feedback from AT employees regarding typical activities executed during daily work. The hierarchical structure of these tasks is outlined in Table 4.1.

[1] https://github.com/SaraLMS/Unsupervised-HAR-Models-for-Characterizing-Office-Tasks

Table 4.1: Hierarchical structure of office activities

Main activities	Sub-activities
Sitting	Working on a computer
Standing	Still Conversing Making/drinking coffee/tea Moving folders/objects inside a tall cabinet
Walking	Slow speed Medium speed Fast speed Upstairs/downstairs

4.2 Datasets

As part of this thesis, two labeled datasets were acquired. The first dataset consists of the human activities described in Table 4.1 acquired from 19 subjects. This dataset is designed to develop unsupervised HAR models that determine which activities people are performing in office environments. The second dataset comprises data from two subjects carrying out their daily work tasks during a whole workday, and was acquired for evaluation of the developed HAR models.

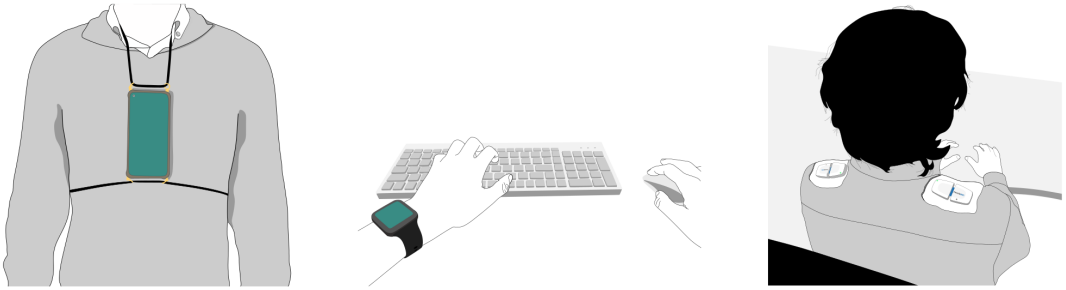
4.2.1 Dataset for HAR model training

4.2.1.1 Experimental Setup

The acquisitions were conducted in an office environment with a group of 20 healthy volunteers, comprising 14 women and 6 men, aged between 19 and 54 years (age: 27.0 ± 8.7 years). Data from subject P017 was removed since, in the sitting acquisition, the subject was not working on a computer. The purpose of the study and the acquisition protocol were thoroughly explained to the participants, and each were provided with an informed consent form, shown in Annex I. The same setup as the PrevOccupAI project was utilized which was approved by the Universidade Nova de Lisboa Ethics Committee and conducted in accordance with the Declaration of Helsinki [11].

4.2.1.2 Sensor Setup

The same sensor placement as the one used in the PrevOccupAI project data acquisitions was employed [11]. The used sensors setup, shown in Figure 4.1, comprises a Xiaomi Redmi Note 9 smartphone placed on the subjects chest, an OPPO 41 mm smartwatch placed on the subjects non-dominant wrist, and two muscleBANs (PLUX Wireless Biosignals) attached to the left and right Trapezius. Data collection was carried out using the PrevOccupAI application [68].



(a) Smartphone (b) Smartwatch (c) MuscleBANs

Figure 4.1: Sensor placement adapted from [11].

The smartphone and the smartwatch run the Android operating system and were used to acquire tri-axial ACC, GYR, and MAG data. ACC and GYR were acquired at 100 Hz and MAG at 50 Hz (restricted by the operating system).

The muscleBAN contains an Electromygraphy (EMG) sensor, as well as a tri-axial ACC and MAG. The data was acquired with a sampling rate of 1000 Hz. The muscleBAN was placed in accordance with the SENIAM guidelines [69]. Before placing the sensor, potential hair at the placement area was removed and the skin was subsequently cleaned with alcohol.

4.2.1.3 Acquisition Protocol

Before acquiring the HAR data, subjects were required to perform an initial recording of a Maximum Voluntary Contraction (MVC) on the trapezius muscles. While seated, they pushed their shoulders upward to resist a downward force applied for approximately five seconds. This was done twice with a 10 seconds rest interval between attempts.

At the beginning of each acquisition, subjects were instructed to perform ten jumps. For jumping, subjects were asked to position their arms parallel to their bodies with their palms facing the hips and fingers pointing toward the floor. Subjects stood straight (body perpendicular to the floor) during the entire jumping sequence. Once the correct posture was executed, subjects stood still for ten seconds, followed by ten vertical short jumps and another ten seconds stop. This jumping sequence was performed to later synchronise the different devices as the jumps would be registered on the ACC of each device.

Following HTA (Section 4.1), a total of 15 minutes, 30 minutes, and 20 minutes were acquired for sub-activities that can be associated to sitting, standing, and walking, respectively. To facilitate the data acquisitions, sub-activities more similar in nature were grouped together. Thus, five acquisition sessions were devised. Each session had clearly defined sections in which the subject was asked to perform a specific sub-activity. To later allow for segmenting each activity, short pauses were introduced between sub-activities that were consecutively acquired. These pauses were ten second stops for tasks involving walking, and ten seconds stops with a jump in the middle for activities performed while

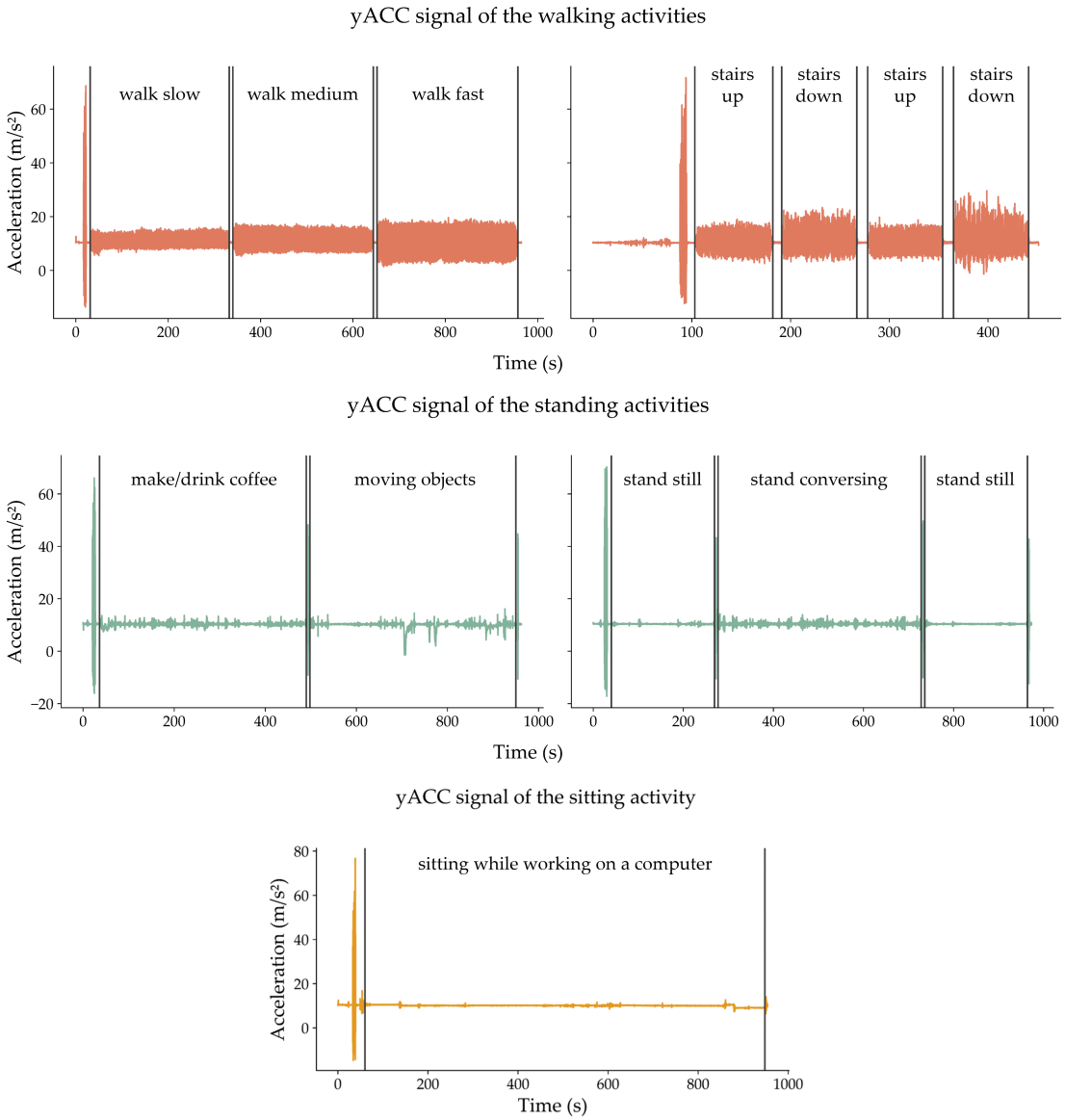


Figure 4.2: Smartphone's y-axis ACC signals for the five recordings.

standing. The sitting activity was continuously performed without any pauses.

The first acquisition session included walking at slow, medium, and fast speeds. Subjects were instructed to walk each speed at their personal slow, medium, and fast pace. Each walking speed was acquired for five minutes (totaling 15 minutes). The second session comprised of walking stairs up and down. This was performed at the subjects preferred walking speed. The acquisition was carried out in four segments of one minute and 15 seconds (totaling five minutes) alternating between going up and down the stairs. The y-axis ACC signals for the these two walking recordings are shown in 4.2.

The third session consisted of two tasks, each lasting seven minutes and 30 seconds, totaling 15 minutes, using a tall cabinet. In the first task, subjects pretended to make and drink coffee or tea, using a cabinet shelf set up with a coffee machine, cups, tea bags, and

a water bottle as a makeshift coffee station. In the second task, subjects moved folders, documents, and other objects within the cabinet, grabbing and examining items before returning them. For the fourth session, subjects stood still and stood while conversing for seven minutes and 30 seconds each (totaling 15 minutes). For standing still, subjects were instructed to stand as if waiting for a meeting or an elevator. This task was divided into two segments for the subjects' comfort. For the standing while conversing, subjects stood while engaging in an active conversation, with multiple topics provided to ensure they remained actively talking. For the two standing sessions, the y-axis ACC signals are presented in 4.2.

The fifth session comprised of sitting while working on a computer for 15 minutes, as shown in Figure 4.2. Subjects were instructed to works as they would normally do at their desks, allowing for phone or stationery use and interaction with others if needed.

4.2.2 Dataset for evaluation of HAR models in a real-world office setting

To evaluate the performance of the purposed models trained on the labeled HAR dataset, a second dataset was acquired. This includes data from two subjects, one female (23 years old) and one male (34 years old). The sensor setup used was the same as described in Section 4.2.1.2.

Data was acquired during a whole workday while the subjects carried out their daily work tasks, comprising six hours and 30 minutes each. The acquisitions were scheduled as in the PrevOccupAI project, described Section 1.1. The main activities were annotated resorting to an application that the subjects used. Three labels were available: sitting, standing, and walking. Subjects selected the activity they were about to perform, therefore saving the initial timestamp of said activity.

4.3 Signal Pre-processing

As mentioned in Section 1.1, this thesis focuses on the usage of smart devices, namely smartphones and smartwatches, due to their higher accessibility compared to other wearable sensors. Therefore, although EMG data an its corresponding MVC was collected, this data was not further used for the development of the HAR models.

The overall pre-processing pipeline for the ACC, GYR, and MAG signals from both the smartphone and smartwatch are shown the Figure 4.3.

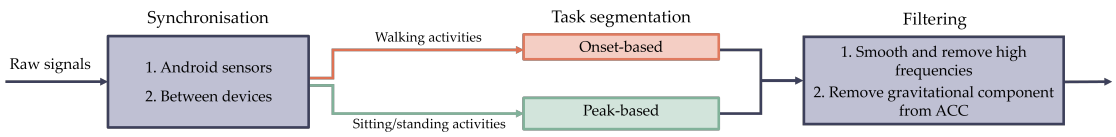


Figure 4.3: Pre-processing pipeline.

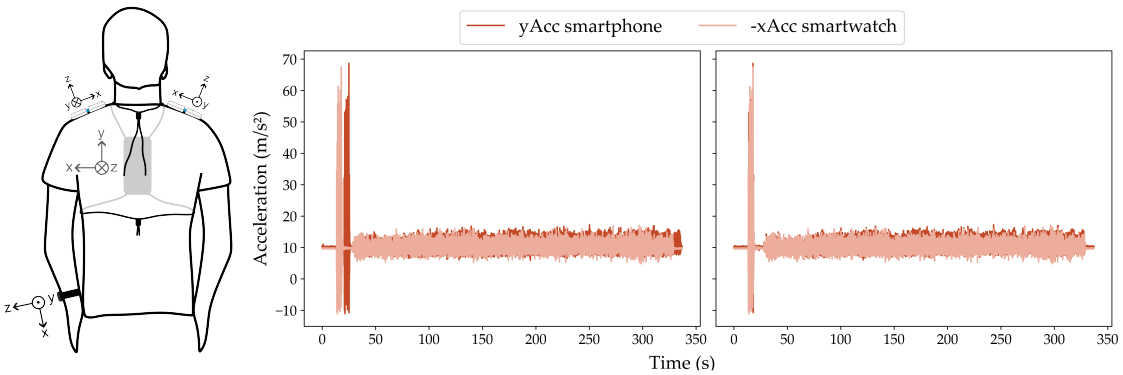
4.3.1 Alignment, Resampling, and Synchronisation

4.3.1.1 Aligning and Resampling Sensors Within a Device

The Android operating system is not primarily designed for continuous data acquisition. As Android optimizes for battery consumption and available processing power, some inconsistencies in the data acquisition process occur. These are mainly sensors not acquiring at a fixed sampling rate, and sensors not starting and stopping to collect data simultaneously, as the system adapts the sampling procedure depending on the current processing needs. Therefore, the signals were cropped according to the last sensor to start and the first to stop acquiring data. Additionally, to ensure all three sensors (ACC, GYR, and MAG) of the smartphone and smartwatch are sampled equidistantly and at the same sampling frequency, the signals were up-sampled to 100 Hz using quadratic interpolation. This process ensures a consistent amount of data from all sensors.

4.3.1.2 Synchronising Between Devices

To utilize the combined sensor data from both the smartphone and smartwatch, these signals need to be synchronised. For this, as described in Section 4.2.1.3, subjects were instructed to jump ten times at the start of each acquisition for synchronisation purposes. These vertical jumps were most evident on the y-axis of the smartphone's ACC and the x-axis of the smartwatch's ACC. The full cross-correlation was computed between these signals, obtaining the shift in samples. This value was cropped from the data of the device that started acquiring first, thereby aligning the signals from both devices. Figure 4.4a illustrates the sensor axes, and Figures 4.4b and 4.4c show the ACC signals before and after synchronisation, respectively.



(a) Sensor axis (b) ACC signals before synchronisation (c) ACC signals after synchronisation

Figure 4.4: Sensor axis and raw segments of walking accelerometer signals before and after synchronisation between devices.

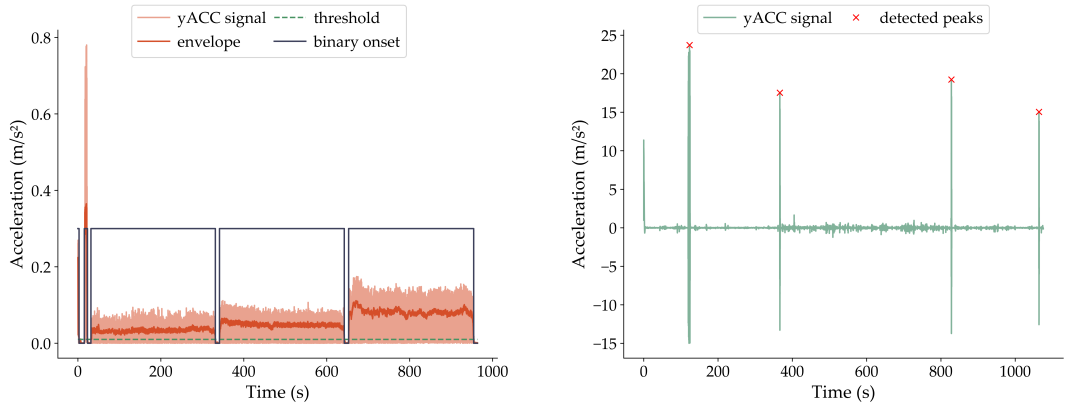
4.3.2 Task Segmentation

As described in Section 4.2.1.3, subjects performed short stop segments to separate tasks within the same recording. These stops are evident in the y-axis of the ACC signal, which was used for segmentation. The first step to identify the start and stop points of each sub-activity was filtering the signal. A median filter with a window length of eleven samples was applied, followed by a Butterworth low-pass filter with a cutoff frequency of 20 Hz. To isolate the gravitational component, another Butterworth low-pass filter with a 0.3 Hz cutoff frequency was applied, and this component was then subtracted from the signal.

An onset-based segmentation approach was employed to extract individual sub-activities for the walking activities (walking slow, medium, fast, upstairs, and downstairs). First, the absolute value of the signal was computed, followed by the application of a root-mean-square filter with a window length of 100 samples to obtain the signal's envelope. The signal was then binarized, setting values above 0.01 m/s^2 to one and values below to zero. The first order discrete difference was calculated to identify the start points (where the difference is one) and stop points (where the difference is minus one). These values were subsequently validated to remove any incorrect detections, primarily caused by the synchronisation jumps. The onset-based task segmentation steps of a walking recording are shown in Figure 4.5a, where the binary onset was scaled to 0.3 m/s^2 for visualization purposes.

To extract sub-activities from standing activities (still, conversing, making/drinking coffee/tea, and moving folders/objects) and to isolate the sitting activity, a peak-based segmentation approach was developed. The *findpeaks* function from SciPy was applied with a peak height of 7 m/s^2 and a minimum distance of 15 000 samples for standing still, conversing, and sitting, and 40 000 samples for the other standing sub-activities. These thresholds were adapted for some subjects. This function detected the peaks from the synchronization jumps and the jumps in the middle of the stop segments. Since these peaks are roughly centered within the ten-second stop segments, the start and stop points were set five seconds before and after each peak. For the first peak, which pertains to the synchronisation jumps, the start was set to 15 seconds after the peak. For the sitting activity, where there were no separation segments, this method was used solely to remove synchronisation jumps. Figure 4.5b shows the peaks found on a standing recording using the peak-based task segmentation.

After isolating each sub-activity, segments were trimmed at both the start and end to remove any residual portions related to synchronisation and stop segments. The trimming duration varied based on the task length, with shorter tasks being trimmed less. For walking upstairs and downstairs, 2.5 seconds were trimmed from each end. For walking slow, medium, and fast speeds, five seconds were trimmed. Standing sub-activities were trimmed by ten seconds, and the sitting activity was trimmed by 30 seconds to remove the sitting down and standing up transitions.



(a) Onset-based segmentation of walking activities

(b) Peak-based segmentation of standing activities

Figure 4.5: Onset- and peak-based task segmentation.

4.3.3 Filtering

After the sub-tasks were segmented, each sub-task signal was filtered to prepare it for feature extraction. The ACC, GYR, and MAG signals for both the smartphone and smartwatch were filtered using the same filters used for the initial step of the task segmentation (Section 4.3.2), which is the same filtering pipeline as described in [70]. The median filter is used for smoothing the signals, while the Butterworth filter removes unnecessary higher frequencies, as daily human activities are commonly characterized by lower frequencies up to 15 Hz.

4.4 Feature Extraction

The TSFEL package was used for feature extraction. Different window sizes, used in [45, 53, 57], were tested, but it was experimentally found that 1.5 seconds, approximately one walking cycle, with 50 % overlap performed the best. The extracted features and correspondent domains, listed in Table 4.2, are in accordance with the TSFEL package. These features were selected based on [46, 57], as well as an analysis of the acquired signals and correspondent Fourier Transforms.

Each window was labeled with a class corresponding to the main activity and a subclass corresponding to the sub-activity. This process resulted in roughly 5 191 instances for each

Table 4.2: Extracted features and respective domains from TSFEL

Statistical	Frequency
Maximum	Median frequency
Minimum	Spectral centroid
Mean	Spectral entropy
Median	
Variance	
Standard deviation	
Interquartile range	

Table 4.3: Number of instances per subclass and main class

Main class	Subclass	Number of instances	
		Subclass	Main class
Sitting	Working on a computer	1199	1199
Standing	Still	599	2396
	Conversing	599	
	Making/drinking coffee/tea	599	
	Moving folders/objects inside a tall cabinet	599	
Walking	Slow speed	399	1596
	Medium speed	399	
	Fast speed	399	
	Upstairs/downstairs	399	

subject, without considering the trimming of the signals described in Section 4.3.2 and some sensors stopping the acquisition early, totaling 98 629 total instances over all subjects. The number of instances for each main class and subclass, per subject, are shown in Table 4.3. Since the dataset is originally imbalanced, the instances were later adjusted to ensure that each main class contained a similar number of windows. The subclasses within each main class were also balanced, with a slight overrepresentation of standing still and stairs. This occurred since these sub-activities were divided into different segments, as described in Section 4.2.1.3, and processed individually.

4.5 Experiments

In the course of this study, the collected data was used for several experiments, which will be explained in the following sections. The unsupervised ML models used in this work were KMeans, GMM, and AGG Ward's linkage. These models are frequently used for HAR and allow for the specification of the number of clusters. As described in Section 4.1, the objective of this work is to distinguish the three main activities from the sub-activities in accordance with HTA. Therefore, for all clustering models the number of clusters was set to three and it was explored whether the sub-activities fall into the clusters representing the main activities. For each of these models, the following clustering approaches were implemented:

- **Subject specific models (SMMs)**: Each subject is clustered individually using a feature set specifically optimized for that subject. The SMMs are used to explore how unsupervised learning schemes perform on each subject individually and to study if there are particular differences between subjects.

- **One-stage general model (1GM)**: Data from all subjects is combined into a single dataset and clustered together. This approach evaluates whether training a clustering model on multiple subjects (subject-independent model) allows it to be effectively applied to each subject individually.

- **Two-stage general model (2GM)**: Each subject is clustered individually using a common feature set. This model explores whether there is a feature set that can be generally applied for HAR.

The contribution of the two smart devices data on the clustering was also studied. For this, each model was implemented using three sensor modalities:

- **Modality smartphone (M_P)**: Using only smartphone sensors.
- **Modality smartwatch (M_W)**: Using only smartwatch sensors.
- **Modality smartphone and smartwatch (M_{P+W})**: Using both smartphone and smartwatch sensors.

Finally, to evaluate the complexity of the input data, the activities were split into two groups:

- **Basic activities**: Composed of three sub-activities, namely standing still, walking medium, and sitting. Considering all sub-activities, these were chosen as the most basic forms of standing, walking, and sitting, commonly studied in the literature.
- **All activities**: Composed of all nine sub-activities shown in Table 4.1. This system is more complex as it captures all various forms of performing the main activities.

4.5.1 Experiment 1 - Feature Selection

The initial experiment conducted on the labeled HAR dataset involved identifying feature sets that achieved the highest ARI accuracy for the models outlined in the previous section. This analysis was performed using the first 80 % of the subjects' data. Two feature selection methods were utilized: a one-stage method for the SMMs and 1GM, and a two-stage method for the 2GM.

4.5.1.1 One-stage Feature Selection

The initial step in this feature selection process involves normalization and elimination of features with low variance and high correlation. For the SMMs, these thresholds were adjusted for each subject to achieve the highest ARI while minimizing the number of features. Regarding the 1GM, considering the complexity of the data from all subjects combined, only features with less than 1 % variance and over 99 % correlation were removed, to retain a broader range of features. The remaining features underwent a forward feature selection approach. Features were initially shuffled and then added iteratively to a sub-dataset that was then passed to the model for clustering. If the addition of a feature did not improve the ARI, it was removed. This procedure was repeated ten times to account for the randomness introduced by the shuffling, thereby ensuring that various combinations were tested. This method produced multiple feature sets along with

Algorithm 1: One-stage feature selection.

Input: Dataset (features in columns, data points in rows) D, variance threshold v_{thresh}, correlation threshold c_{thresh}, iterations n, clustering model M
Output: Feature sets F_{list}, accuracies ARI_{list}

```
D ← Normalize(D);
D ← Remove Low Variance Features(D, v_thresh);
D ← Remove Correlated Features(D, c_thresh);
F_list, ARI_list ← [ ], [ ];
for i ← 1 to n do
    shuffled_features ← Shuffle(D.columns);
    best_ARI ← 0;
    sub_dataset ← [ ];
    best_features_list ← [ ];
    foreach feature in shuffled_features do
        Append feature to sub_dataset;
        Cluster(M, sub_dataset);
        if ARI increases at least 1% then
            best_ARI ← ARI;
            best_features_list ← feature.get_name;
        end
        else
            Remove feature from sub_dataset;
        end
    end
    if best_features_list ∉ F_list then
        Append best_features_list to F_list;
        Append best_ARI to ARI_list;
    end
end
return F_list, ARI_list;
```

their corresponding ARI, as detailed in Algorithm 1. Among the tested feature sets, the one with the highest ARI and fewest features was selected.

4.5.1.2 Two-stage Feature Selection

For the 2GM, an additional feature selection method was implemented. First, the best feature sets are found for each subject using the one-stage feature selection approach described in Algorithm 1. The variance threshold was set to 5 % for basic activities and 1 % for all activities, while the correlation threshold was set to 99 % for both. Then, from the subject-specific feature sets, the n most common features across all subjects were selected to form a final feature set. This feature combination was then used to cluster each subject individually. Figure 4.6 illustrates the two-stage feature selection process for finding the three most common features (orange, yellow, and green) across all subjects. To determine the most suitable number of features for the final feature set, values of $n = 4, 5, 6, 7, 8$ for basic activities, and values of $n = 8, 10, 12, 14$ for all activities, were tested. The selection of these n values was based on an intermediate number of features (between the minimum and maximum) found for the SMMs. Since the one-stage feature selection produces different feature sets with each iteration, this two-stage method was repeated five times for each value of n. The final feature set was selected based on the highest ARI achieved.

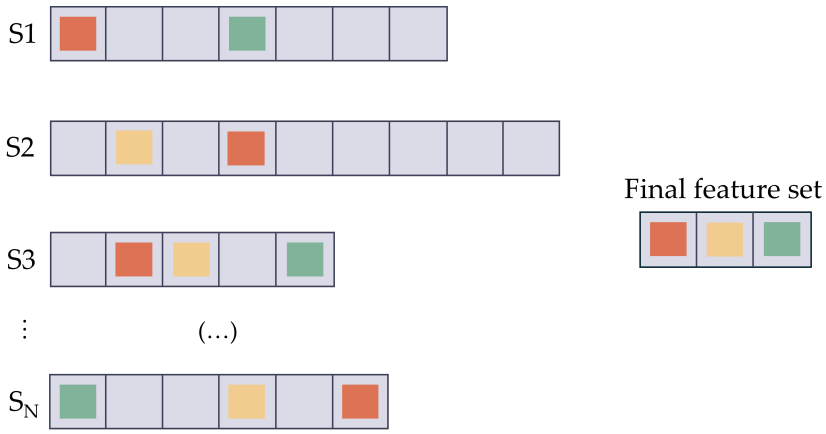


Figure 4.6: Two-stage feature selection scheme.

4.5.2 Experiment 2 - Cluster Stability

Experiment 2 was performed to assess whether the trained clusters were representative enough to be re-used on new, unseen data. Demonstrating this would suggest that the clusters exhibit stability over time, indicating that a person's activity patterns (such as sitting behavior throughout the day or across different days) remain consistent. In an entirely unsupervised scenario, if clustering has been performed and the clusters have been identified for specific activities, they could be reused without the need to re-cluster each time new data is introduced.

For KMeans and GMM, the clusters were initially formed using the first 80 % of the data, and the final 20 % was then assigned to these clusters. Unlike these models, AGG does not support predicting new data without reconstructing the clusters, so the remaining 20 % of the data was clustered anew. For the SMMs and 2GM, the data from each subject was trained and tested individually, while for the 1GM the clusters were formed using 80 % of all subjects' data combined and tested on the remaining 20 % of each subject individually. This data splitting is supposed to simulate having two recordings, one for training and one for testing.

As the described approach has similarities to supervised learning schemes, a comparison to a RF was carried out. For each of the best performing SMMs, 1GM, and 2GM, a RF was trained and tested on the same 80/20 % data split. The features used for the RF were the same as in the respective best performing clustering models. The RF hyperparameters (number of trees and tree depth) for both basic and all activities using grid search. The dataset from the 1GM was used for this optimization since it includes data from all subjects.

4.5.3 Experiment 3 - Cluster Imbalance

The previous experiments were conducted on a balanced dataset. However, in a real-world office setting, workers typically spend more time sitting than walking or standing, leading to class imbalances that can affect clustering performance. To study the impact

of class imbalance on the ARI, the 2GM for basic activities with M_P was applied. Four scenarios were tested, with sitting comprising 33 % (balanced dataset), 50 %, 70 %, and 90 % of the data. The remaining data for walking and standing tasks were equally distributed. In all of these scenarios, the clustering was performed on 100 % of the data.

The imbalanced datasets were constructed by including all instances from the sitting task, while the walking medium and standing still tasks were added in equal proportions to achieve the desired overall distributions. Given that each of these two tasks made up only 25 %, 15 %, or 5 % of the data in the different scenarios, 20 subsets of these proportions were tested, resulting in a average ARI for each of these conditions. The 20 subsets were obtained by partitioning the dataset, with overlap, into subsets. The window size was equal to the chosen percentage.

4.6 Evaluation on real-world labeled dataset

The HAR models were evaluated on the dataset described in Section 4.2.2. The pre-processing (signal alignment, resampling, and filtering) and feature extraction methods used were the same that were applied for the HAR dataset, described in Sections 4.3 and 4.4. Clustering was performed using the 2GM for all activities with M_P. Given the more natural and unrestricted scenario, an analysis of the silhouette coefficient was conducted to determine the optimal number of clusters for each clustering model. The clustering model with the highest silhouette coefficient was used for the clustering. The data was clustered twice: once with three clusters, consistent with the previous experiments on the labeled HAR dataset, and once using the number of clusters suggested by the silhouette coefficient analysis.

To better understand the clustering results, an interpretable clustering approach was implemented [71], where a decision tree model was trained using the cluster labels as the target class labels. This method enables the examination of feature importance at each split in the tree, providing insights into which features are most influential in distinguishing between clusters. The depth of the tree was set to three to limit the analysis to the most significant features, which are typically those near the top of the tree. Finally, Kernel density estimation (KDE) plots were utilized to analyse the distributions of the most important features for the decision tree.

5

Results

The results of this thesis are presented in this chapter. It begins with a description of the results obtained for the proposed HAR models across the three different experiments, described in Chapter 4. Lastly, these models are tested on the real-world labeled dataset.

5.1 Experiment 1 - Feature Selection

5.1.1 One-stage Feature Selection

The distribution of subjects by the number of features needed to achieve the highest ARI for the SMMs is shown in Table 5.1. When comparing modalities over the three models for the basic activities, the feature sets for M_P and M_{P+W} require two to four features for the majority of subjects. In contrast, M_W generally requires more features than the other two modalities. For M_W, although most subjects require up to seven features, some subjects need more than eight. There are eight instances in which a subject needs more than 13 features, when summing up the number of subjects across all modalities and models. For all activities, the number of features needed increases across all clustering models and sensor modalities. Overall, more than half of the subjects' feature sets require over eight features. Additionally, there are 37 instances where more than 13 features are needed to distinguish all activities. For both basic and all activities, at least one MAG feature is present in all feature sets across the three sensor modalities and clustering models. For feature sets with more than two features, there are usually multiple features, such as the MAG maximum, minimum, or mean. The most common axis of the MAG for these

Table 5.1: Distribution of subjects by number of features selected for the subject specific models for basic and all activities.

Number of Features	KMeans						AGG						GMM					
	M_P		M_W		M_{P+W}		M_P		M_W		M_{P+W}		M_P		M_W		M_{P+W}	
	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All
2 - 4	13	—	7	—	13	—	12	—	8	1	14	—	14	—	10	5	9	—
5 - 7	3	4	4	1	3	1	3	6	2	9	5	2	—	7	2	6	10	2
8 - 12	3	9	8	16	2	4	3	12	8	8	—	12	2	10	5	7	—	12
> 13	—	6	—	2	1	14	1	1	1	1	—	5	3	2	2	1	—	5

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch.

Table 5.2: ARI (%) results of Experiment 1 for subject specific models for basic and all activities.

Subject	KMeans						AGG						GMM					
	M_P		M_W		M_{P+W}		M_P		M_W		M_{P+W}		M_P		M_W		M_{P+W}	
	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All
P001	86.6	76.3	94.7	72.3	99.4	88.7	82.9	77.0	100	73.2	99.4	92.3	97.5	93.8	95.6	52.6	95.9	98.0
P002	87.7	88.6	84.0	66.7	95.9	97.4	96.8	93.4	79.4	67.0	100	99.1	100	96.6	62.0	68.5	100	98.7
P003	100	84.4	100	75.0	100	97.2	100	96.1	100	81.5	100	99.2	100	85.2	99.7	74.0	100	97.6
P004	78.3	76.4	100	66.2	100	83.9	76.6	64.0	100	72.4	100	78.9	94.7	94.6	99.7	77.3	100	97.2
P005	100	88.5	95.9	47.1	100	82.2	100	97.5	95.6	51.1	100	85.1	100	95.9	95.9	49.6	100	86.9
P006	100	78.7	99.3	81.0	100	91.4	100	81.0	99.7	69.6	100	89.3	100	91.8	97.3	71.4	100	91.3
P007	100	86.8	81.1	74.3	100	98.6	100	77.3	63.9	75.8	100	99.9	100	85.7	63.9	65.2	100	95.7
P008	100	84.1	86.9	62.0	100	84.0	100	82.0	75.6	60.3	100	84.2	100	77.6	81.1	51.3	100	79.3
P009	100	77.4	98.4	69.5	100	75.4	100	79.2	98.1	62.8	100	74.8	100	78.7	96.2	64.5	100	78.2
P010	98.6	91.0	97.0	62.7	98.8	86.9	99.5	93.8	98	62.6	99.5	83.8	98.8	86.9	97.0	62.6	99.5	85.6
P011	100	94.0	97.8	51.5	100	98.5	100	81.7	96.8	52.3	100	76.9	100	90.6	94.4	58.0	100	89.8
P012	100	76.6	96.2	45.3	100	76.5	100	77.5	96.9	48.2	100	79.7	99.7	71.3	93.5	48.8	100	80.5
P013	100	88.5	99.7	80.9	100	96.2	100	97.4	100	82.2	100	96.6	100	97.3	95.6	90.1	100	97.5
P014	100	76.3	95.3	63.8	100	80.0	100	69.5	97.1	56.8	100	77.3	100	70.3	91.6	51.3	100	80.2
P015	100	90.6	90.8	52.8	100	91.3	100	97.0	89.4	57.3	100	97.5	98.1	90.9	80.7	58.5	98.1	97.5
P016	74.8	54.0	90.7	54.6	98.1	60.0	77.3	55.3	92.2	52.1	98.1	55.9	71.7	55.1	91.4	50.0	98.7	61.2
P018	96.2	91.3	87.6	47.1	98.8	95.8	97.8	95.8	91.2	37.2	97.5	97.7	97.8	92.9	74.1	42.0	95.7	94.8
P019	100	78.4	66.1	65.8	100	82.0	100	79.4	69.2	71.6	100	78.2	100	75.9	67.8	65.3	99.7	79.8
P020	100	81.7	99.0	58.2	100	83.2	100	96.1	98.1	67.5	100	98.3	96.5	95.1	95.9	63.8	100	93.6
Mean	95.9	82.3	92.7	63.0	**99.5**	**86.8**	96.4	83.7	91.6	63.2	**99.7**	**86.6**	97.6	85.6	88.1	61.3	**99.3**	**88.6**
SD	7.7	8.9	8.4	10.8	1.0	9.7	7.7	12.1	10.9	11.6	0.7	11.4	6.3	11.1	12.1	11.6	1.3	9.7

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch.

features is the z-axis. The feature sets obtained for the SMMs are shown in Annex II.

The ARI results corresponding to the number of features shown in Tables 5.1 and the sets presented in Annex II, can be found in Table 5.2. For basic activities, the mean ARI is generally above 90 % across all modalities and models, except for the GMM with M_W, which achieves 88.1 %. Subject specific ARI values range from a minimum of 62.0 % (for Subject P002 with GMM and M_W) to 100 %. The AGG with M_{P+W} achieves the highest mean ARI at 99.7 % with the lowest Standard deviation (SD) of 0.7 %. The M_{P+W} consistently achieves the highest mean ARI, exceeding 99 % across all clustering models. Compared to M_{P+W}, mean model performance decreases for M_W and M_P by 6.8 % and 3.6 % for KMeans, 8.1 % and 3.3 % for AGG, and 11.2 % and 1.7 % for GMM, respectively. Comparing M_P and M_W, M_P performs better or similar across all subjects, except for P004, and P016, where it performs worse than M_W for the three clustering models.

For all activities, the lowest ARI is 37.2 % (Subject P018 for AGG M_W), while the highest is 99.9 % (Subject P007 for AGG M_{P+W}). The model with the best performance is GMM with M_{P+W}, with a mean ARI of 88.6 % and a SD of 9.7 %. M_{P+W} remains the best sensor modality across all clustering models, though, compared to basic activities, the mean ARI drops by 12.7 % for KMeans, 13.1 % for AGG, and 10.3 % for GMM. Compared to M_{P+W}, the mean ARI for M_W and M_P decreases by 23.8 % and 4.5 % for KMeans, 23.4 % and 2.9 % for AGG, and 27.3 % and 3.0 % for GMM. Additionally, the SDs generally increase across all clustering models and sensor modalities.

For the 1GM feature selection, the number of features and ARI results for the three clustering models for the feature sets in Annex II, are shown in Table 5.3. Unlike the SMMs, the number of features needed to achieve the best ARI does not differ as much, overall, between basic and all activities. Additionally, M_W consistently requires the fewest features across both basic and all activities, which is the opposite of the SMMs results.

Table 5.3: ARI (%) and number of features (#F) of Experiment 1 for the one-stage general model.

Activities	KMeans						AGG						GMM					
	M_P		M_W		M_{P+W}		M_P		M_W		M_{P+W}		M_P		M_W		M_{P+W}	
	ARI	#F	ARI	#F	ARI	#F	ARI	#F	ARI	#F	ARI	#F	ARI	#F	ARI	#F	ARI	#F
Basic	66.0	12	56.7	11	**77.2**	9	54.0	4	47.3	2	56.6	7	59.6	8	45.6	3	62.9	4
All	53.6	15	29.4	7	52.2	16	**56.2**	6	37.1	3	51.1	8	56.0	7	36.5	3	53.7	11

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch.

Regarding the clustering results for the 1GM, for basic activities, KMeans M_{P+W} achieves the highest accuracy at 77.2 %. M_{P+W} also performs the best for AGG, reaching 56.6 %, and for GMM, achieving 62.9 %. When evaluating all activities, accuracy decreases across all clustering models and sensor modalities. The highest ARI is 56.2 % with AGG M_P. M_W achieves low ARI values reaching 29.4 % for KMeans, 37.1 % for AGG, and 26.5 % for GMM.

5.1.2 Two-stage Feature Selection

The 2GM feature selection results for basic activities are depicted in Figure 5.1. The highest mean ARI for basic activities are as follows: KMeans with M_P using eight features (91.4 %) and M_{P+W} using six features (91.6 %); AGG with M_{P+W} using eight features (92.1 %); and GMM with M_P using eight features (90.7 %) and M_{P+W} using seven features (93.4 %), which was the highest ARI overall. The M_W resulted in the lowest scores of all models, achieving ARIs of 83.8 % (with seven features) for KMeans, 82.5 % (with eight features) for AGG, and 79.2 % (with eight features) for GMM. The 2GM results for all activities are shown in Figure 5.2. Compared to the more static basic activities, the overall ARI is lower, similar to the SMMs and 1GM results, with more features required, as observed with the SMMs. M_W consistently demonstrated lower performance across all clustering models, while M_P and M_{P+W} produced similar results. For KMeans, the highest ARI with M_P is with ten features (67.1 %), M_W is with 14 features (50.0 %), and M_{P+W} is

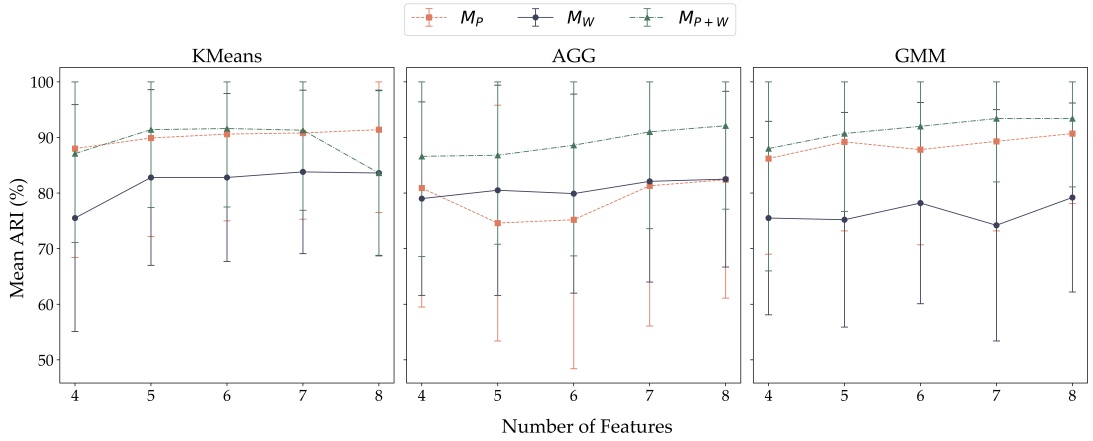


Figure 5.1: Mean ARI (%) over the number of features for the two-stage general model for basic activities. Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **AGG** - Agglomerative Clustering; **GMM** - Gaussian Mixture Model.

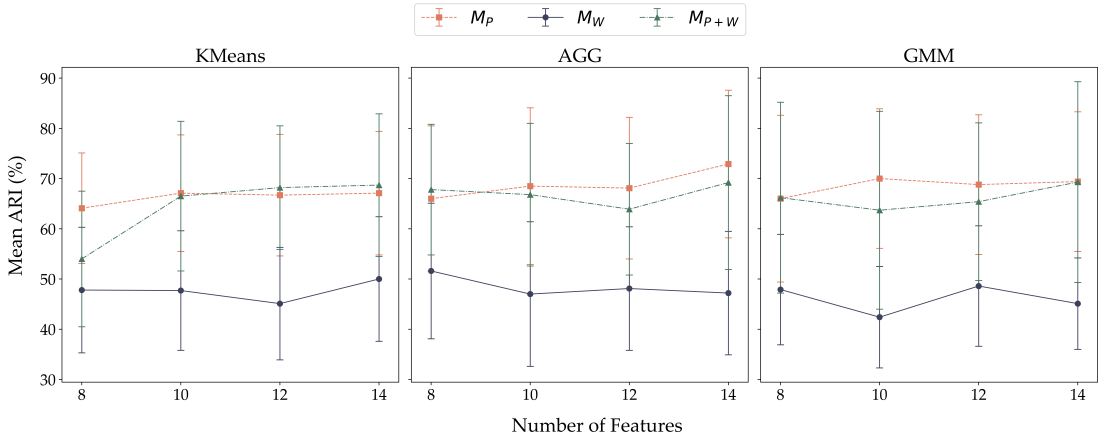


Figure 5.2: Mean ARI (%) over the number of features for the two-stage general model for all activities. Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **AGG** - Agglomerative Clustering; **GMM** - Gaussian Mixture Model.

with 14 features (68.2 %). For AGG, the highest ARI with M_P is with 14 features (72.9 %), which is the highest ARI across all clustering models and modalities. With M_W, the best score is with eight features (51.6 %), and for M_{P+W} is with 14 features (69.2 %). Lastly, for GMM, the highest ARI is achieved with M_P using ten features (70.0 %), M_W using 12 features (48.6 %), and M_{P+W} using 14 features (69.3 %). Overall, the SDs for both basic and all activities are higher compared to the SMMs.

For both basic and all activities, the models with fewer features are chosen when the ARI improvement was only around 2 % or less. The number of features and respective ARI selected for each clustering model and modality are shown in Annex II. These were the models used for the 2GM in the following experiments.

5.2 Experiment 2 - Cluster Stability

5.2.1 SSM Results

The clustering results of Experiment 2 for the SMMs are presented in Table 5.5. For basic activities, the highest mean ARI is 94.4 % with KMeans M_{P+W}. This modality also reaches 91.2 % with AGG and 90.0 % with GMM. Comparing to the M_{P+W} results of Experiment 1, the mean ARI drops 5.1 %, 8.5 %, and 9.3 % for KMeans, AGG, and GMM, respectively. The maximum subject-specific ARI is 100 %, while the minimum is 0.0 %

Table 5.4: Selected number of features (#F) and respective ARI (%) for the two-stage general model.

Activities	KMeans						AGG						GMM					
	M_P		M_W		M_{P+W}		M_P		M_W		M_{P+W}		M_P		M_W		M_{P+W}	
	ARI	#F	ARI	#F	ARI	#F	ARI	#F	ARI	#F	ARI	#F	ARI	#F	ARI	#F	ARI	#F
Basic	89.9	5	82.8	5	91.4	5	80.9	4	80.5	5	91.0	7	89.2	5	78.2	6	93.4	7
All	67.1	10	47.8	8	66.5	10	72.9	14	51.6	8	67.8	8	70.0	10	47.9	8	69.3	14

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch.

Table 5.5: ARI (%) results of Experiment 2 for the subject specific models for basic and all activities.

Subject	KMeans						AGG						GMM					
	M_P		M_W		M_{P+W}		M_P		M_W		M_{P+W}		M_P		M_W		M_{P+W}	
	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All
P001	100	65.4	90.1	54.1	93.6	70.7	83.4	53.4	83.6	18.8	71.3	44.4	57.8	24.9	97.5	50.5	100	17.7
P002	100	94.3	55.8	68.9	98.7	92.9	100	95.8	44	48.5	100	91.6	55.1	39.3	51.3	65.3	98.7	0.1
P003	100	70.0	88.1	52.2	100	92.8	100	90.2	100	40.9	100	94.7	100	55.3	100	37.2	100	15.6
P004	74.9	36.3	100	66.9	100	76.3	63.5	32.5	100	54.9	100	73.5	57.1	56.8	97.5	76.6	100	27.7
P005	100	67.5	91.3	60.1	100	86.0	100	47.6	100	49.4	100	42.0	100	70.0	40.0	32.9	63.2	56.9
P006	100	55.3	98.7	66.9	100	87.5	100	63.4	97.4	56.8	100	67.2	100	49.1	57.8	24.9	97.4	47.9
P007	100	80.6	70.1	65.9	100	94.1	100	35.0	63.7	64.3	100	59.4	100	79.7	63.7	65.5	100	58.9
P008	100	82.5	77.3	52.6	100	94.0	100	88.2	73.3	62.1	100	73.8	100	52.9	0.0	46.8	56.1	53.3
P009	100	66.7	95.0	60.9	100	74.6	100	67.2	92.6	65.5	100	58.7	100	40.8	91.4	47.6	100	43.9
P010	98.1	48.8	62.0	47.2	82.2	89.8	68.1	77.9	91.5	60.4	91.5	43.2	77.2	0.2	57.2	44.9	61.6	40.6
P011	100	79.1	63.2	66.6	90.2	67.2	77.7	45.0	63.7	20.9	63.7	47.2	100	44.0	63.2	22.1	100	43.9
P012	100	71.6	65.8	30.0	100	68.9	100	60.7	66.9	15.8	60.5	50.4	97.5	42.8	10.5	41.2	100	79.3
P013	87.9	61.9	57.6	77.3	63.5	95.4	98.7	94.4	95.0	58.0	100	98.1	100	54.5	100	37.8	100	48.0
P014	100	53.5	72.9	61.1	100	79.6	100	61.6	63.2	54.1	100	88.9	100	50.6	67.9	61.8	100	54.2
P015	100	85.6	100	50.8	100	90.5	100	77.4	100	40.4	100	98.6	97.5	37.2	64.7	37.9	97.5	94.5
P016	62.4	48.9	86.5	52.5	98.7	61.9	44.3	54.7	77.2	58.1	93.8	56.3	53.9	9.8	75.1	46.3	98.7	62.0
P018	53.4	57.6	40.9	25.1	69.3	70.6	51.5	84.8	59.6	19.6	76.5	68.1	56.8	53.9	60.2	5.2	50.3	51.0
P019	92.4	79.1	90.0	73.7	97.5	77.8	96.2	72.7	86.6	37.6	100	64.7	84.7	65.4	55.4	47.1	96.2	44.3
P020	100	60.3	97.5	65.5	100	66.1	100	41.1	97.5	55.4	75.0	64.5	51.2	29.5	100	46.5	90.1	0.6
Mean	93.1	66.6	79.1	57.8	**94.4**	**80.9**	88.6	65.5	81.9	46.4	91.2	67.6	83.6	45.1	66.0	44.1	90.0	44.2
SD	13.6	14.3	17.3	13.1	10.6	10.9	18.0	19.5	17.0	16.1	13.6	16.5	20.1	18.9	27.9	16.2	16.9	23.3

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch.

(Subject P008, GMM with M_W). For all activities, the maximum subject-specific ARI is 98.8 % with AGG (Subject P015 with M_{P+W}), while the lowest is 0.1 % with GMM (Subject P002 with M_{P+W}). The highest mean ARI is 80.9 % with KMeans M_{P+W}. M_{P+W}, which was the best-performing modality in Experiment 1, shows ARI values of 80.9 %, 67.6 %, and 44.2 % for KMeans, AGG, and GMM, respectively. These represent decreases of 5.9 % for KMeans, 19.0 % for AGG, and 44.4 % for GMM. In both basic and all activities, the SDs for all sensor modalities and clustering models are generally much higher compared to Experiment 1.

5.2.2 1GM Results

Table 5.6 shows the results for the 1GM cluster stability experiment. The best performing models for both basic and all activities remain the same as in Experiment 1. The 1GM reaches 72.5 % mean ARI with KMeans M_{P+W}, for basic activities, with a decrease of 4.7 % compared to Experiment 1. For all activities, AGG with M_P was the best model with 55.3 %, similar to Experiment 1, with a drop of 0.9 %. The SDs for the three clustering models and sensor modalities reach, in most cases, lower values when compared to the SMMs results (Table 5.5).

Table 5.6: Mean ARI (%) results of Experiment 2 for the one-stage general model.

Activities	KMeans			AGG			GMM		
	M_P	M_W	M_{P+W}	M_P	M_W	M_{P+W}	M_P	M_W	M_{P+W}
Basic	67.3 ± 15.2	71.6 ± 18.5	**72.5 ± 15.7**	52.5 ± 4.0	50.8 ± 7.4	58.2 ± 11.3	53.2 ± 6.3	45.6 ± 11.2	59.0 ± 19.9
All	52.3 ± 4.7	34.7 ± 8.8	52.0 ± 4.8	**55.3 ± 18.7**	26.2 ± 10.4	39.7 ± 23.7	41.7 ± 10.4	15.2 ± 11.8	49.4 ± 13.6

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch.

5.2.3 2GM Results

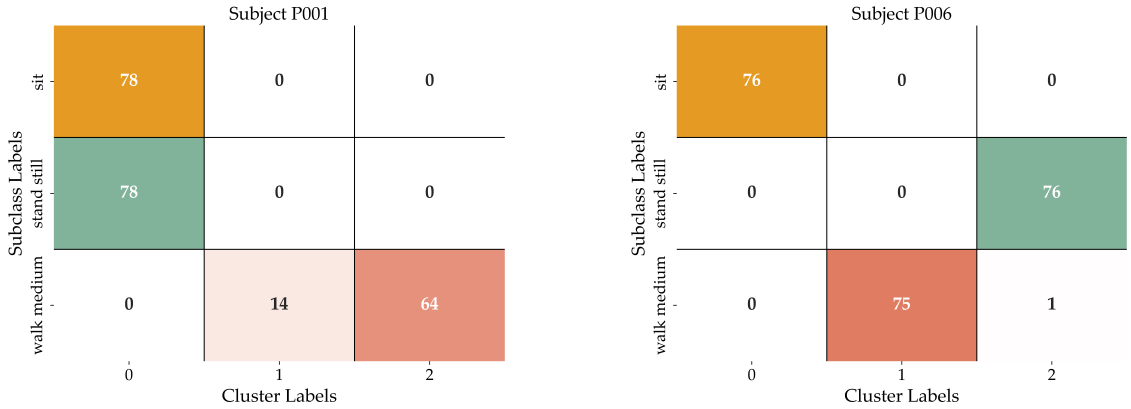
The Experiment 2 clustering results for the 2GM are shown in Table 5.7. For this Experiment, the number of features used for clustering were those of the simpler models, shown in Table 5.4. Similar to Experiment 1, KMeans achieves a mean ARI above 90 % with M_P and M_{P+W}. Although KMeans with M_{P+W} achieves the highest score, M_P is considered the best modality for basic activities since adding the smartwatch sensors only increases 0.3 %. Both KMeans and AGG show overall drops in mean ARI results. For M_P and M_{P+W} these drops are less than 1.5 % comparing to Experiment 1. For M_W, the mean ARI decreased 5.9 % and 3.6 % for KMeans and AGG, respectively. GMM experiences larger drops in mean ARI, with 15.9 % for M_P, 17.1 % for M_W, and 18.3 % for M_{P+W}. For all activities, the results follow the same trend as the SMMs, with a general decrease in mean ARI. The best performance is AGG with M_P, reaching 67.0 % (dropped by 0.8 % compared to Experiment 1).

Given that KMeans with M_P has the best ARI for basic activities, two confusion matrices for a low ARI subject (P001) and a subject with high ARI (P006) are shown in Figure 5.3. Comparing both subjects, it can be observed that for P001, shown in Figure 5.3a, the sitting and standing still points are merged into one cluster, while the walking medium instances are mostly in their own cluster. On the contrary, for subject P006, shown in Figure 5.3b, all three sub-activities are completely separated. The model with the highest performance for all activities is AGG with M_P. Confusion matrices for subjects corresponding to a low ARI (P016) and high ARI (P015) are presented in Figure 5.4. In the confusion matrix of subject P016, shown in Figure 5.4a, all instances of sitting, standing still, standing while conversing (stand gestures), and walking slow are grouped into the same cluster. For subject P015, shown in Figure 5.4b, the three main activities are separated.

Table 5.7: ARI (%) results of Experiment 2 for the two-stage general model for basic and all activities.

Subject	KMeans						AGG						GMM					
	M_P		M_W		M_{P+W}		M_P		M_W		M_{P+W}		M_P		M_W		M_{P+W}	
	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All	Basic	All
P001	49.5	66.6	88.9	55.2	91.4	59.5	44.8	69.9	100	41.0	100	43.3	49.4	0.0	84.9	52.0	98.7	39.2
P002	100	95.0	56.5	64.1	100	94.0	100	95.8	66.4	68.6	100	90.9	92.6	0.0	43.3	0.0	70.2	54.2
P003	100	71.6	100	40.2	100	50.2	100	84.3	100	68.8	100	47.2	100	52.1	91.5	7.0	100	54.3
P004	63.1	40.3	100	59.9	65.2	53.2	64.1	22.6	100	58.5	100	47.2	65.8	55.7	100	27.3	67.0	28.6
P005	63.7	58.3	73.3	30.1	98.7	48.0	100	58.3	71.7	63.9	100	55.0	56.9	55.4	56.5	0.0	63.7	56.8
P006	98.7	55.4	89.8	60.8	98.7	62.8	100	67.2	97.4	37.6	100	66.7	56.9	55.0	86.1	46.9	100	1.5
P007	100	57.2	63.7	64.3	100	89.6	100	67.2	63.7	61.3	100	54.8	98.7	46.2	56.9	45.0	100	58.5
P008	100	82.7	76.8	49.5	56.9	75.8	100	89.8	76.8	46.7	77.7	91.8	56.1	50.1	12.0	33.9	49.0	54.1
P009	100	65.4	71.5	61.1	100	64.1	100	66.5	92.6	63.6	100	63.6	95.0	4.6	89.1	49.6	64.9	34.9
P010	97.2	67.5	62.9	34.0	62.2	55.1	68.1	55.0	58.1	38.3	63.3	67.4	91.7	49.6	58.4	23.0	61.1	54.0
P011	100	66.6	64.3	60.8	100	67.2	98.7	67.3	64.3	55.9	64.3	50.3	100	49.5	63.2	61.2	68.7	96.3
P012	100	65.3	63.2	33.7	91.6	67.6	100	68.1	57.4	43.6	58.2	43.0	63.3	20.2	18.4	35.6	56.9	79.0
P013	100	87.1	95.0	0.6	100	55.3	100	67.5	100	63.6	100	67.6	100	0.1	62.7	0.0	100	46.0
P014	100	61.3	85.3	44.1	100	66.0	100	62.1	61.2	53.9	100	66.6	94.9	50.5	46.9	41.7	70.2	65.8
P015	100	80.9	100	17.2	100	61.4	49.1	98.3	100	55.5	100	97.3	37.3	0.0	68.0	0.0	77.2	51.1
P016	67.3	35.7	72.1	21.5	65.0	50.8	44.4	32.1	33.8	35.1	98.7	34.0	59.3	35.7	50.3	32.1	93.8	54.3
P018	70.7	53.9	46.6	0.3	85.7	48.2	51.0	70.0	38.4	32.5	60.7	62.4	56.8	51.0	44.6	7.7	48.1	45.6
P019	100	73.5	51.9	64.1	100	61.6	42.2	64.9	80.0	62.0	90.2	51.1	84.7	75.5	27.6	20.9	76.7	2.2
P020	100	46.1	100	41.4	100	64.2	45.6	65.4	100	50.6	92.6	63.6	32.8	18.2	100	0.8	60.8	51.8
Mean	**90.0**	64.8	76.9	42.3	90.3	62.9	79.4	**67.0**	76.9	52.7	89.8	61.2	73.3	35.2	61.1	25.5	75.1	48.8
SD	16.7	15.0	17.4	20.4	15.0	12.3	24.8	27.7	21.6	11.5	15.5	16.8	22.4	23.8	25.9	20.2	17.7	21.6

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch.



(a) Subject with low ARI (49.5 %). (b) Subject with high ARI (98.7 %).

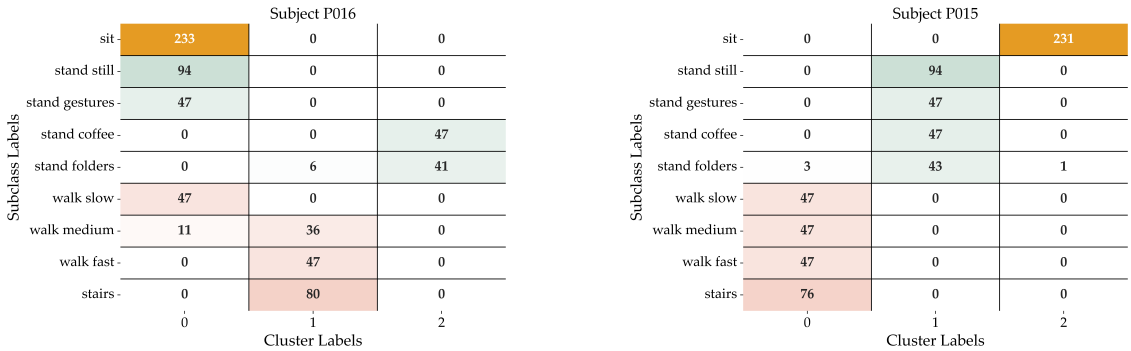
Figure 5.3: Confusion matrices for the two-stage general model for basic activities using KMeans with modality smartphone.

5.2.4 Comparative RF

The optimized parameters for the RF, found using grid search, are shown in Annex III. The RF outperforms the best clustering models for both basic and all activities. Nevertheless, for basic activities the SMMs and the 2GM have comparative results with the clustering. Unlike the clustering models, the RF does not experience significant decreases in accuracy when moving from basic to all activities, consistently achieving over 95 % for the SMMs and 2GM. However, as with the clustering models, SMMs are the top-performing models for the RF, followed by the 2GM and the 1GM, which had noticeably lower accuracy. These results are shown in Table 5.8.

5.3 Experiment 3 - Cluster Imbalance

To assess the impact of cluster imbalance on a well-performing model, the 2GM for basic activities with M_P was selected. For this experiment, the clustering was performed



(a) Subject with low ARI (32.1 %). (b) Subject with high ARI (98.3 %)

Figure 5.4: Confusion matrices for the two-stage general model for all activities using AGG with modality smartphone.

Table 5.8: Comparison of the best clustering models mean ARI (%) and RF mean accuracy (%) on the subject specific models (SSMs), one-stage (1GM) and two stage (2GM) general models.

Models	Basic activities		All activities	
	Clustering	RF	Clustering	RF
SSM	94.4 ± 10.6[a]	**99.5 ± 1.9**	80.9 ± 10.9[a]	**97.4 ± 4.4**
1GM	72.5 ± 15.7[a]	**85.3 ± 15.6**	55.3 ± 18.7[b]	**74.0 ± 10.0**
2GM	90.0 ± 16.7[c]	**98.7 ± 3.2**	67.0 ± 27.7[b]	**95.9 ± 5.6**

[a]KMeans M_{P+W}, [b]AGG M_P, [c]KMeans M_P. Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch.

on 100 % of the data. The influence of data imbalance on the ARI for the three clustering models is shown in Figure 5.5. When the dataset is balanced (33 % sitting class), the mean ARI results are 92.5 % for KMeans, 78.3 % for AGG, and 89.9 % for GMM. When the proportion of sitting class increases to 50 %, the mean ARI scores are 52.9 % for KMeans, 83.4 % for AGG, and 50.9 % for GMM. Comparing to the balanced results, the scores drop 39.6 % for KMeans and 39.0 % for GMM, while AGG increases 5.1 %. With 70 % of sitting class, the mean ARI for KMeans, GMM, and AGG reach 67.1 % (-25.4 %), 81.9 % (+3.6 %), and 64.8 % (-25.1 %), respectively. Finally, when the sitting class represents 90 % of the data, the mean ARI declines to 63.8 % (-28.7 %) for KMeans, 65.9 % (-12.4 %) for AGG, and 62.9 % (-27.0 %) for GMM. The SD for each clustering model is generally high, with the highest being 29.8 % for AGG at 90 % sitting class and the lowest at 11.4 % for KMeans at 50 % sitting class. SDs are particularly high when the sitting class reaches 90 %, with values of 24.3 %, 29.8 %, and 24.1 % for KMeans, AGG, and GMM, respectively. Overall, the SD tends to increase with a higher class imbalance.

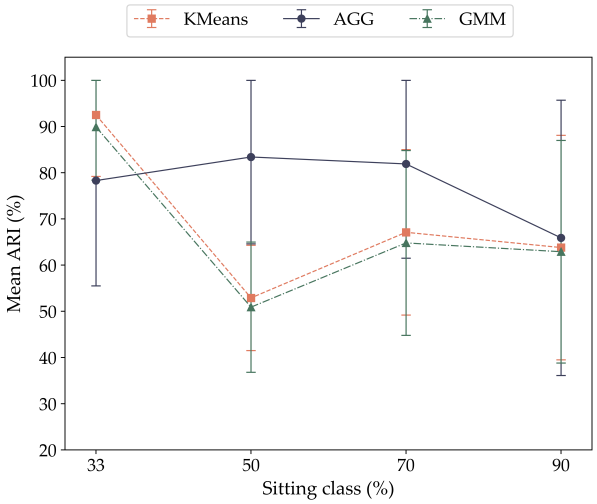
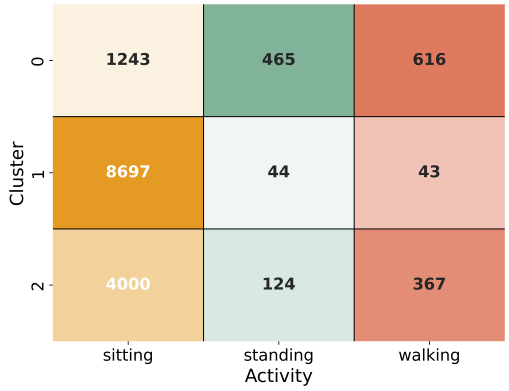
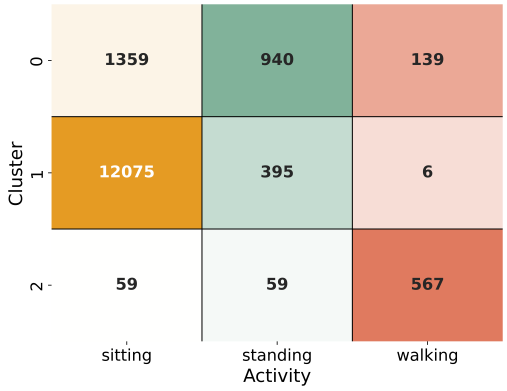


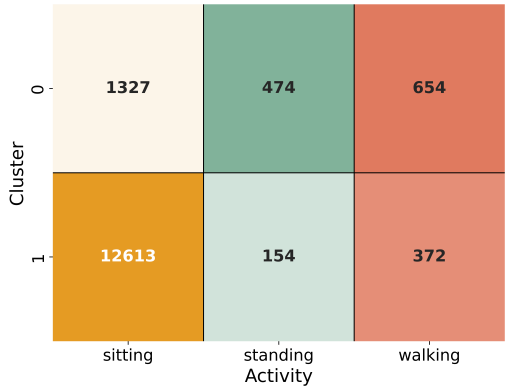
Figure 5.5: Influence of data imbalance on the performance of the two-stage general model with modality smartphone for basic activities. Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch.



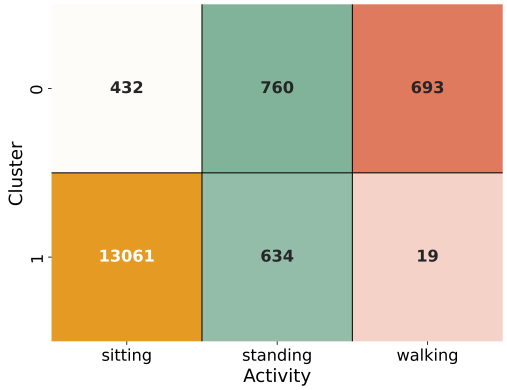
(a) Subject 1 - three clusters (16.9 % ARI).



(b) Subject 2 - three clusters (51.9 % ARI).



(c) Subject 1- two clusters (41.2 % ARI).



(d) Subject 2 - two clusters (63.0 % ARI).

Figure 5.6: Confusion matrices of subject 1 and 2 with two and three clusters.

5.4 Evaluation on a real-world labeled dataset

To evaluate the best performing models in a real-world scenario, a labeled dataset was acquired from two subjects, with a total recording time of six hours and 30 minutes each. The class distribution annotations revealed that subject 1 spent 89.4 % of the time sitting, 4.1 % standing, and 6.6 % walking, while subject 2 spent 86.5 % sitting, 8.9 % standing, and 4.6 % walking.

The 2GM with M_P was selected for this experiment based on the results of Experiment 2, since, as seen in Table 5.8, it was the best-performing general model for basic and all activities. A silhouette coefficient analysis was conducted on the collected data. The highest scores obtained are with KMeans and two clusters, reaching 61.3 % and 57.1 % for subject 1 and 2, respectively. Based on these results, the 2GM with M_P for KMeans was used for clustering the new data into three clusters, as explained in Section 4.5, and into two clusters, as this achieved the highest silhouette scores.

Figure 5.6 presents the confusion matrices and ARI for both subjects with three and two clusters. For subject 1, the confusion matrix with three clusters is shown in Figure 5.6a. For this subject, three clusters achieve an ARI of 16.9 %, with sitting instances being the

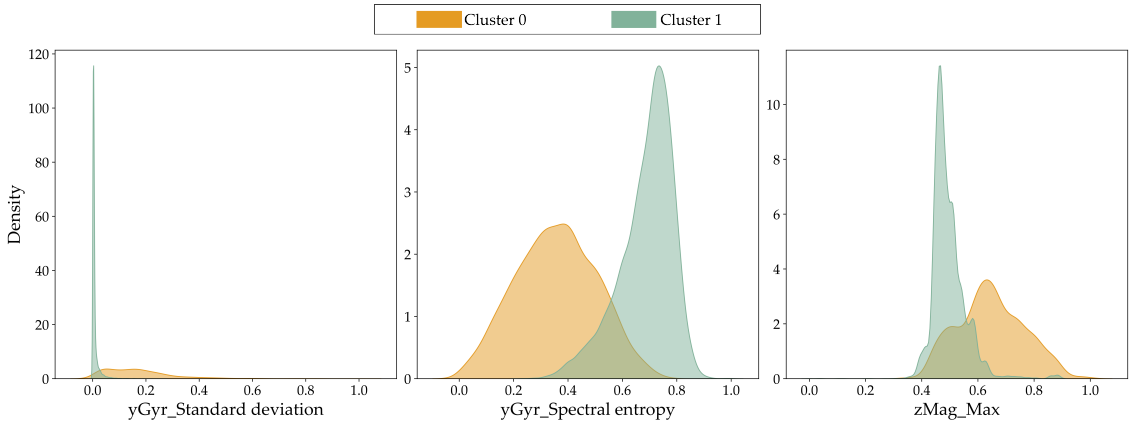


Figure 5.7: KDE of the most relevant features in the decision tree.

most represented across the three clusters. Although cluster 1 contains the largest number of sitting instances, it cannot be determined which cluster belongs to each activity, since most walking and standing instances do not fall into their separate clusters either. For subject 2, the confusion matrix with three clusters is presented in Figure 5.6b. While each class is primarily associated with a distinct cluster, namely walking in cluster 2, sitting in cluster 1, and standing in cluster 0, there is still considerable overlap between clusters, meaning that the instances are being assigned to the wrong cluster. Despite this, the ARI for subject 2 is 51.9 %, considerably higher than for subject 1. The confusion matrices with two clusters are shown in Figure 5.6c for subject 1 and in Figure 5.6d for subject 2. For both subjects, the majority of standing and walking instances are found in cluster 0, while most sitting instances are in cluster 1. Subject 1 has 1327 sitting, 474 standing, and 654 walking instances in cluster 0, and 12 613 sitting, 154 standing, and 372 walking instances in cluster 1. Subject 2 has 432 sitting, 760 standing, and 693 walking instances in cluster 0, and 13 061 sitting, 634 standing, and 19 walking instances in cluster 1. With two clusters the ARI improves for both subjects, reaching 41.2 % for subject 1 and 63.0 % for subject 2.

To get an insight into which features are relevant for separation into the two clusters, a decision tree was trained using the obtained cluster labels (cluster 0 and cluster 1) as target class labels. This is done with subject 2 as this subject achieved a higher ARI. The maximum depth of the tree is set to three and from the grid search, an accuracy score of 97.3 % is obtained with a minimum number of samples per leaf of 20 and a minimum number of samples to split a node of ten. The decision tree is shown up to a depth of two in Annex IV. Three features appear in the decision tree: the SD of the GYR y-axis, the spectral entropy of the GYR y-axis, and the maximum of the MAG z-axis. The KDE of these features are shown in Figure 5.7. The KDEs in cluster 0 show a broader range of values than in cluster 1 for the three features. Additionally, the shapes of the distributions differ in the two clusters.

6

Discussion

In this section, the results presented in Chapter 5 are discussed. Based on the discussion, the limitations of the proposed HAR system are elaborated.

6.1 Experiment 1 - Feature Selection

The feature selection results for the SMMs presented in Table 5.1 show that, for basic activities, M_P and M_{P+W} require fewer features than M_W. This can be explained by the fact that the smartwatch, worn on the wrist, has greater freedom of movement, allowing for more variability compared to the more stable, chest-mounted smartphone. As a result, more features are needed to characterize the smartwatch's movements. When considering all sub-activities, the number of features increases considerably to account for the added complexity and variability. As shown in Annex II and mentioned in Section 5.1.1, for the three sensor modalities and clustering models, and for both basic and all activities, the feature sets have at least one MAG feature, such as the minimum, maximum, or mean. This highlights the essential role of the MAG sensor in these models. MAG readings tend to be higher in magnitude near electromagnetic sources like computers, monitors, and routers. Consequently, during sitting activities, the MAG typically records higher absolute values compared to other activities, such as standing, performed farther from these sources. However, some limitations were observed in the data from subjects P001 and P016. For subject P001, the standing still was recorded outside, thus the shielding effects of the building's ferrous materials on the Earth's magnetic field [72] were not present, resulting in higher readings. A similar effect was observed for P016, as the subject was standing still while on their smartphone.

Regarding the clustering results of the SMMs for basic activities, shown in Table 5.2, these achieve high mean ARI across all three sensor modalities and clustering models, with the highest being AGG with M_{P+W}. These results are consistent with those in [45] and [46], where clustering basic activities also achieved high accuracy. In this scenario, M_W performs the worst and requires more features, whereas M_P performs similarly to M_{P+W} with a similar number of features. Therefore, it can be concluded that the smartphone

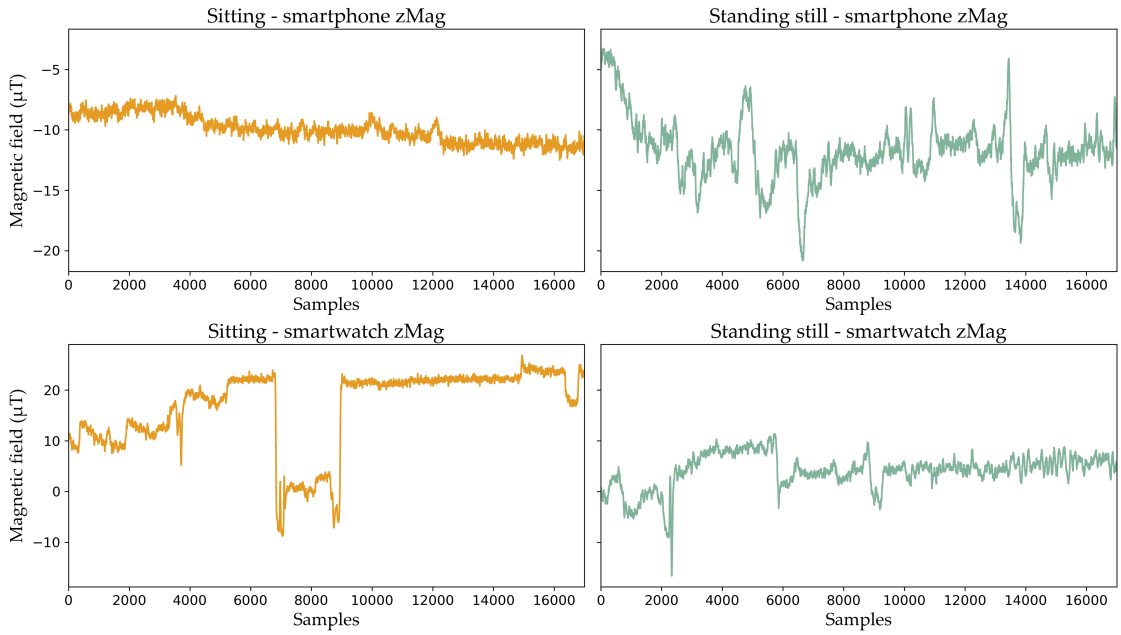


Figure 6.1: Raw segments from the z-axis magnetometer signals, sitting and standing still activities, from both the smartphone and smartwatch for subject P004.

sensors provide more relevant information than smartwatch sensors, likely due to the higher variability of smartwatch data, as mentioned above.

Although the smartphone generally provides more useful information, for two subjects (P004 and P016), with basic activities, M_P performs worse than M_W for the three clustering models. In order to understand this confusion, it is worth analysing the activities being recognized as well as the sensor placement. Sitting and standing still are relatively idle activities compared to walking, which has a distinct, repetitive pattern easily detected with just an ACC, as shown in [45]. When the smartphone is placed on the chest, the ACC readings for sitting and standing still are similar and show low variance, making it difficult to distinguish between the two. The GYR may capture some differences since standing requires constant balance adjustments, resulting in slight variations in angular velocity, unlike sitting with the back supported by a chair. However, as explained earlier, the MAG sensor is likely the most effective at distinguishing these two activities. Figure 6.1 illustrates the z-axis MAG readings for the sitting and standing still tasks from both the smartphone and smartwatch for subject P004. The MAG signals from the smartphone oscillate around similar values for the sitting and standing still. This similarity adds confusion if one would rely solely on the smartphone's MAG sensor. However, the smartwatch signals not only have a different range of values between sitting and standing still, but also exhibit higher magnetic field readings in the sitting activity compared to the smartphone's MAG. This could be explained by the subject sitting at a desk with the computer farther away, while their hands remain static and closer to the magnetic source (e.g., during typing, reading, or writing). Nevertheless, in most cases, if the subject is sitting with the computer closer

to the chest, the smartphone's MAG is enough to capture the different readings between sitting and standing still, as shown in Figure 6.2 with the signals from subject P003.

The models' performance drop considerably with all sub-activities, which can be attributed to the increased complexity of the input data. M_{P+W} performs the best, similar to basic activities, and GMM achieves the highest ARI. While KMeans assumes circular clusters, AGG and GMM allow for more complex geometric shapes. AGG does not assume specific cluster shapes, and GMM can accommodate oval-shaped clusters. Given the better performance of AGG and GMM for basic and all activities, respectively, it can be assumed that the activities do not necessarily form circular clusters. M_W performs worse than the other two modalities and experienced the most significant drop in ARI from basic to all activities. The addition of activities that allow subjects to express higher variability, such as standing while conversing, moving objects inside a cabinet, and making/drinking coffee, can introduce confusion into the models, specially when using smartwatch sensors. The arms have more freedom of movement, and upper body movement does not necessarily correlate to the activities being performed. This leads to higher variance in smartwatch data compared to smartphone data, resulting in a more dispersed feature space, which makes identifying geometric shapes more difficult. The increase in SDs from basic to all activities further highlights the differences between subjects, with some individuals displaying more active behaviours than others while performing the same activity. This variability makes clustering certain subjects more challenging compared to others.

The feature selection results for the 1GM are shown in Table 5.3. The ARI for basic activities is considerably lower than the SMMs. This indicates that a general model that combines data of multiple subjects is not able to successfully cluster the data. This can be attributed to inter-subject variability, meaning that the difference between subjects with respect to execution of the examined activities is too large. This holds true for both basic and all activities. When considering a basic activity like standing still, people already express different behaviors. Some people are more static when standing while others are more likely to move their arms and upper body. Similar behavioral differences can be observed for executing computer work while sitting. This overall variability leads to an overlap of the clusters in the feature space. These differences become even more

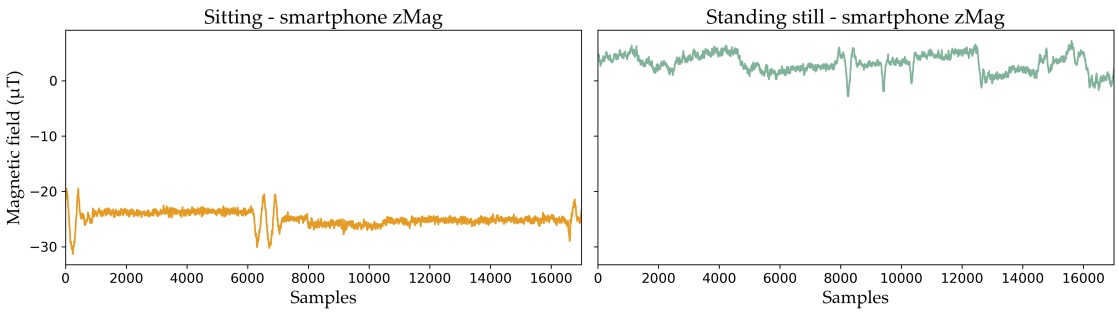


Figure 6.2: Raw segments from the z-axis magnetometer signals, sitting and standing still activities, from the smartphone subject P003.

pronounced when considering all activities, explaining the further drop in performance of the models.

For the 2GM, the number of features and clustering results for the three sensor modalities are shown in Figure 5.1 for basic activities, and in Figure 5.2 for all activities. Overall, with basic activities, these results are similar to the SMMs. However, the high SDs indicate that some subjects can be clustered accurately when a model is finetuned for them, but if their way of performing activities differs significantly from others, the common feature set is not representative. On the other hand, subjects who perform activities in a more static and consistent manner are easier to cluster with a common feature set.

Based on the feature selection experiment results, several conclusions can be drawn. Firstly, the proposed clustering models rely heavily on the MAG sensor. Indoor location techniques [73], which map the interior of buildings by identifying specific areas based on magnetic field readings, could be useful to determine a person's location within the building. This information, if provided to the models as an extra feature, could be useful for recognizing office workers' activities (i.e., if in the hallway, the person is probably either standing or walking). Since PrevOccupAI was not primarily designed for HAR, adding an extra IMU on the thigh should be considered to distinguish sitting from standing without relying on the MAG, which can be heavily affected by external factors. Changing the placement of the smartwatch to the thigh could also be considered, if the device is solely being used for movement tracking. Finally, since the handcrafted features are not able to produce accurate results with all sub-activities, the implementation of an autoencoder for unsupervised feature extraction, as done in [56–58], should be explored.

6.2 Experiment 2 - Cluster Stability

The performance of the SMMs, presented in Table 5.5, decrease for both basic and all activities and KMeans now outperforms AGG and GMM, which were the best models for Experiment 1. Additionally, the increase in SD also suggests that more subjects achieve lower scores. Analysing the results from KMeans with M_P, eleven subjects maintained the 100 % ARI from Experiment 1. This indicates that these subjects likely express more static behaviours and lower variability throughout the recording. On the contrary, subjects P004 and P016, achieved 78.3 % and 74.8 % in Experiment 1 while in Experiment 2 achieved 74.9 % and 62.4 %, respectively. This suggests that these subjects show more variability over time (intra-subject variability), as the first 80 % of data is less similar to the last 20 %. Furthermore, external factors could also change over time and impact clustering performance, such as the environment (inside or outside the building) or electromagnetic sources around the person. Even though the recordings were relatively short, this might indicate that, for these subjects, the shapes and position of the clusters might change over time. However, this can only be explored with data recorded over longer periods or across different days.

AGG and GMM experience greater drops in ARI compared to KMeans, especially for all activities. Since with AGG, clusters are formed based on pairwise distances within the original dataset, it is not designed for handling new, unseen data points without recalculating the clusters. Therefore, training with only the remaining 20 % of the data might not be representative enough, especially involving all activities. Regarding the GMM, the 44.4% ARI drop for all activities likely derives from underfitting. This can occur if the model is not able to capture the complexity of the data, possibly due to the incorrect assumption of Gaussian distributions. This leads to poorly fitted Gaussians, resulting in inaccurate predictions of unseen data. A combination of these factors resulted in subject P008 achieving 0 % for basic activities with GMM and M_W. This can be explained by the variability of smartwatch data, as described in Section 6.1, and different behaviours within the same recording. However, since KMeans and AGG obtained higher scores, the main factor is likely the underfitting of the GMM. The ARI metric also contributes to the 0 %, as it adjusts for randomly correct assignments.

The 1GM results, presented in Table 5.6, also show drops in ARI, though less significant since this model already performed poorly with the first 80 % of the data. The SDs are lower compared to the SMMs, indicating that the 1GM performs poorly across all subjects. The results for the 2GM, presented in Section 5.2.3, show similar trends to the SMMs for the three clustering models and sensor modalities, for both basic and all activities. The confusion matrices give additional insight on how well the model separates the clusters. These show that sitting and standing activities are often confused, and walking slow can be mistaken with these two if performed very slowly. This reinforces the need for one extra sensor and more advanced feature extraction techniques, as stated in Section 6.1.

As mentioned in Section 4.5.2, the data used in Experiment 2 is from the same continuous recording. The obtained results indicate that collecting data to form clusters, and subsequently using these to assign new data points to the established clusters, could lead to poor performance, especially when the time between the training data and the testing data increases (i.e., data from a different day).

The results of the RF, presented in Table 5.8, show that it outperforms the clustering models, similar to findings in [57, 58]. This can be explained by the fact that the RF is composed of decision trees, which are capable of finding non-linear relationships between features [74]. Furthermore, supervised models have the objective to learn the direct relationship between the data and the corresponding label [26], while unsupervised models do not optimize for that objective. For both unsupervised and supervised methods, the best models are the SMMs, since each subject has a feature set which was optimized for them. However, a limitation of this model is that, in a fully unsupervised scenario, an optimized feature set can not be easily obtained as there are no labels for evaluation. The 1GM is also not viable, as it was the worst-performing model for both the clustering models and the RF. Considering the 2GM results, it seems there is a common feature set that could be applied for all subjects. However, to apply this model for all activities, it is worth exploring more robust feature extraction techniques, as stated in Section 6.1.

6.3 Experiment 3 - Cluster Imbalance

The results for Experiment 3 are shown in Figure 5.5. Generally, the datasets with imbalanced classes perform worse than the balanced datasets, with performance decreasing as the imbalance increases. However, there are notable exceptions. AGG shows an improvement in ARI when the dataset consists of 50 % sitting class, while KMeans and GMM improve from 50 % to 70 % sitting class. This can be attributed to the fact that, to achieve the remaining proportions for the standing still and walking medium, majority of these instances are removed. As described in Section 4.5.3, this was done in subsets, meaning that different portions of the standing still and walking medium clusters were tested. If some of these instances were originally (with a balanced dataset) overlapping with the sitting cluster, by removing them, it can actually enhance clustering results. Nevertheless, this improvement is limited. With increasing imbalance, the sitting cluster becomes dominant in terms of size and amount of data points. When imbalance reaches 90 % sitting class, some underrepresented instances, if close to it, may be incorrectly merged with the larger sitting cluster.

This experiment shows that class imbalance has a major impact on clustering performance. Thus, when designing unsupervised HAR systems, class imbalance has to be considered and models that robustly handle these imbalances have to be explored. Equilibrium KMeans is an example of a model designed to handle imbalanced data [75]. This adaptation of KMeans introduces a mechanism that repels centroids, with larger clusters experiencing stronger repulsion. This approach overcomes the "uniform effect" of traditional KMeans, which tends to form clusters of similar sizes, even when the input data contains groups of varying sizes.

6.4 Evaluation on a real-word labeled dataset

The 2GM with M_P, which was the best general model and modality for both basic and all activities, as shown in Table 5.8, was used to validate the proposed framework on actual office work data from two subjects. A silhouette coefficient analysis was performed on this data, which showed higher scores with KMeans and two clusters. As mentioned in Section 4.6, clustering was initially performed with three clusters, followed by two clusters based on the results of the silhouette analysis.

The confusion matrices and correspondent ARI with three clusters for subject 1 and 2 are shown in Figure 5.6a and 5.6b, respectively. The performance of the 2GM with KMeans and M_P dropped considerably when comparing the results of Experiment 1 and Experiment 2. Several factors can contribute to these results. In these recordings, subjects were free to perform their usual work activities without any restrictions on how these should be carried out. Additionally, the long duration of the recordings (over six hours) allowed subjects to express more variability when compared to the shorter acquisitions from the HAR dataset used for model training. Since both subjects spent roughly 90 % of

their time sitting, the class imbalance, as seen in Section 6.3, can also impact clustering performance. The results are particularly poor for subject 1, likely due to the spread of sitting instances over the three clusters, which was not as significant for subject 2. This can indicate that subject 1 expresses more varying behaviours while seated. Nevertheless, these results, along with the silhouette score being higher with two clusters, indicate that the clustering models are not able to accurately distinguish between sitting, standing, and walking in this data.

For two clusters, the confusion matrices and ARI for subject 1 and 2 are shown in Figure 5.6c and 5.6d, respectively. When clustering the data into two clusters, for both subjects, most sitting instances are in cluster 1, which is the dominant cluster, considering the roughly 90 % of sitting class. Although cluster 1 also has standing and walking instances, most of these are present in cluster 0. This suggests that cluster 1 represents a non-active cluster. On the contrary, cluster 0 is comprised mostly of walking and standing instances indicating it forms a cluster containing active activities. Some sitting instances appear in the active cluster, likely since subjects may engage in more dynamic behaviors while seated, such as moving around in the chair, frequently adjusting their posture, or conversing with others. A similar rationale can be applied to the standing activity. For walking, instances may be placed in the non-active cluster if the subject walks very slowly, as seen in Section 6.2.

To interpret the clustering results and to study feature importance, a decision tree, shown in Annex IV, was trained with the two cluster labels (cluster 0 and cluster 1) as target class labels. By analysis of the features present in the tree, these further indicate that MAG features are the most relevant for distinguishing the sitting activity, while GYR features are sufficient for identifying more active activities such as walking. The ACC features, as these do not appear in the first splits of the tree, can be considered less relevant, likely due to the sensor's placement on the chest, as discussed in Section 6.1. The KDE plots, presented in Figure 5.7, show considerable differences between the distributions of the features in the active and non-active clusters. Namely, in the active cluster, the distributions are much wider and spread out than in the non-active cluster, which can be explained by the wider range of motion of active activities. Additionally, it can be seen that some distributions, specially in the active cluster, are not Gaussian, which aligns with the assumptions made for the GMM in Section 6.2.

Although the proposed models cannot accurately distinguish between sitting, standing, and walking, these can differentiate between active and non-active behaviors. This could be useful to detect prolonged static behaviors, which are linked to Work-related musculoskeletal disorders (WRMDs). To further improve this distinction, it is worth exploring contrastive learning schemes for unsupervised feature extraction [76]. These methods differentiate between similar and dissimilar points by analysing their representations within a learned embedding space.

7

Conclusion and Outlook

The objective of this thesis was to implement unsupervised learning models for Human activity recognition (HAR) in office environments, with the purpose of utilizing them for identification of work-related musculoskeletal disorders (WRMDs) risk factors in the future. This work proposed a new HAR dataset comprising nine office activities acquired from 19 subjects using a smartphone and a smartwatch, which was used for model training. KMeans, Agglomerative Clustering (AGG), and Gaussian Mixture Model (GMM) were employed in three different approaches. Subject specific models (SSMs), used to evaluate how unsupervised learning schemes with optimized feature sets perform for each subject. The one-stage general model (1GM), created to study if there is similarity in subjects' data that allows them to be clustered as a group. Lastly, the two-stage general model (2GM), implemented for assessing if there is a common feature set for HAR in office environments. The impact of the complexity of the input data on the models' performance was studied by considering only basic activities (sitting, standing still, and walking medium speed) and all nine activities. Three experiments were conducted with this HAR system.

Experiment 1 consisted of feature selection. Less features were required to cluster the basic activities than all activities. The feature sets obtained with modality smartwatch (M_W) were bigger than modality smartphone (M_P) and modality smartphone and smartwatch (M_{P+W}). Magnetometer (MAG) features are included in all sets and are the most effective at distinguishing sitting from standing, due to higher magnetic field readings in front of computers. For basic activities, the SSMs achieved over 99 % mean Adjusted Rand Index (ARI) while the 2GM exceeded 90 %. This occurred for all three clustering models with M_{P+W}. The lowest scores for basic activities were obtained with the 1GM reaching 77 % ARI with KMeans and M_{P+W}. Adding all activities impacts clustering performance with the SSMs decreasing to 88.6 % ARI (GMM with M_{P+W}), the 2GM to 72.9 % (AGG with M_P), and the 1GM to 56.2 % (AGG with M_P). M_W produced the worst scores and required more features due to the greater freedom of arm movement, leading to higher variability and thus overlapping clusters in the feature space. Although M_{P+W} achieved overall the highest results, the minimal difference compared to M_P led to the latter being considered the best modality. Cluster stability was studied in Experiment 2 by constructing the

clusters on the initial 80 % of the data and predicting the remaining 20 %. The variability of subjects' behaviours over time (intra-subject variability) impacted clustering results for the three clustering models and sensor modalities, specially with all activities. The best clustering models were compared to a Random Forest (RF), trained with the same feature sets, which obtained higher accuracy. Lastly, in Experiment 3, the datasets containing basic activities and M_P were imbalanced to compare the models' performances when sitting class was overrepresented. It was concluded that cluster imbalance has a major impact on clustering performance.

The proposed models were tested on data from two subjects performing actual work tasks during a workday. Subjects spent roughly 90 % of their time sitting which, similar to Experiment 3, impacted clustering performance. The clustering models were not able to accurately distinguish between sitting, standing, and walking. However, clustering the data into two clusters produced higher ARI and silhouette coefficient. In this scenario, a non-active cluster contained most of the sitting instances, while an active cluster contained mainly standing and walking instances. Recognizing active and non-active behaviours can give insights into potential risk factors for WRMDs. Considering the results mentioned above, the objectives of this thesis, proposed in Section 1.2, were achieved to the following extent:

- **(O1):** A diverse dataset was collected for HAR in office environments.
- **(O2):** KMeans, AGG, and GMM were tested on three different experiments.
- **(O3):** The feature sets found allow the models to accurately cluster basic activities, with performance drops being observed for all activities.
- **(O4):** A subject-independent model showed low performance as the difference between subjects with respect to execution of the activities is too high.
- **(O5):** The stability of clusters is affected by intra-subject variability over time and data imbalance impacts clustering performance.
- **(O6):** When applying the models to real-world data, one active and one non-active cluster yielded better results.

Several improvements should be considered. Firstly, indoor location techniques could be employed to add context regarding where the person is inside the office building. This information could be added as an extra feature to the models. Regarding the sensor setup, it is not optimal for distinguishing sitting from standing. Therefore, adding an extra sensor strapped to the thigh could reduce the models' reliance on the MAG sensor, which is often affected by external factors. The handcrafted features are not representative enough to cluster all activities accurately. Thus, unsupervised feature extraction using autoencoders or contrastive learning should be considered. Lastly, models that can robustly handle data imbalances should be explored for application in real-world office environments.

Bibliography

[1] J. M. Lourenço. *The NOVAthesis LaTeX Template User's Manual*. NOVA University Lisbon. 2021. URL: https://github.com/joaomlourenco/novathesis/raw/main/template.pdf (cit. on p. i).

[2] European Agency for Safety and Health at Work et al. *Work-related musculoskeletal disorders – Prevalence, costs and demographics in the EU*. Publications Office, 2019. DOI: 10.2802/66947 (cit. on p. 1).

[3] H. Zerguine et al. "Online office ergonomics training programs: A scoping review examining design and user-related outcomes". In: *Safety Science* 158 (2023). ISSN: 0925-7535. DOI: https://doi.org/10.1016/j.ssci.2022.106000 (cit. on p. 1).

[4] European Agency for Safety and Health at Work et al. *Estimating the cost of work-related accidents and ill-health – An analysis of European data sources*. Publications Office, 2017. DOI: 10.2802/566789 (cit. on p. 1).

[5] D. Srinivasan and S. E. Mathiassen. "Motor variability - An important issue in occupational life". In: *Work*. Vol. 41. 2012, pp. 2527–2534. DOI: 10.3233/WOR-2012-0493-2527 (cit. on p. 1).

[6] N. Owen et al. "Sedentary Behavior and Public Health: Integrating the Evidence and Identifying Potential Solutions". In: *Annual Review of Public Health* 41 (2020), pp. 265–287. ISSN: 1545-2093. DOI: https://doi.org/10.1146/annurev-publhealth-040119-094201 (cit. on p. 1).

[7] A. B. Bakker and J. D. de Vries. "Job Demands–Resources theory and self-regulation: new explanations and remedies for job burnout". In: *Anxiety, Stress, & Coping* 34.1 (2021), pp. 1–21. DOI: 10.1080/10615806.2020.1797695 (cit. on p. 1).

[8] L. Punnett and D. H. Wegman. "Work-related musculoskeletal disorders: the epidemiologic evidence and the debate". In: *Journal of Electromyography and Kinesiology* 14.1 (2004), pp. 13–23. ISSN: 1050-6411. DOI: https://doi.org/10.1016/j.jelekin.2003.09.015 (cit. on p. 1).

[9] European Commission and Directorate-General for Research and Innovation. *Horizon Europe strategic plan 2021-2024*. Publications Office of the European Union, 2021. DOI: 10.2777/083753 (cit. on p. 1).

[10] J. Park et al. "Sedentary Lifestyle: Overview of Updated Evidence of Potential Health Risks". In: *Korean Journal of Family Medicine* 41 (2020-11), pp. 365–373. DOI: 10.4082/kjfm.20.0165 (cit. on p. 1).

[11] E. Oliosi et al. "Week-long Multimodal Data Acquisition of Occupational Risk Factors in Public Administration Workers". In: *2023 19th International Conference on Intelligent Environments, IE 2023 - Proceedings*. IEEE, 2023. DOI: 10.1109/IE57519.2023.10179099 (cit. on pp. 1, 10, 11).

[12] V. Dentamaro et al. "Human activity recognition with smartphone-integrated sensors: A survey". In: *Expert Systems with Applications* 246 (2024). ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2024.123143. URL: https://www.sciencedirect.com/science/article/pii/S0957417424000083 (cit. on pp. 2, 6).

[13] A. O. Ige and M. H. Mohd Noor. "A survey on unsupervised learning for wearable sensor-based activity recognition". In: *Applied Soft Computing* 127 (2022). ISSN: 1568-4946. DOI: https://doi.org/10.1016/j.asoc.2022.109363 (cit. on pp. 2, 6, 7).

[14] M. S. Ryoo et al. "Privacy-preserving human activity recognition from extreme low resolution". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 31. 1. 2017, pp. 4255–4262. DOI: https://doi.org/10.1609/aaai.v31i1.11233 (cit. on p. 2).

[15] N. Gupta et al. "Human activity recognition in artificial intelligence framework: a narrative review". In: *Artificial intelligence review* 55 (2022), pp. 4755–4808. ISSN: 0269-2821. DOI: 10.1007/s10462-021-10116-x (cit. on p. 3).

[16] S. Seenath and M. Dharmaraj. "Conformer-Based Human Activity Recognition Using Inertial Measurement Units". In: *Sensors* 23.17 (2023). ISSN: 1424-8220. DOI: 10.3390/s23177357 (cit. on p. 3).

[17] D.-A. Nguyen and N.-A. Le-Khac. "Sok: Behind the accuracy of complex human activity recognition using deep learning". In: *2024 International Joint Conference on Neural Networks (IJCNN)*. IEEE. 2024, pp. 1–8. DOI: 10.1109/IJCNN60899.2024.10650322 (cit. on pp. 3, 6).

[18] C.-C. Yang and Y.-L. Hsu. "A Review of Accelerometry-Based Wearable Motion Detectors for Physical Activity Monitoring". In: *Sensors* 10.8 (2010), pp. 7772–7788. ISSN: 1424-8220. DOI: 10.3390/s100807772 (cit. on p. 3).

[19] T. Seel, M. Kok, and R. S. McGinnis. "Inertial Sensors—Applications and Challenges in a Nutshell". In: *Sensors* 20.21 (2020). ISSN: 1424-8220. DOI: 10.3390/s20216221 (cit. on p. 3).

[20] M. Grossi. "A sensor-centric survey on the development of smartphone measurement and sensing systems". In: *Measurement* 135 (2019), pp. 572–592. ISSN: 0263-2241. DOI: https://doi.org/10.1016/j.measurement.2018.12.014 (cit. on p. 3).

[21] K. Chen and D. Bassett. "The Technology of Accelerometry-Based Activity Monitors: Current and Future". In: *Medicine and science in sports and exercise* 37 (2005-12), S490–500. DOI: 10.1249/01.mss.0000185571.49104.82 (cit. on p. 3).

[22] V. M. N. Passaro et al. "Gyroscope Technology and Applications: A Review in the Industrial Perspective". In: *Sensors* 17.10 (2017). ISSN: 1424-8220. DOI: 10.3390/s17102284 (cit. on p. 3).

[23] S. Stančin and S. Tomažič. "On the Interpretation of 3D Gyroscope Measurements". In: *Journal of Sensors* 2018.1 (2018). DOI: https://doi.org/10.1155/2018/9684326 (cit. on p. 3).

[24] G. Ouyang and K. Abed-Meraim. "A Survey of Magnetic-Field-Based Indoor Localization". In: *Electronics* 11.6 (2022). ISSN: 2079-9292. DOI: 10.3390/electronics11060864 (cit. on p. 3).

[25] P. Zhu et al. "Improving Indoor Pedestrian Dead Reckoning for Smartphones under Magnetic Interference Using Deep Learning". In: *Sensors* 23.23 (2023). ISSN: 1424-8220. DOI: 10.3390/s23239348 (cit. on p. 3).

[26] R. Choi et al. "Introduction to Machine Learning, Neural Networks, and Deep Learning". In: *Translational vision science technology* 9 (2020), p. 14. URL: https://tvst.arvojournals.org/article.aspx?articleid=2762344 (cit. on pp. 4, 36).

[27] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. URL: http://www.deeplearningbook.org (cit. on p. 4).

[28] S. Badillo et al. "An Introduction to Machine Learning". In: *Clinical Pharmacology & Therapeutics* 107.4 (2020), pp. 871–885. DOI: https://doi.org/10.1002/cpt.1796 (cit. on pp. 4, 5).

[29] S. Aghabozorgi, A. Seyed Shirkhorshidi, and T. Ying Wah. "Time-series clustering – A decade review". In: *Information Systems* 53 (2015), pp. 16–38. ISSN: 0306-4379. DOI: https://doi.org/10.1016/j.is.2015.04.007 (cit. on pp. 4, 5).

[30] L. Kaufman and P. Rousseeuw. *Finding Groups in Data: An Introduction To Cluster Analysis*. 1990-01. ISBN: 0-471-87876-6. DOI: 10.2307/2532178 (cit. on p. 4).

[31] scikit-learn. *KMeans*. 2024. URL: https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html#sklearn.cluster.KMeans.predict (cit. on p. 4).

[32] D. Müllner. "Modern hierarchical, agglomerative clustering algorithms". In: *ArXiv* (2011). DOI: https://doi.org/10.48550/arXiv.1109.2378 (cit. on p. 4).

[33] C. Biernacki, G. Celeux, and G. Govaert. "Assessing a mixture model for clustering with the integrated completed likelihood". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.7 (2000), pp. 719–725. DOI: 10.1109/34.865189 (cit. on p. 5).

[34] J.-O. Palacio-Niño and F. B. Galiano. "Evaluation Metrics for Unsupervised Learning Algorithms". In: *ArXiv* (2019). DOI: https://doi.org/10.48550/arXiv.1905.05667 (cit. on p. 5).

[35] M. Warrens and H. van der Hoef. "Understanding the Adjusted Rand Index and Other Partition Comparison Indices Based on Counting Object Pairs". In: *Journal of Classification* 39 (2022), pp. 487–509. ISSN: 0176-4268. DOI: 10.1007/s00357-022-09413-z (cit. on p. 5).

[36] scikit-learn. *rand_score*. 2024. URL: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.rand_score.html#rbedd61930922-1 (cit. on p. 5).

[37] scikit-learn. *adjusted_rand_score*. 2024. URL: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.adjusted_rand_score.html (cit. on p. 5).

[38] scikit-learn. *silhouette_score*. 2024. URL: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.silhouette_score.html (cit. on p. 5).

[39] C. Jobanputra, J. Bavishi, and N. Doshi. "Human Activity Recognition: A Survey". In: *Procedia Computer Science* 155 (2019), pp. 698–703. ISSN: 1877-0509. DOI: 10.1016/j.procs.2019.08.100 (cit. on p. 6).

[40] O. C. Ann and L. B. Theng. "Human activity recognition: A review". In: *2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014)*. 2014, pp. 389–393. DOI: 10.1109/ICCSCE.2014.7072750 (cit. on p. 6).

[41] O. D. Lara and M. A. Labrador. "A Survey on Human Activity Recognition using Wearable Sensors". In: *IEEE Communications Surveys Tutorials* 15.3 (2013), pp. 1192–1209. DOI: 10.1109/SURV.2012.110112.00192 (cit. on p. 6).

[42] P. Ariza Colpas et al. "Unsupervised Human Activity Recognition Using the Clustering Approach: A Review". In: *Sensors* 20.9 (2020). ISSN: 1424-8220. DOI: https://doi.org/10.3390/s20092702 (cit. on p. 6).

[43] A. Wang et al. "HierHAR: Sensor-Based Data-Driven Hierarchical Human Activity Recognition". In: *IEEE Sensors Journal* 21.3 (2021), pp. 3353–3365. DOI: 10.1109/JSEN.2020.3023860 (cit. on p. 6).

[44] A. Omolaja, A. Otebolaku, and A. Alfoudi. "Context-Aware Complex Human Activity Recognition Using Hybrid Deep Learning Models". In: *Applied Sciences* 12.18 (2022). ISSN: 2076-3417. DOI: 10.3390/app12189305 (cit. on p. 6).

[45] I. Machado. et al. "Human Activity Recognition from Triaxial Accelerometer Data - Feature Extraction and Selection Methods for Clustering of Physical Activities". In: *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing (BIOSTEC 2014) - BIOSIGNALS*. INSTICC. SciTePress, 2014, pp. 155–162. ISBN: 978-989-758-011-6. DOI: 10.5220/0004749801550162 (cit. on pp. 6, 7, 16, 32, 33).

[46] Y. Kwon, K. Kang, and C. Bae. "Unsupervised learning for human activity recognition using smartphone sensors". In: *Expert Systems with Applications* 41.14 (2014), pp. 6067–6074. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2014.04.037 (cit. on pp. 6, 7, 16, 32).

[47] D. Anguita et al. "A Public Domain Dataset for Human Activity Recognition using Smartphones". In: *The European Symposium on Artificial Neural Networks*. 2013. URL: https://api.semanticscholar.org/CorpusID:6975432 (cit. on p. 6).

[48] A. Stisen et al. "Smart Devices are Different: Assessing and MitigatingMobile Sensing Heterogeneities for Activity Recognition". In: *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*. Association for Computing Machinery, 2015, pp. 127–140. ISBN: 9781450336314. DOI: 10.1145/2809695.2809718 (cit. on p. 6).

[49] O. Banos et al. "mHealthDroid: A Novel Framework for Agile Development of Mobile Health Applications". In: *Ambient Assisted Living and Daily Activities*. Springer International Publishing, 2014, pp. 91–98. DOI: https://doi.org/10.1007/978-3-319-13105-4_14 (cit. on pp. 6, 7).

[50] T. Stiefmeier et al. "Wearable Activity Tracking in Car Manufacturing". In: *IEEE Pervasive Computing* 7.2 (2008), pp. 42–50. DOI: 10.1109/MPRV.2008.40 (cit. on p. 7).

[51] A. Reiss and D. Stricker. "Introducing a New Benchmarked Dataset for Activity Monitoring". In: *2012 16th International Symposium on Wearable Computers*. 2012, pp. 108–109. DOI: 10.1109/ISWC.2012.13 (cit. on p. 7).

[52] O. Baños et al. "A benchmark dataset to evaluate sensor displacement in activity recognition". In: *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. Association for Computing Machinery, 2012, pp. 1026–1035. ISBN: 9781450312240. DOI: 10.1145/2370216.2370437 (cit. on p. 7).

[53] L. F. Mejia-Ricart, P. Helling, and A. Olmsted. "Evaluate action primitives for human activity recognition using unsupervised learning approach". In: *2017 12th International Conference for Internet Technology and Secured Transactions (ICITST)*. 2017, pp. 186–188. DOI: 10.23919/ICITST.2017.8356374 (cit. on pp. 7, 16).

[54] D. Trabelsi et al. "An Unsupervised Approach for Automatic Activity Recognition Based on Hidden Markov Model Regression". In: *IEEE Transactions on Automation Science and Engineering* 10.3 (2013), pp. 829–835. DOI: 10.1109/TASE.2013.2256349 (cit. on pp. 7, 8).

[55] L. Bai et al. "Motion2Vector: unsupervised learning in human activity recognition using wrist-sensing data". In: *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. Association for Computing Machinery, 2019, pp. 537–542. ISBN: 9781450368698. DOI: 10.1145/3341162.3349335 (cit. on p. 8).

[56] A. Abedin et al. "Towards deep clustering of human activities from wearables". In: *ISWC '20:Proceedings of the 2020 ACM International Symposium on Wearable Computers*. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–6. ISBN: 9781450380775. DOI: 10.1145/3410531.3414312 (cit. on pp. 8, 35).

[57] T. Sheng and M. Huber. "Unsupervised Embedding Learning for Human Activity Recognition Using Wearable Sensor Data". In: *The Thirty-Third International FLAIRS Conference (FLAIRS-33)*. 2023. DOI: 10.48550/arXiv.2307.11796 (cit. on pp. 8, 16, 35, 36).

[58] H. Ma et al. "Unsupervised Human Activity Representation Learning with Multi-task Deep Clustering". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 5.1 (2021). DOI: 10.1145/3448074 (cit. on pp. 8, 35, 36).

[59] G. Van Rossum and F. L. Drake. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009. ISBN: 1441412697 (cit. on p. 9).

[60] C. R. Harris et al. "Array programming with NumPy". In: *Nature* 585.7825 (2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2 (cit. on p. 9).

[61] W. McKinney. "Data Structures for Statistical Computing in Python". In: *Proceedings of the 9th Python in Science Conference*. Ed. by S. van der Walt and J. Millman. 2010, pp. 56–61. DOI: 10.25080/Majora-92bf1922-00a (cit. on p. 9).

[62] P. Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2 (cit. on p. 9).

[63] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830 (cit. on p. 9).

[64] M. Barandas et al. "TSFEL: Time Series Feature Extraction Library". In: *SoftwareX* 11 (2020). ISSN: 2352-7110. DOI: https://doi.org/10.1016/j.softx.2020.100456 (cit. on p. 9).

[65] P. W. Biosignals. *Synchronising data from multiple Android sensor files into one file*. 2024. URL: http://notebooks.pluxbiosignals.com/notebooks/Categories/Other/android_file_sync_rev.html (cit. on p. 9).

[66] P. W. Biosignals. *Device Synchronisation - Cable, Light and Sound Approaches*. 2024. URL: http://notebooks.pluxbiosignals.com/notebooks/Categories/Pre-Process/synchronisation_rev.html (cit. on p. 9).

[67] N. A. Stanton. "Hierarchical task analysis: Developments, applications, and extensions". In: *Applied Ergonomics* 37.1 (2006), pp. 55–79. ISSN: 0003-6870. DOI: https://doi.org/10.1016/j.apergo.2005.06.003 (cit. on p. 9).

[68] S. Silva. et al. "Assessing Occupational Health with a Cross-platform Application based on Self-reports and Biosignals". In: *Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2022) - HEALTHINF*. INSTICC. SciTePress, 2022, pp. 549–556. ISBN: 978-989-758-552-4. DOI: 10.5220/0010846700003123 (cit. on p. 10).

[69] H. J. Hermens et al. "Development of recommendations for SEMG sensors and sensor placement procedures". In: *Journal of Electromyography and Kinesiology* 10.5 (2000), pp. 361–374. ISSN: 1050-6411. DOI: https://doi.org/10.1016/S1050-6411(00)00027-4 (cit. on p. 11).

[70] D. Anguita et al. "A Public Domain Dataset for Human Activity Recognition Using Smartphones". In: *ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium)*. i6doc.com, 2013. ISBN: 9782874190810 (cit. on p. 16).

[71] H. Yang, L. Jiao, and Q. Pan. "A Survey on Interpretable Clustering". In: *Chinese Control Conference, CCC*. IEEE, 2021-07, pp. 7384–7388. DOI: 10.23919/CCC52363.2021.9549986 (cit. on p. 21).

[72] V. Gerginov, F. C. S. da Silva, and D. Howe. "Prospects for magnetic field communications and location using quantum sensors". In: *Review of Scientific Instruments* 88.12 (2017). ISSN: 0034-6748. DOI: 10.1063/1.5003821 (cit. on p. 32).

[73] D. Moreira et al. "Human Activity Recognition for Indoor Localization Using Smartphone Inertial Sensors". In: *Sensors* 21.18 (2021). ISSN: 1424-8220. DOI: 10.3390/s21186316 (cit. on p. 35).

[74] L. Auret and C. Aldrich. "Interpretation of nonlinear relationships between process variables by use of random forests". In: *Minerals Engineering* 35 (2012), pp. 27–42. ISSN: 0892-6875. DOI: https://doi.org/10.1016/j.mineng.2012.05.008 (cit. on p. 36).

[75] Y. He. "Imbalanced Data Clustering using Equilibrium K-Means". In: *arXiv preprint arXiv:2402.14490* (2024). DOI: https://doi.org/10.48550/arXiv.2402.14490 (cit. on p. 37).

[76] K. Takenaka, K. Kondo, and T. Hasegawa. "Segment-Based Unsupervised Learning Method in Sensor-Based Human Activity Recognition". In: *Sensors* 23.20 (2023). DOI: 10.3390/s23208449 (cit. on p. 38).

I

Consent Form

Consentimento informado, esclarecido e livre para participação em estudos de investigação

Por favor, leia com atenção a seguinte informação. Se achar que algo está incorreto ou que não está claro, não hesite em solicitar mais informações.

Título do estudo: *Inter- and intrasubject similarity using unsupervised learning schemes for evaluation of working patterns.*

Enquadramento: No âmbito da tese de mestrado *Inter- and intrasubject similarity using unsupervised learning schemes for evaluation of working patterns.*

Explicação do estudo: Este estudo envolve a recolha de dados através de sensores de movimento e posição a partir do *smartwatch*, *smartphone* e de dois dispositivos *muscleBAN*. A partir do smartphone vai ser recolhido o nível de ruído, não sendo detetáveis quaisquer palavras ou sons de fundo que permitam identificar conversas. O sinal de eletromiografia também será recolhido a partir dos dispositivos *muscleBAN*. A aquisição será realizada no laboratório do LIBPhys e no departamento de física da NOVA school of Science and Technology. Serão realizadas as seguintes atividades: trabalhar sentado numa secretária, caminhar, mexer objetos dentro de um armário e permanecer de pé. No total, as acquisições terão a duração de uma hora e cinco minutos. Será realizada uma sessão prévia para apresentação do projeto, explicação do protocolo da recolha de dados e esclarecimento de potenciais dúvidas. É importante mencionar que neste relatório integrará os seus dados pessoais, dos quais é o único proprietário, pelo que não deve ser partilhado com ninguém.

Condições: É importante referir que terá toda a liberdade para recusar a participação no estudo ou retirar o seu consentimento, suspendendo a participação em qualquer momento. Ao longo de todo o processo não terá qualquer custo por participar neste estudo. Os participantes neste estudo comprometem-se a cumprir os requisitos abaixo:

- Jovens adultos com idade superior a 18 anos;
- Sujeitos sem patologia associada (neurológica, ortopédica, reumática, oncológica ou cardiorrespiratória);
- Sujeitos que não tomem drogas psicotrópicas;
- Sujeitos não medicados;
- Sujeitos capazes de utilizar o equipamente necessário.

Confidencialidade e anonimato: Os dados pessoais (i.e., biossinais e dados demográficos) recolhidos serão utilizados apenas no contexto desta investigação, não sendo divulgados em nenhum outro contexto e poderão ser tornados públicos (publicações científicas e conferências).

A qualquer momento, poderá exercer os seus direitos, nomeadamente o direito de solicitar mais informações sobre o tratamento dos seus dados pessoais, o direito de retificação e de apagamento de dados, de remoção do seu consentimento ou de oposição às atividades de tratamento, entre outros, de acordo com o disposto no Regulamento Geral sobre a Proteção de Dados.

Pode exercer os seus direitos contactando Sara Leonor Martins dos Santos, slm.santos@campus.fct.unl.pt. Caso entenda que o tratamento dos seus dados pessoais não cumpre com a legislação aplicável têm, ainda, o direito de apresentar uma

reclamação junto da autoridade supervisora competente. Em Portugal, pode contactar a Comissão Nacional de Proteção de Dados (CNPD) em https://www.cnpd.pt.

Identificação do Investigador que irá recolher os dados:

Nome: Sara Leonor Martins dos Santos

Profissão: Aluna de mestrado em engenharia biomédica

Contato telefónico: 925 138 582

Endereço eletrónico: slm.santos@campus.fct.unl.pt

Investigador

Confirmo que expliquei à pessoa abaixo indicada, de forma adequada e inteligível, os procedimentos necessários ao estudo referido neste documento. Respondi a todas as questões que me foram colocadas e assegurei-me de que houve um período de reflexão suficiente para a tomada da decisão.

(Assinatura Legível) _____

(Sara Leonor Martins dos Santos)

Data: ___/___/_____

Participante

Por favor, leia com atenção todo o conteúdo deste documento. Não hesite em solicitar mais informações se não estiver completamente esclarecido(a). Verifique se todas as informações estão corretas. Se concordar com o procedimento descrito e com todas as informações fornecidas acerca do estudo e pretender continuar, por favor, assine em baixo.

Declaro ter compreendido os objetivos que me foram propostos e explicados. Foi-me concedida a oportunidade de esclarecer todas as dúvidas sobre o assunto e para todas elas obtive uma resposta esclarecedora. Tive tempo suficiente para refletir sobre esta proposta, pelo que declaro que autorizo/não autorizo (riscar o que não interessa) participar no estudo indicado, bem como os procedimentos diretamente relacionados que sejam necessários no meu próprio interesse e justificados por razões fundamentadas. Compreendo que a minha participação no estudo é voluntária e que posso retirar o consentimento em qualquer momento.

Nome: _____

(Assinatura Legível) _____

Data: ___/___/_____

II

Feature Sets

II.1 Feature sets for the subject specific models (SSMs)

II.1.1 Basic activities

Table II.1: SSMs feature sets KMeans, modality smartphone for basic activities

Subject	Feature set
P001	zGyr Standard deviation, zMag Max, yAcc Min, xMag Max
P002	zMag Max, zGyr Median frequency, yAcc Interquartile range, xAcc Min, yGyr Spectral entropy, yMag Spectral centroid, yGyr Median frequency, zGyr Spectral centroid
P003	yMag Max, yAcc Min
P004	xAcc Max, yGyr Interquartile range, zAcc Interquartile range, zGyr Max, zMag Max, yGyr Spectral entropy, zGyr Spectral centroid, yAcc Spectral entropy
P005	xGyr Variance, yAcc Max, zMag Max, yMag Min, yMag Mean
P006	yAcc Max, zMag Max
P007	zMag Max, yAcc Min
P008	zMag Mean, yAcc Interquartile range, yAcc Min
P009	zMag Mean, yAcc Max
P010	xMag Max, yAcc Min, zMag Max, zGyr Standard deviation
P011	yAcc Min, xMag Max
P012	xGyr Variance, xMag Max, xGyr Min, zAcc Max, yAcc Max
P013	xMag Max, zAcc Standard deviation
P014	yAcc Standard deviation, zMag Max
P015	xGyr Standard deviation, xMag Max, yAcc Max

Continued on next page

Table II.1 SSMs feature sets, KMeans modality smartphone for basic activities

Subject	Feature set
P016	zMag Standard deviation, xAcc Min, yAcc Spectral entropy, xMag Spectral centroid, xAcc Median frequency, yAcc Min, yMag Max, zGyr Spectral entropy, yAcc Interquartile range, xMag Mean, zAcc Median frequency, yMag Mean
P018	zMag Min, yGyr Max, zAcc Min, xMag Mean, zAcc Max
P019	xAcc Standard deviation, xMag Max, zAcc Interquartile range
P020	yAcc Min, xMag Max, yAcc Max

Legend: **SSMs**- Subject specific models.

Table II.2: SSMs feature sets Kmeans, modality smartwatch for basic activities

Subject	Feature set
P001	xAcc wear Variance, zAcc wear Median frequency, xMag wear Max, yMag wear Min
P002	xGyr wear Interquartile range, xMag wear Min, yGyr wear Spectral entropy, zGyr wear Variance, xAcc wear Standard deviation, zMag wear Max, yMag wear Mean, zMag wear Spectral entropy, yAcc wear Spectral entropy, xAcc wear Spectral centroid, zAcc wear Median frequency
P003	yGyr wear Variance, zMag wear Max, zMag wear Standard deviation, zMag wear Spectral centroid, xMag wear Max
P004	xMag wear Max, yMag wear Max, yMag wear Variance
P005	yMag wear Max, xAcc wear Min, xMag wear Min
P006	zGyr wear Max, zGyr wear Min, xMag wear Max, yMag wear Spectral centroid
P007	yAcc wear Variance, xGyr wear Standard deviation, yMag wear Min, xAcc wear Min, yMag wear Variance, zGyr wear Variance, yAcc wear Spectral centroid, zAcc wear Interquartile range, yGyr wear Median frequency, zGyr wear Spectral entropy, zMag wear Spectral centroid, yMag wear Max
P008	xMag wear Spectral entropy, yMag wear Max, zGyr wear Spectral entropy, zGyr wear Max, zGyr wear Variance, yGyr wear Standard deviation, yMag wear Spectral centroid, yGyr wear Spectral entropy, xAcc wear Standard deviation, xAcc wear Spectral centroid, zMag wear Min, zMag wear Spectral centroid
P009	xMag wear Max, yMag wear Spectral centroid, zGyr wear Standard deviation, zGyr wear Interquartile range, zMag wear Min

Continued on next page

Table II.2 SSMs feature sets KMeans, modality smartwatch for basic activities

Subject	Feature set
P010	zGyr wear Max, zGyr wear Interquartile range, zMag wear Mean, xMag wear Min
P011	zGyr wear Interquartile range, xMag wear Max
P012	xMag wear Spectral entropy, xGyr wear Spectral entropy, xGyr wear Min, xAcc wear Variance, yMag wear Mean, xMag wear Mean, xMag wear Min, yMag wear Spectral centroid
P013	zGyr wear Interquartile range, zMag wear Max, yMag wear Interquartile range
P014	zMag wear Min, yAcc wear Interquartile range, zGyr wear Standard deviation, zGyr wear Median frequency, xAcc wear Median frequency, xMag wear Mean, zGyr wear Variance, xMag wear Spectral entropy, xGyr wear Interquartile range, xMag wear Min
P015	yAcc wear Standard deviation, xAcc wear Median frequency, yMag wear Median frequency, xMag wear Interquartile range, xMag wear Mean, yGyr wear Max, zGyr wear Interquartile range, yGyr wear Standard deviation, yMag wear Standard deviation, zAcc wear Spectral centroid
P016	zGyr wear Interquartile range, yMag wear Mean, yMag wear Interquartile range, xMag wear Mean, zGyr wear Median
P018	yAcc wear Standard deviation, xMag wear Interquartile range, zGyr wear Interquartile range, yGyr wear Max, xMag wear Mean, xMag wear Spectral entropy, yGyr wear Interquartile range, yGyr wear Standard deviation, yMag wear Mean, xMag wear Max, yMag wear Interquartile range
P019	zGyr wear Variance, yMag wear Mean, zGyr wear Median frequency, yMag wear Interquartile range, zAcc wear Interquartile range, zGyr wear Interquartile range, yGyr wear Interquartile range, xGyr wear Interquartile range
P020	yMag wear Min, xMag wear Spectral centroid, xMag wear Max, yMag wear Interquartile range, xGyr wear Interquartile range, zGyr wear Min, zGyr wear Interquartile range

Legend: **SSMs**- Subject specific models; **wear**- watch.

Table II.3: SSMs feature sets KMeans, modality smartphone and -watch for basic activities

Subject	Feature set
P001	xGyr wear Standard deviation, zMag Max, zGyr wear Interquartile range, yMag wear Max, zAcc Standard deviation, zMag wear Spectral centroid, yMag Max
P002	yMag Spectral centroid, yMag Standard deviation, zMag Median frequency, xAcc wear Standard deviation, xMag wear Min, xGyr Standard deviation, xGyr wear Median frequency, yGyr Median frequency, xMag Median frequency, zMag wear Spectral entropy, yGyr Interquartile range, yAcc wear Median frequency, zMag Max, yGyr wear Spectral centroid, yMag wear Min, xMag wear Mean
P003	xMag wear Max, yGyr wear Standard deviation
P004	yMag wear Min, xMag wear Max
P005	yMag Mean, yAcc Max, yMag Min
P006	zGyr Standard deviation, zMag Max
P007	zMag Max, yAcc Min
P008	zGyr wear Standard deviation, xMag Max, yAcc Max
P009	xGyr Standard deviation, zMag Mean
P010	yMag Mean, xMag Max, xGyr wear Interquartile range, xGyr Min, xMag wear Min, zMag wear Min
P011	yGyr Standard deviation, zMag Max, yMag Min, xMag Max
P012	xAcc Standard deviation, xMag Max, xGyr Min, yAcc Max
P013	yAcc Interquartile range, zMag wear Mean
P014	zMag Max, yAcc Standard deviation
P015	xMag Max, xGyr Standard deviation, yAcc Max
P016	xMag Max, zGyr Max, xGyr Variance, xAcc Variance, yAcc Max, yAcc Interquartile range, yMag wear Max, zAcc Median frequency, xMag wear Max, xMag Min, zGyr Median frequency, yAcc Variance
P018	xAcc Interquartile range, zAcc Interquartile range, yAcc Min, xMag wear Max, zMag wear Spectral centroid, yGyr Interquartile range, yAcc wear Max, zMag Min
P019	yAcc Max, zGyr wear Interquartile range, yMag wear Mean, xMag Max, xGyr Standard deviation
P020	xAcc Max, xGyr Standard deviation, xMag wear Min, xGyr wear Interquartile range

Legend: **SSMs**- Subject specific models; **wear**- watch.

Table II.4: SSMs feature sets AGG, modality smartphone for basic activities

Subject	Feature set
P001	xAcc Spectral entropy, zAcc Standard deviation, yMag Mean, yGyr Spectral entropy, yAcc Spectral centroid, xAcc Standard deviation, zMag Max, yMag Max
P002	zGyr Variance, xAcc Max, xGyr Variance, zGyr Standard deviation, xGyr Interquartile range, zGyr Min, yMag Mean, zMag Standard deviation, zMag Max
P003	zAcc Interquartile range, zMag Max
P004	yMag Spectral centroid, xMag Spectral centroid, yAcc Median frequency, yMag Spectral entropy, xGyr Min, xGyr Spectral entropy, yGyr Spectral entropy
P005	zGyr Max, yAcc Min, yMag Min
P006	yAcc Max, zMag Max
P007	zMag Max, yAcc Max
P008	yGyr Spectral centroid, yMag Spectral entropy, yGyr Median frequency, xGyr Min, zGyr Median, yAcc Interquartile range, zAcc Spectral entropy, zGyr Median frequency, xMag Median frequency, xMag Standard deviation, zGyr Spectral entropy, zMag Max, yAcc Median
P009	xAcc Standard deviation, yAcc Max, zMag Mean
P010	zAcc Max, yAcc Spectral centroid, zAcc Median frequency, yGyr Spectral entropy, yAcc Min, xMag Spectral centroid, xAcc Spectral entropy, xGyr Max, yMag Max, xMag Max
P011	yAcc Max, xMag Max
P012	yMag Max, xMag Max, xAcc Min, yAcc Interquartile range
P013	zMag Max, zAcc Max
P014	zMag Max, yAcc Standard deviation
P015	xGyr Standard deviation, xMag Max, yAcc Max
P016	yGyr Standard deviation, yMag Mean, xAcc Variance, yAcc Max, xGyr Min, zGyr Spectral entropy, zMag Mean
P018	yAcc Max, yAcc Median frequency, xAcc Variance, zMag Min
P019	zAcc Min, xAcc Variance, xMag Min
P020	zMag Mean, yGyr Max, zGyr Standard deviation, xMag Max, zGyr Max

Legend: **SSMs**- Subject specific models.

Table II.5: SSMs feature sets, AGG modality smartwatch for basic activities

Subject	Feature set
P001	zGyr wear Median frequency, yAcc wear Median frequency, xGyr wear Standard deviation, yMag wear Min, zGyr wear Spectral entropy, xAcc wear Min, xMag wear Max, zMag wear Spectral centroid, yAcc wear Standard deviation, zMag wear Max, xGyr wear Variance
P002	xAcc wear Min, yMag wear Median frequency, yAcc wear Spectral entropy, xMag wear Min, zGyr wear Max, yMag wear Mean, xGyr wear Standard deviation, zGyr wear Standard deviation
P003	zMag wear Max, yGyr wear Standard deviation, xMag wear Max
P004	xGyr wear Min, xGyr wear Spectral entropy, yMag wear Max, xMag wear Max
P005	zAcc wear Min, xGyr wear Min, xAcc wear Min, xMag wear Min, xMag wear Max, xMag wear Spectral centroid
P006	yMag wear Spectral centroid, zGyr wear Min, xMag wear Max, zMag wear Max
P007	yGyr wear Interquartile range, xMag wear Mean
P008	xAcc wear Interquartile range, zGyr wear Median frequency, zMag wear Max, yGyr wear Median frequency, zMag wear Spectral centroid, zGyr wear Spectral entropy, xGyr wear Max, zAcc wear Median frequency, yMag wear Spectral centroid
P009	zMag wear Mean, zGyr wear Min, xMag wear Mean, xMag wear Max
P010	zGyr wear Interquartile range, zMag wear Min, yMag wear Standard deviation, xMag wear Min
P011	xMag wear Max, zGyr wear Interquartile range
P012	yGyr wear Interquartile range, zAcc wear Min, xGyr wear Interquartile range, yMag wear Max, yGyr wear Spectral entropy, zMag wear Spectral centroid, xMag wear Max, xAcc wear Interquartile range, zMag wear Spectral entropy, xMag wear Min, yMag wear Mean, zGyr wear Spectral centroid
P013	yMag wear Min, xMag wear Mean, yMag wear Interquartile range, zGyr wear Interquartile range
P014	zMag wear Mean, yMag wear Standard deviation, yGyr wear Spectral centroid, xMag wear Min, yMag wear Spectral centroid, yMag wear Interquartile range, yAcc wear Spectral entropy, xMag wear Mean, xGyr wear Standard deviation, zGyr wear Standard deviation, yMag wear Mean

Continued on next page

Table II.5 SSMs feature sets AGG, modality smartwatch for basic activities

Subject	Feature set
P015	yGyr wear Interquartile range, yGyr wear Variance, yMag wear Spectral centroid, yMag wear Max, xGyr wear Standard deviation, zGyr wear Median frequency, yGyr wear Spectral entropy, xMag wear Max, yAcc wear Standard deviation, xGyr wear Median, zGyr wear Median
P016	zGyr wear Max, xAcc wear Standard deviation, yMag wear Mean, yMag wear Interquartile range, zAcc wear Max, xMag wear Mean, zGyr wear Interquartile range, xMag wear Spectral centroid, xGyr wear Interquartile range
P018	zGyr wear Mean, xGyr wear Min, zGyr wear Standard deviation, xMag wear Max, zMag wear Min, xAcc wear Spectral entropy, yGyr wear Standard deviation, yMag wear Max, xAcc wear Max, xMag wear Spectral entropy, zMag wear Spectral entropy, zGyr wear Median frequency, yAcc wear Standard deviation, xMag wear Mean, yAcc wear Median
P019	yMag wear Median frequency, xGyr wear Standard deviation, zGyr wear Interquartile range, zAcc wear Standard deviation, xMag wear Max, yMag wear Mean, yMag wear Min, xGyr wear Max
P020	xMag wear Max, yAcc wear Spectral centroid, zMag wear Mean, zGyr wear Standard deviation, zGyr wear Min

Legend: **SSMs**- Subject specific models; **wear**- watch.

Table II.6: SSMs feature sets AGG, modality smartphone and -watch for basic activities

Subject	Feature set
P001	xGyr Standard deviation, yMag wear Mean, xAcc wear Standard deviation
P002	xAcc wear Interquartile range, zAcc Min, xMag wear Min, yMag wear Spectral centroid, zMag Max, zAcc wear Max, yAcc wear Min
P003	xMag wear Max, zAcc Interquartile range
P004	xMag wear Max, yMag wear Mean
P005	zGyr Interquartile range, yAcc Max, yMag Min
P006	zMag Max, xGyr Min
P007	zMag Max, yAcc Min
P008	yAcc Max, xAcc Spectral entropy, xAcc Max, zAcc Interquartile range, xMag Spectral centroid, zMag Max
P009	xMag wear Mean, zMag Mean, yAcc Max
P010	yMag Mean, zGyr Max, xGyr Standard deviation, zMag wear Min, xGyr wear Standard deviation

Continued on next page

Table II.6 SSMs feature sets, AGG modality smartphone and -watch for basic activities

Subject	Feature set
P011	zGyr Standard deviation, xMag wear Max
P012	zAcc Max, xGyr Variance, xMag wear Min, zAcc Standard deviation, xGyr Max, yMag Max
P013	yAcc Interquartile range, zMag Max
P014	yAcc Standard deviation, zMag Max
P015	xMag Max, xGyr Standard deviation, yAcc Max
P016	xMag wear Min, yMag wear Max, xGyr Max, yAcc Interquartile range, yAcc Min, xMag Min
P018	xGyr Min, xMag wear Max, yGyr Standard deviation, yAcc Min
P019	yAcc Interquartile range, xMag Min, zGyr wear Variance
P020	yAcc Min, xMag Max

Legend: **SSMs**- Subject specific models; **wear**- watch.

Table II.7: SSMs feature sets GMM, modality smartphone for basic activities

Subject	Feature set
P001	yGyr Spectral centroid, zMag Max, yGyr Standard deviation, zAcc Spectral entropy, yMag Mean, zGyr Spectral entropy, xMag Max, zMag Median frequency, xMag Spectral centroid, xAcc Interquartile range, yGyr Median frequency, xAcc Standard deviation
P002	yGyr Min, zGyr Interquartile range, zAcc Min, yAcc Max, zGyr Spectral entropy, yGyr Spectral entropy, yMag Max, zMag Max, xAcc Median frequency, yAcc Interquartile range, xMag Max, yGyr Max, yMag Spectral centroid, zAcc Median frequency
P003	yAcc Max, yMag Max
P004	zAcc Median frequency, xAcc Max, xAcc Interquartile range, zGyr Median, xAcc Median frequency, zGyr Spectral entropy, xAcc Spectral entropy, yAcc Spectral entropy, xMag Max, zGyr Spectral centroid, zGyr Interquartile range, zMag Max, yMag Spectral entropy, xGyr Mean
P005	zMag Max, yMag Min, xGyr Max
P006	yAcc Min, zMag Max
P007	zMag Max, xAcc Standard deviation
P008	yAcc Max, zMag Mean, yAcc Min
P009	yAcc Min, zMag Mean, zGyr Standard deviation
P010	yAcc Interquartile range, zMag Max, xMag Max, yAcc Max
P011	zGyr Standard deviation, yGyr Standard deviation, xMag Max
P012	xGyr Min, xMag Max, xAcc Max, yAcc Max

Continued on next page

Table II.7 SSMs feature sets, GMM modality smartphone for basic activities

Subject	Feature set
P013	zMag Max, yAcc Max
P014	yAcc Standard deviation, zMag Max
P015	zAcc Standard deviation, xGyr Standard deviation, xMag Max
P016	xGyr Max, xGyr Spectral entropy, yMag Max, xMag Spectral entropy, zMag Max, zAcc Median frequency, yGyr Median frequency, yAcc Spectral entropy, yGyr Spectral centroid, zGyr Spectral centroid, yGyr Interquartile range
P018	zAcc Interquartile range, yMag Mean, zGyr Standard deviation, xMag Min, zGyr Median frequency, yAcc Spectral entropy, yAcc Variance, zAcc Spectral centroid, yGyr Variance, yGyr Median frequency, yGyr Spectral centroid, zMag Mean, xGyr Median frequency, yAcc Median frequency, xGyr Interquartile range
P019	xGyr Standard deviation, yAcc Min, xMag Mean
P020	xGyr Min, xMag Max

Legend: **SSMs**- Subject specific models.

Table II.8: SSMs feature sets, GMM modality smartwatch for basic activities

Subject	Feature set
P001	yMag wear Max, xAcc wear Standard deviation, xMag wear Min, xMag wear Max
P002	xAcc wear Interquartile range, xMag wear Min, yAcc wear Spectral entropy, yMag wear Min
P003	yGyr wear Standard deviation, xMag wear Max
P004	xMag wear Max, yMag wear Spectral centroid
P005	zAcc wear Median frequency, xAcc wear Interquartile range, xAcc wear Spectral entropy, xMag wear Min, yGyr wear Spectral entropy, xMag wear Spectral centroid, zMag wear Max, xMag wear Max
P006	xMag wear Max
P007	xMag wear Min, xMag wear Max
P008	zAcc wear Interquartile range, zAcc wear Spectral entropy, zMag wear Max, yMag wear Spectral centroid, zGyr wear Spectral entropy, zGyr wear Standard deviation, yMag wear Mean, zGyr wear Median, xGyr wear Median frequency, zGyr wear Variance, zAcc wear Median frequency
P009	xMag wear Min, zGyr wear Standard deviation, xMag wear Max, zMag wear Mean

Continued on next page

Table II.8 SSMs feature sets, GMM modality smartwatch for basic activities

Subject	Feature set
P010	xMag wear Min, yMag wear Interquartile range, zMag wear Min, zMag wear Mean, zGyr wear Min
P011	zGyr wear Interquartile range, xMag wear Max
P012	xAcc wear Max, xMag wear Mean, zMag wear Min
P013	zGyr wear Min, yMag wear Max, xMag wear Mean
P014	zMag wear Spectral entropy, xAcc wear Max, yMag wear Min, yAcc wear Spectral entropy, xAcc wear Median frequency, xMag wear Min, zAcc wear Spectral entropy, zGyr wear Max, xAcc wear Spectral entropy
P015	zAcc wear Standard deviation, zAcc wear Median frequency, xMag wear Max, yMag wear Standard deviation, xMag wear Mean, yAcc wear Median frequency, yMag wear Spectral entropy
P016	yMag wear Mean, zGyr wear Max, zGyr wear Spectral centroid, xMag wear Max, zGyr wear Interquartile range, zGyr wear Spectral entropy, zMag wear Mean, xAcc wear Spectral entropy, xMag wear Mean
P018	zGyr wear Spectral entropy, zMag wear Min, zGyr wear Spectral centroid, yMag wear Variance, xMag wear Interquartile range, xMag wear Max, xMag wear Spectral centroid, zGyr wear Variance, yMag wear Median, yGyr wear Standard deviation, yMag wear Standard deviation, xAcc wear Max, zGyr wear Max, zMag wear Interquartile range, yAcc wear Interquartile range
P019	yAcc wear Spectral entropy, xMag wear Mean, yGyr wear Spectral entropy, yMag wear Mean, xGyr wear Interquartile range, yGyr wear Max, zGyr wear Median, yAcc wear Min, xGyr wear Median frequency
P020	zGyr wear Max, xMag wear Max

Legend: **SSMs**- Subject specific models; **wear**- watch.

Table II.9: SSMs feature sets, GMM modality smartphone and -watch for basic activities

Subject	Feature set
P001	yAcc wear Max, yMag Spectral centroid, xMag Spectral centroid, yAcc Max, yMag Min, zMag Max, xMag wear Mean
P002	xGyr Max, xAcc wear Min, yAcc Min, xMag wear Min, zMag Max, yMag wear Max
P003	yAcc Min, xMag wear Max
P004	zGyr Max, yMag wear Mean, yAcc Spectral entropy, xMag Max, xMag wear Max

Continued on next page

Table II.9 SSMs feature sets, GMM modality smartphone and -watch for basic activities

Subject	Feature set
P005	zGyr Interquartile range, yAcc Interquartile range, yMag Mean, xMag Max
P006	zAcc Standard deviation, zGyr wear Min, zMag Max
P007	xMag wear Mean, zAcc Standard deviation, zMag Max
P008	xAcc Min, yAcc Max, yMag wear Mean, xMag Max, zAcc Max, yAcc Interquartile range
P009	yAcc Interquartile range, xMag wear Mean, zMag Mean
P010	yAcc Max, xMag Max, zGyr wear Interquartile range, zMag Max, zMag wear Min
P011	xGyr Interquartile range, xMag Max, yAcc Min
P012	xAcc Standard deviation, xGyr Min, xMag wear Min, xMag Max, zAcc Max, yMag Max
P013	zMag Max, zAcc Max, zAcc Standard deviation
P014	yAcc Standard deviation, zMag Max
P015	xGyr Standard deviation, xMag Max, zAcc Standard deviation
P016	yAcc Standard deviation, zMag wear Max, yMag wear Min, xMag Mean, xMag wear Max, zMag wear Min
P018	xMag wear Max, yGyr Standard deviation, xAcc Min, yMag wear Max, yAcc Min
P019	xAcc Standard deviation, zGyr Interquartile range, xGyr Min, xMag wear Min, xMag Min
P020	yMag Min, zMag Min, zAcc Spectral entropy, xMag wear Mean, xAcc Max, zGyr Min

Legend: **SSMs**- Subject specific models; **wear**- watch.

II.1.2 All Activities

Table II.10: SSMs feature sets KMeans, modality smartphone for all activities

Subject	Feature set
P001	zAcc Min, yAcc Mean, yGyr Spectral entropy, zAcc Median frequency, yGyr Interquartile range, xGyr Max, yGyr Standard deviation, xMag Min, zMag Min, zMag Max, xMag Max, xMag Spectral centroid, xMag Mean
P002	zAcc Median frequency, xAcc Min, zMag Max, xAcc Max, yGyr Standard deviation, yAcc Standard deviation, xGyr Min, yGyr Median frequency

Continued on next page

Table II.10 SMMs feature sets KMeans, modality smartphone for all activities

Subject	Feature set
P003	xGyr Spectral centroid, zGyr Standard deviation, yAcc Standard deviation, zAcc Spectral centroid, yMag Max, xMag Mean, xMag Min, xMag Max
P004	zMag Min, xMag Max, zGyr Max, yGyr Max, yGyr Interquartile range, xMag Mean, zGyr Median frequency, yAcc Max
P005	zGyr Max, yAcc Interquartile range, yAcc Max, xMag Max, yGyr Min
P006	yMag Spectral centroid, xMag Max, xGyr Standard deviation, yAcc Min, xAcc Max, zGyr Median frequency, yGyr Median frequency, yGyr Spectral entropy, zMag Max, yAcc Spectral entropy
P007	xGyr Max, yMag Standard deviation, zAcc Min, xAcc Variance, yAcc Min, xAcc Median frequency, xMag Spectral entropy, zMag Standard deviation, zGyr Median frequency, zMag Max, xAcc Interquartile range, xMag Min
P008	yGyr Mean, zAcc Variance, xMag Max, yGyr Spectral entropy, yAcc Interquartile range, zMag Max, yGyr Interquartile range, yAcc Spectral entropy, yAcc Max, yMag Max
P009	zMag Min, zMag Standard deviation, yGyr Interquartile range, yMag Standard deviation, yGyr Spectral entropy, zGyr Min, yAcc Standard deviation, xGyr Standard deviation, xGyr Median frequency, zAcc Median frequency, yGyr Median frequency, xGyr Min, zGyr Spectral entropy
P010	yGyr Variance, xMag Max, zAcc Min, zAcc Variance, xGyr Interquartile range, yAcc Max
P011	zAcc Median frequency, yGyr Min, xAcc Min, zMag Max, xMag Spectral centroid, zMag Mean, xGyr Min, zGyr Min, xAcc Spectral entropy, zMag Spectral entropy, zGyr Spectral entropy, zMag Spectral centroid, yMag Max, xMag Max, yAcc Median frequency, yGyr Standard deviation, yAcc Interquartile range, zMag Min, zGyr Median frequency
P012	xAcc Spectral centroid, zMag Spectral centroid, zMag Min, xGyr Max, xAcc Variance, yAcc Spectral entropy, xMag Max, yMag Spectral entropy, xAcc Max, zAcc Interquartile range, xGyr Standard deviation, yAcc Max
P013	yAcc Standard deviation, zGyr Min, xMag Mean, yAcc Interquartile range, xGyr Interquartile range, yGyr Standard deviation, zAcc Median frequency
P014	yGyr Max, yGyr Standard deviation, yAcc Max, xMag Max, zMag Spectral centroid, xGyr Standard deviation

Continued on next page

Table II.10 SMMs feature sets KMeans, modality smartphone for all activities

Subject	Feature set
P015	xGyr Interquartile range, yMag Standard deviation, yGyr Min, yGyr Interquartile range, yMag Mean, zMag Spectral entropy, yMag Spectral entropy, zMag Min, zMag Max, xMag Max, zMag Spectral centroid, zAcc Median frequency, yGyr Standard deviation, zAcc Spectral entropy
P016	xAcc Standard deviation, yGyr Interquartile range, yAcc Max, zMag Mean, zGyr Spectral entropy, yMag Mean, yMag Max, zAcc Median frequency
P018	yAcc Median, zMag Mean, zMag Min, xGyr Interquartile range, zAcc Standard deviation, xGyr Max, xGyr Median frequency, yGyr Spectral entropy, yAcc Median frequency, yGyr Standard deviation, xAcc Standard deviation
P019	zGyr Spectral centroid, yAcc Max, yGyr Spectral entropy, xGyr Max, zMag Standard deviation, zMag Min, xGyr Interquartile range, zGyr Min, yGyr Standard deviation, yGyr Median frequency, zMag Spectral centroid, zGyr Standard deviation, xMag Max, zAcc Median frequency, zGyr Interquartile range
P020	zMag Spectral centroid, yAcc Min, zAcc Min, yMag Standard deviation, yGyr Standard deviation, zGyr Min, xGyr Variance, xMag Max, zGyr Spectral entropy, yMag Interquartile range, xAcc Max, zGyr Standard deviation, zAcc Standard deviation, xMag Spectral entropy, yGyr Spectral entropy, yAcc Standard deviation

Legend: **SSMs**- Subject specific models.

Table II.11: SSMs feature sets KMeans, modality smartwatch for all activities

Subject	Feature set
P001	xAcc wear Spectral entropy, xAcc wear Standard deviation, zGyr wear Max, xAcc wear Median frequency, zGyr wear Median frequency, xMag wear Max, zMag wear Max, yMag wear Max
P002	zGyr wear Min, zAcc wear Median frequency, zMag wear Median frequency, yMag wear Min, yAcc wear Standard deviation, xGyr wear Spectral entropy, xMag wear Mean, xAcc wear Median frequency, yGyr wear Spectral entropy

Continued on next page

Table II.11 SSMs feature sets KMeans, modality smartwatch for all activities

Subject	Feature set
P003	yMag wear Min, yAcc wear Standard deviation, zAcc wear Max, yMag wear Interquartile range, yGyr wear Spectral entropy, xMag wear Min, zGyr wear Min, xMag wear Median frequency, xMag wear Spectral entropy, zGyr wear Interquartile range, xMag wear Max, xAcc wear Standard deviation
P004	xAcc wear Spectral entropy, xMag wear Standard deviation, yMag wear Max, yMag wear Interquartile range, zAcc wear Interquartile range, xMag wear Mean, zMag wear Mean, zAcc wear Median frequency, zMag wear Median frequency
P005	yMag wear Spectral entropy, zGyr wear Min, xAcc wear Spectral entropy, xGyr wear Interquartile range, xMag wear Min, yMag wear Spectral centroid, yGyr wear Spectral entropy, yMag wear Min
P006	zGyr wear Min, yGyr wear Spectral entropy, zGyr wear Spectral entropy, zMag wear Max, xGyr wear Median frequency, yMag wear Min, xMag wear Max, zGyr wear Max, zMag wear Mean, xAcc wear Median frequency
P007	xAcc wear Interquartile range, xMag wear Min, yGyr wear Spectral entropy, xMag wear Max, xAcc wear Median frequency
P008	yGyr wear Max, yGyr wear Min, xMag wear Max, zAcc wear Spectral entropy, xAcc wear Median frequency, zGyr wear Standard deviation, yGyr wear Median frequency, yMag wear Max, yGyr wear Spectral entropy, zMag wear Max, zGyr wear Spectral entropy
P009	xMag wear Spectral entropy, zGyr wear Max, xAcc wear Max, zMag wear Standard deviation, xMag wear Max, xAcc wear Median frequency, yAcc wear Spectral entropy, yMag wear Min, zAcc wear Spectral entropy, zAcc wear Standard deviation, xMag wear Min
P010	yMag wear Standard deviation, zGyr wear Max, yGyr wear Interquartile range, zAcc wear Interquartile range, xAcc wear Median frequency, xMag wear Max, xMag wear Mean, xMag wear Min, yAcc wear Spectral centroid, xAcc wear Standard deviation, yMag wear Spectral entropy
P011	xAcc wear Median frequency, yMag wear Median frequency, zMag wear Max, zGyr wear Spectral entropy, xMag wear Max, xMag wear Spectral entropy, zGyr wear Standard deviation, xGyr wear Max, xGyr wear Interquartile range
P012	xGyr wear Interquartile range, zGyr wear Standard deviation, xMag wear Min, zGyr wear Spectral entropy, yAcc wear Spectral entropy, xAcc wear Min, xMag wear Max, xMag wear Mean

Continued on next page

Table II.11 SSMs feature sets KMeans, modality smartwatch for all activities

Subject	Feature set
P013	zMag wear Mean, zMag wear Max, zGyr wear Min, yAcc wear Spectral entropy, zGyr wear Max, xMag wear Spectral centroid, yMag wear Min, zMag wear Spectral centroid, yMag wear Mean, xMag wear Min
P014	zGyr wear Max, xMag wear Max, yGyr wear Min, xGyr wear Spectral entropy, xAcc wear Max, yAcc wear Spectral entropy, zMag wear Spectral entropy, zAcc wear Interquartile range, yMag wear Standard deviation
P015	zMag wear Spectral entropy, xMag wear Spectral centroid, zGyr wear Interquartile range, yAcc wear Min, xGyr wear Standard deviation, xGyr wear Max, zGyr wear Max, xMag wear Mean, xMag wear Max, zGyr wear Spectral entropy, yMag wear Spectral entropy, zAcc wear Max
P016	yGyr wear Spectral entropy, zMag wear Spectral entropy, xGyr wear Interquartile range, yMag wear Spectral entropy, xMag wear Spectral entropy, xMag wear Min, zGyr wear Interquartile range, xAcc wear Spectral centroid, zGyr wear Min, yMag wear Min, xAcc wear Standard deviation, xGyr wear Spectral centroid, zGyr wear Spectral entropy, zMag wear Max
P018	yMag wear Min, yAcc wear Standard deviation, yGyr wear Min, zMag wear Min, xGyr wear Interquartile range, zMag wear Spectral entropy, zMag wear Spectral centroid, zMag wear Interquartile range, xMag wear Mean, zGyr wear Interquartile range, xMag wear Max, xMag wear Min, yGyr wear Spectral entropy
P019	yGyr wear Max, yGyr wear Standard deviation, yMag wear Mean, zGyr wear Spectral entropy, xGyr wear Max, yMag wear Min, yAcc wear Median frequency, xGyr wear Median frequency, xMag wear Min, xAcc wear Min
P020	zAcc wear Standard deviation, zGyr wear Max, xAcc wear Min, zAcc wear Median frequency, xMag wear Max, yGyr wear Interquartile range, zMag wear Max, zMag wear Spectral entropy, yMag wear Spectral entropy

Legend: **SSMs**- Subject specific models; **wear**- watch.

Table II.12: SSMs feature sets KMeans, modality smartphone and -watch for all activities

Subject	Feature set
P001	yGyr Median frequency, yAcc Max, xAcc Median frequency, yMag wear Max, xGyr wear Min, xAcc Min, xMag Max, xMag Mean, xMag wear Mean, xGyr wear Standard deviation, zMag wear Mean, zMag wear Spectral entropy, zMag wear Max, xMag Spectral centroid, zMag Min, zMag wear Min
P002	zGyr Interquartile range, zGyr wear Min, xGyr Min, yAcc Max, yGyr Standard deviation, zMag Min
P003	zGyr Median frequency, xGyr Standard deviation, yAcc Min, yMag wear Min, yMag wear Standard deviation, zMag Spectral centroid, xMag wear Mean, xMag wear Spectral entropy, yAcc wear Standard deviation, xMag Max, xMag Mean, yMag Spectral centroid, xMag Min
P004	yMag wear Median frequency, yMag wear Standard deviation, zAcc Median frequency, xAcc Min, zAcc wear Interquartile range, xGyr wear Standard deviation, xGyr Spectral entropy, yMag wear Max, xMag Mean, xMag Spectral centroid, zGyr wear Spectral entropy, xMag wear Min, zMag Mean, xMag Max, zAcc Interquartile range, yMag wear Mean, yAcc Max, zMag Spectral centroid, xMag Min
P005	xAcc Min, zMag wear Max, xAcc Max, yGyr Spectral entropy, yMag wear Mean, zAcc wear Spectral centroid, xMag Max, xGyr Median frequency, yMag wear Median frequency, yMag wear Max, xGyr Min, yMag Max, zMag Spectral centroid, yMag Mean, yMag Min, xMag Mean
P006	zGyr wear Median, yGyr Max, zMag wear Min, yAcc wear Min, yGyr Standard deviation, zGyr Standard deviation, xAcc Standard deviation, yGyr Spectral entropy, yMag wear Mean, zMag wear Max, yAcc Standard deviation, xMag Median frequency, yMag Spectral centroid
P007	yAcc wear Spectral centroid, zGyr wear Interquartile range, xMag wear Min, xAcc Max, zAcc Interquartile range, yGyr Standard deviation, xGyr Spectral entropy, yGyr Spectral entropy, zMag Min, yAcc Standard deviation, zAcc wear Min, xAcc wear Standard deviation, xMag wear Max
P008	xGyr Spectral entropy, yAcc Max, xMag Spectral entropy, zGyr Median frequency, xGyr Interquartile range, zGyr wear Spectral entropy, zAcc wear Min, xMag Spectral centroid, yAcc wear Max, yGyr wear Median frequency, zMag wear Min, xMag wear Max, zMag Max, yGyr Spectral entropy, yMag Max, zAcc wear Median frequency, yMag wear Spectral centroid

Continued on next page

Table II.12 SSMs feature sets, KMeans modality smartphone and -watch for all activities

Subject	Feature set
P009	xMag wear Spectral entropy, yAcc Median frequency, yGyr Min, xMag Max, zGyr Max, yGyr Interquartile range, xMag wear Mean, xGyr wear Max, zAcc wear Spectral entropy, xGyr Median frequency, zGyr wear Interquartile range, zMag Median frequency, zMag wear Spectral centroid, zMag Spectral centroid, xGyr Spectral entropy, yMag wear Max, yAcc Max, xAcc Median frequency
P010	zAcc Spectral centroid, yGyr wear Min, zAcc Spectral entropy, xAcc Variance, xGyr Min, zMag Max, zAcc Standard deviation, xMag Spectral centroid, xAcc Max, zMag Spectral entropy, xMag wear Median frequency, zGyr Median frequency, xAcc wear Spectral entropy, xMag wear Mean, xMag wear Spectral entropy, zMag wear Min, xMag Max
P011	xAcc Max, xMag wear Max, zGyr Min, yAcc Standard deviation, xMag Max, zMag wear Max, yMag wear Min, zMag Min, yMag Max, yGyr Spectral entropy
P012	xAcc wear Max, zGyr Interquartile range, yAcc Spectral entropy, yMag wear Mean, zGyr Min, xAcc Min, xGyr Interquartile range, yAcc Max, yMag Max, zMag Mean, zMag Max
P013	xMag wear Spectral centroid, yMag wear Max, xMag wear Max, xMag wear Spectral entropy, yGyr wear Spectral entropy, zGyr wear Interquartile range, yGyr Min, xMag Max, xGyr wear Max, yAcc wear Median frequency, yAcc wear Spectral entropy, zAcc Median frequency, xMag Mean, yMag wear Min, zAcc Min, zAcc Interquartile range, zMag wear Spectral centroid, zAcc Standard deviation, zMag wear Min
P014	zAcc wear Min, yGyr wear Interquartile range, xGyr Max, zAcc wear Standard deviation, xMag wear Max, yAcc Min, yGyr wear Spectral centroid, xGyr wear Max, yGyr Spectral entropy, xMag Spectral entropy, xGyr wear Spectral entropy, xAcc wear Median frequency, xMag Max, xMag Spectral centroid, zAcc wear Spectral entropy, xGyr Standard deviation
P015	zMag Min, zAcc Spectral entropy, zGyr Max, xMag wear Spectral entropy, xGyr Min, zGyr wear Spectral entropy, yMag wear Min, zMag Standard deviation, xMag wear Max, zGyr Interquartile range, xGyr wear Standard deviation
P016	xGyr Spectral entropy, yMag wear Max, yAcc Interquartile range, xAcc Min, zMag Max, xGyr wear Spectral entropy, yGyr Interquartile range, xMag wear Max, yMag wear Mean, yGyr Standard deviation, zMag Spectral entropy

Continued on next page

Table II.12 SSMs feature sets, KMeans modality smartphone and -watch for all activities

Subject	Feature set
P018	zMag wear Standard deviation, xMag wear Standard deviation, zGyr Max, xGyr Min, zAcc Min, xGyr Interquartile range, zMag Max, xAcc Interquartile range, zMag Min, yAcc Standard deviation, yGyr wear Spectral entropy, yAcc Median frequency, zMag wear Max, xGyr Median frequency
P019	xAcc Min, zMag wear Mean, xAcc Standard deviation, zAcc Median frequency, xMag wear Min, xGyr wear Interquartile range, xGyr wear Standard deviation, zMag Spectral entropy, yMag wear Spectral centroid, zGyr Standard deviation, zGyr Median frequency, zMag wear Spectral entropy, yMag Min, yMag wear Mean, yMag wear Min, yGyr Median frequency, yAcc Interquartile range
P020	zAcc wear Spectral centroid, xGyr Variance, yGyr wear Spectral entropy, zGyr Min, xAcc Max, yMag wear Median frequency, zGyr Interquartile range, zAcc Interquartile range, zAcc wear Spectral entropy, zAcc Standard deviation, zAcc Median frequency, xMag wear Max, yAcc Max, yMag wear Spectral entropy, xMag Max, zGyr Spectral entropy, yMag Spectral entropy, xMag Spectral entropy

Legend: **SSMs**- Subject specific models; **wear**- watch.

Table II.13: SSMs feature sets, AGG modality smartphone for all activities

Subject	Feature set
P001	zMag Min, yAcc Spectral entropy, xAcc Standard deviation, yGyr Standard deviation, zAcc Median frequency, xMag Spectral centroid, yGyr Spectral entropy, yAcc Mean
P002	yMag Max, zAcc Median frequency, zMag Min, xGyr Min, zGyr Spectral centroid, yAcc Max, yGyr Standard deviation, xMag Min
P003	xMag Min, yAcc Max, xMag Spectral centroid, xAcc Min, yAcc Median frequency, yGyr Max, xGyr Median frequency, yMag Max
P004	zAcc Interquartile range, xGyr Spectral entropy, xMag Mean, zMag Standard deviation, yAcc Max, xGyr Median frequency, zGyr Max
P005	xAcc Min, yMag Max, zGyr Median frequency, yMag Min, yAcc Min, yAcc Standard deviation, xMag Mean, xAcc Max, xAcc Interquartile range

Continued on next page

Table II.13 SSMs feature sets, AGG modality smartphone for all activities

Subject	Feature set
P006	zGyr Spectral entropy, yMag Spectral entropy, yGyr Median frequency, yGyr Spectral entropy, xGyr Min, xAcc Spectral centroid, zMag Max, xMag Standard deviation, yAcc Variance, xGyr Median frequency, xMag Spectral centroid, xMag Median frequency, zAcc Standard deviation
P007	xAcc Max, zGyr Median frequency, yAcc Interquartile range, zMag Max, xAcc Spectral entropy, zGyr Spectral entropy, xGyr Variance, yGyr Interquartile range
P008	yGyr Mean, zMag Max, yAcc Spectral entropy, xGyr Standard deviation, xGyr Variance, yAcc Max
P009	xMag Interquartile range, xAcc Standard deviation, zMag Min, xGyr Min, yGyr Spectral entropy, yAcc Min, xMag Spectral entropy, yAcc Max, yGyr Median frequency
P010	zGyr Median frequency, xGyr Min, zGyr Variance, zMag Max, yAcc Interquartile range, yGyr Spectral entropy, xMag Max, zAcc Median frequency, zMag Standard deviation, xMag Spectral centroid, yAcc Max
P011	yAcc Min, zGyr Spectral entropy, xAcc Standard deviation, yGyr Spectral entropy, zMag Max, xAcc Interquartile range, zMag Mean, yAcc Max
P012	xGyr Variance, yGyr Standard deviation, zMag Mean, zGyr Max, xMag Mean, zMag Min
P013	zGyr Standard deviation, yAcc Interquartile range, xGyr Median frequency, zGyr Spectral entropy, yAcc Standard deviation, xMag Mean, zGyr Max, zAcc Interquartile range, zAcc Spectral entropy, yMag Max, yGyr Standard deviation
P014	xAcc Median frequency, xGyr Interquartile range, yAcc Standard deviation, xMag Spectral centroid, xGyr Spectral centroid, zMag Max, xAcc Max
P015	zAcc Standard deviation, yAcc Max, zMag Max, yGyr Standard deviation, yGyr Spectral entropy, xAcc Variance, xMag Max, zAcc Spectral entropy, xAcc Spectral centroid, yMag Spectral centroid
P016	zMag Max, yMag Max, zGyr Interquartile range, xAcc Median frequency, xAcc Variance, yAcc Standard deviation, xGyr Spectral centroid
P018	zMag Min, zMag Mean, xMag Max, yAcc Min, zAcc Min, xGyr Standard deviation, xAcc Standard deviation
P019	zGyr Spectral entropy, yAcc Variance, xMag Max, yGyr Interquartile range, zAcc Interquartile range, zAcc Spectral entropy, zAcc Min, yAcc Median frequency, xGyr Spectral centroid

Continued on next page

Table II.13 SSMs feature sets, AGG modality smartphone for all activities

Subject	Feature set
P020	xAcc Variance, zAcc Min, zMag Standard deviation, xMag Max, xAcc Max, zAcc Standard deviation, zMag Max, yMag Min, yAcc Spectral entropy, yAcc Standard deviation

Legend: **SSMs**- Subject specific models.

Table II.14: SSMs feature sets AGG, modality smartwatch for all activities

Subject	Feature set
P001	xGyr wear Variance, zMag wear Min, yGyr wear Min, yAcc wear Max, zMag wear Max, yAcc wear Spectral centroid, xMag wear Max, xMag wear Spectral entropy, yMag wear Max, yGyr wear Max
P002	yGyr wear Spectral entropy, zGyr wear Max, yMag wear Standard deviation, yAcc wear Standard deviation, yMag wear Min, zAcc wear Min, xMag wear Min, yAcc wear Median frequency, zGyr wear Interquartile range
P003	yAcc wear Median frequency, xGyr wear Max, zGyr wear Max, xMag wear Max, zGyr wear Standard deviation, yMag wear Min, yGyr wear Max
P004	xGyr wear Spectral entropy, xMag wear Spectral centroid, xMag wear Mean, yGyr wear Interquartile range, zMag wear Min, xAcc wear Spectral entropy, xMag wear Min, xMag wear Max
P005	xAcc wear Spectral entropy, zGyr wear Min, yMag wear Mean, xMag wear Min
P006	xGyr wear Spectral entropy, xGyr wear Max, zGyr wear Min, xMag wear Mean, zGyr wear Spectral entropy, yAcc wear Min, yMag wear Min
P007	yMag wear Max, zAcc wear Min, xMag wear Spectral centroid, xGyr wear Interquartile range, xAcc wear Interquartile range, xAcc wear Min, xMag wear Max, zMag wear Min, xAcc wear Standard deviation, xMag wear Min, yGyr wear Spectral entropy
P008	yGyr wear Median frequency, xAcc wear Min, zGyr wear Spectral entropy, xMag wear Max, zGyr wear Interquartile range, yMag wear Mean, yGyr wear Max
P009	yMag wear Mean, zGyr wear Interquartile range, zGyr wear Median frequency, zGyr wear Variance, xMag wear Max, xGyr wear Interquartile range

Continued on next page

Table II.14 SSMs feature sets, AGG modality smartwatch for all activities

Subject	Feature set
P010	xMag wear Spectral entropy, zAcc wear Median frequency, zMag wear Mean, xGyr wear Min, zGyr wear Min, xMag wear Spectral centroid, xMag wear Median frequency, xMag wear Mean, yAcc wear Interquartile range, xAcc wear Standard deviation
P011	zGyr wear Median, xGyr wear Min, zGyr wear Min, xGyr wear Max, zMag wear Max, yMag wear Max
P012	yAcc wear Spectral entropy, zGyr wear Max, zMag wear Min, xMag wear Min, zMag wear Spectral entropy
P013	zMag wear Min, zMag wear Spectral centroid, xMag wear Spectral entropy, zGyr wear Max, xMag wear Min, yGyr wear Standard deviation, xAcc wear Spectral entropy, zMag wear Mean, zGyr wear Standard deviation, yAcc wear Max, xGyr wear Standard deviation, xAcc wear Median frequency, xGyr wear Max
P014	yMag wear Spectral entropy, xGyr wear Standard deviation, zGyr wear Min, xMag wear Max, yAcc wear Spectral entropy, xAcc wear Standard deviation
P015	xMag wear Standard deviation, xMag wear Max, xMag wear Min, zGyr wear Max, yGyr wear Spectral entropy, zMag wear Min
P016	xMag wear Min, yMag wear Mean, zGyr wear Max, zGyr wear Spectral entropy, zGyr wear Standard deviation
P018	zMag wear Spectral entropy, yAcc wear Min, zGyr wear Interquartile range, xAcc wear Standard deviation, xAcc wear Spectral centroid, zMag wear Max, xMag wear Max, xMag wear Spectral entropy
P019	yMag wear Median frequency, xAcc wear Min, zAcc wear Interquartile range, yMag wear Standard deviation, xAcc wear Max, xAcc wear Standard deviation, yMag wear Min, xMag wear Min, yGyr wear Max
P020	zGyr wear Interquartile range, zMag wear Mean, xMag wear Max, yGyr wear Standard deviation, xGyr wear Standard deviation, zMag wear Spectral centroid, zMag wear Max, xGyr wear Spectral entropy, xAcc wear Spectral centroid

Legend: **SSMs**- Subject specific models; **wear**- watch.

Table II.15: SSMs feature sets, AGG modality smartphone and -watch for all activities

Subject	Feature set
P001	zMag wear Min, xAcc wear Max, yGyr wear Spectral entropy, yGyr Spectral entropy, xAcc wear Median frequency, zAcc Spectral entropy, zGyr Max, zMag wear Mean, zMag Min, xMag Mean, xMag wear Spectral centroid, yAcc Variance, xMag wear Min, yGyr wear Max, zGyr Interquartile range, xMag Min, yMag wear Max
P002	xMag wear Spectral centroid, yGyr wear Standard deviation, zAcc Standard deviation, xGyr wear Standard deviation, zAcc wear Spectral entropy, xAcc Min, yMag wear Min, zGyr wear Spectral entropy, xAcc Variance, yGyr Interquartile range, zMag Mean, yGyr Min, yAcc Interquartile range, zGyr wear Interquartile range, zGyr wear Standard deviation
P003	xMag wear Spectral entropy, zGyr Min, zAcc Min, yMag wear Interquartile range, zGyr Max, xMag wear Mean, xMag Min, zGyr wear Median, zMag Mean, xMag Spectral centroid, yAcc Max
P004	xAcc Spectral entropy, xAcc wear Max, xMag wear Min, yMag wear Spectral centroid, xMag Spectral centroid, xMag wear Spectral centroid, zAcc Variance, zMag Standard deviation, zGyr wear Spectral centroid, xGyr wear Spectral entropy, xGyr Median frequency
P005	yMag wear Min, yGyr Standard deviation, zGyr Min, xGyr Max, zAcc wear Standard deviation, xGyr Median frequency, yMag Mean, zGyr Standard deviation, xAcc Standard deviation, xMag Mean, xMag Spectral entropy, xMag Min, xMag Max
P006	zGyr Variance, xAcc Standard deviation, yMag Spectral centroid, zMag wear Min, zMag wear Standard deviation, zGyr Spectral entropy, zAcc Min
P007	yGyr Min, yAcc wear Standard deviation, zAcc Standard deviation, xGyr Interquartile range, zAcc Min, zMag wear Min, yGyr Max, zGyr wear Median frequency, xMag wear Mean, zMag Mean, zGyr wear Max
P008	yAcc Interquartile range, zAcc Median frequency, xAcc Median frequency, xMag Spectral centroid, xGyr Interquartile range, yMag wear Max, xAcc Min, yGyr Spectral entropy, zGyr wear Standard deviation, xMag wear Max, zMag wear Min
P009	xAcc wear Median frequency, yMag wear Max, xAcc Standard deviation, zGyr wear Median, zGyr Standard deviation, xMag wear Mean, yGyr Spectral entropy

Continued on next page

Table II.15 SSMs feature sets, AGG modality smartphone and -watch for all activities

Subject	Feature set
P010	xMag Max, xMag wear Min, zAcc Standard deviation, yGyr wear Median frequency, zMag Spectral entropy, xGyr wear Median, yAcc Max, yAcc Interquartile range, yGyr wear Standard deviation
P011	xGyr Median frequency, xAcc Standard deviation, yAcc wear Median frequency, xGyr wear Min, yGyr Spectral entropy, yGyr Variance, zAcc Standard deviation, xAcc Spectral entropy, zMag wear Spectral centroid, zMag Max, zGyr wear Min, xGyr Max
P012	xAcc wear Max, zMag Min, xGyr wear Spectral entropy, zAcc Standard deviation, xMag wear Mean, zMag Max, zAcc Median frequency, xGyr Min
P013	xMag wear Mean, yMag wear Mean, xGyr wear Spectral entropy, zAcc Interquartile range, xAcc Standard deviation, yAcc wear Standard deviation, xMag Mean, xAcc wear Spectral centroid, zMag wear Max
P014	yMag wear Min, zAcc Median frequency, zGyr Interquartile range, xMag wear Spectral centroid, xMag wear Mean, xGyr Interquartile range, xMag Mean, yAcc Standard deviation, xGyr Standard deviation
P015	yAcc Interquartile range, zMag wear Interquartile range, xAcc Median frequency, xAcc Spectral entropy, yGyr Standard deviation, zMag Max, xGyr Median frequency, yAcc Variance, yGyr Median frequency, xGyr wear Standard deviation, yGyr Max, xMag wear Mean, yAcc wear Standard deviation, xMag Max, yGyr Variance
P016	xGyr wear Min, xAcc Max, yMag Min, yAcc Median frequency, yGyr Standard deviation, xGyr Median frequency, zMag wear Mean, zMag Mean, xGyr Spectral entropy
P018	yGyr wear Spectral entropy, zMag Min, zMag Median frequency, zGyr Interquartile range, xMag Mean, xMag wear Interquartile range, yAcc wear Min, yAcc Min, yMag wear Min, zMag Standard deviation, yAcc Max, zMag Max, xMag wear Min, yGyr wear Min, xMag Spectral centroid
P019	xMag wear Mean, yGyr Standard deviation, xGyr wear Standard deviation, zGyr Median frequency, yGyr Max, xMag Mean, xGyr wear Spectral centroid, yMag wear Mean, yGyr Spectral centroid
P020	zGyr Spectral entropy, yAcc Standard deviation, xMag wear Max, zGyr Median frequency, xGyr Interquartile range, xAcc Standard deviation, zMag Max, xGyr wear Spectral entropy, yMag Min, xMag Max

Legend: **SSMs**- Subject specific models; **wear**- watch.

Table II.16: SSMs feature sets, GMM modality smartphone for all activities

Subject	Feature set
P001	yGyr Interquartile range, yGyr Max, yGyr Spectral entropy, zMag Max, xMag Mean, zAcc Spectral entropy
P002	xMag Spectral entropy, yAcc Variance, xMag Min, yMag Max, yMag Min, zMag Min
P003	xMag Mean, xGyr Median frequency, yGyr Max, zAcc Spectral entropy, zMag Max, yAcc Variance
P004	zAcc Median frequency, yAcc Max, zMag Max, xMag Mean, xGyr Spectral entropy, yAcc Variance
P005	yMag Mean, xMag Mean, yAcc Standard deviation, yAcc Variance, xMag Max, zMag Spectral centroid
P006	xAcc Spectral centroid, yGyr Interquartile range, yGyr Max, zGyr Max, zAcc Median frequency, yMag Spectral centroid, yAcc Variance, yGyr Spectral entropy, xMag Min, xMag Median frequency, zMag Max
P007	zGyr Standard deviation, zGyr Variance, zAcc Median frequency, xMag Standard deviation, yAcc Min, zMag Max, yAcc Interquartile range, yAcc Standard deviation, yAcc Variance, xMag Mean
P008	xAcc Standard deviation, yAcc Max, zGyr Standard deviation, xMag Spectral centroid, yAcc Interquartile range, xMag Max, yAcc Variance, yGyr Min
P009	xGyr Spectral centroid, xAcc Min, zGyr Interquartile range, yAcc Min, yGyr Spectral entropy, xGyr Spectral entropy, zMag Max, yAcc Standard deviation, zAcc Spectral entropy, yAcc Interquartile range, xMag Mean, zMag Median frequency, zMag Standard deviation
P010	zAcc Min, zGyr Spectral centroid, xGyr Spectral entropy, yGyr Interquartile range, yAcc Spectral entropy, zGyr Median frequency, zMag Max, yMag Max, xMag Spectral centroid
P011	zMag Mean, yAcc Standard deviation, yGyr Interquartile range, yMag Max, zGyr Spectral entropy, yAcc Median frequency
P012	yAcc Spectral entropy, zAcc Min, xAcc Spectral centroid, yGyr Median frequency, xAcc Median frequency, zMag Max, xAcc Spectral entropy, zAcc Spectral centroid, yGyr Spectral entropy, zAcc Median frequency
P013	yGyr Spectral centroid, yGyr Standard deviation, zMag Spectral centroid, yAcc Spectral entropy, yGyr Median frequency, xMag Spectral centroid, zAcc Spectral entropy, yMag Max, xMag Min

Continued on next page

Table II.16 SSMs feature sets, GMM modality smartphone for all activities

Subject	Feature set
P014	zGyr Spectral entropy, xAcc Interquartile range, yAcc Median frequency, xAcc Max, yGyr Interquartile range, xAcc Spectral entropy, yMag Min, xMag Spectral centroid, zAcc Median frequency, xMag Min, yAcc Variance
P015	zGyr Spectral entropy, yMag Spectral entropy, yAcc Variance, zGyr Standard deviation, xAcc Median frequency, zMag Max, yMag Min, yGyr Standard deviation, xMag Max, yGyr Spectral entropy
P016	yAcc Min, xAcc Spectral entropy, xAcc Median frequency, yGyr Spectral entropy, xGyr Spectral centroid, yAcc Median frequency, zAcc Median frequency, xMag Mean
P018	xMag Spectral entropy, zMag Min, xGyr Standard deviation, xGyr Spectral centroid, yAcc Spectral centroid, yGyr Spectral entropy, yMag Mean, yAcc Median frequency, zAcc Median frequency
P019	zAcc Spectral centroid, yMag Max, xMag Min, yAcc Median frequency, yAcc Variance
P020	zMag Spectral entropy, yAcc Spectral entropy, zMag Spectral centroid, xGyr Spectral entropy, yGyr Spectral centroid, zAcc Spectral entropy, zAcc Max, yAcc Interquartile range, xMag Max, zMag Mean, xAcc Spectral entropy, yAcc Variance, yAcc Standard deviation, xAcc Median frequency

Legend: **SSMs**- Subject specific models.

Table II.17: SSMs feature sets, GMM modality smartwatch for all activities

Subject	Feature set
P001	zMag wear Spectral centroid, yGyr wear Interquartile range, yAcc wear Min, xGyr wear Variance, xMag wear Mean, xAcc wear Min, zGyr wear Spectral entropy
P002	yGyr wear Spectral centroid, yGyr wear Median frequency, xGyr wear Spectral centroid, xMag wear Spectral entropy, yMag wear Spectral centroid, zGyr wear Interquartile range, zAcc wear Min, xAcc wear Spectral entropy, yMag wear Mean, xMag wear Min, yGyr wear Interquartile range, zAcc wear Median frequency, zAcc wear Standard deviation, xAcc wear Median frequency, xAcc wear Spectral centroid
P003	zGyr wear Min, yGyr wear Spectral entropy, xGyr wear Spectral entropy, yMag wear Interquartile range, xMag wear Max, xMag wear Spectral entropy, yMag wear Mean, yAcc wear Median frequency

Continued on next page

Table II.17 SSMs feature sets, GMM modality smartwatch for all activities

Subject	Feature set
P004	xMag wear Mean, zMag wear Max, yMag wear Max, zGyr wear Spectral entropy
P005	zGyr wear Variance, zMag wear Max, zAcc wear Median frequency, xMag wear Min, yMag wear Mean, yMag wear Max, xMag wear Mean
P006	zGyr wear Median, xMag wear Max, yGyr wear Spectral entropy, zAcc wear Median frequency, zAcc wear Spectral entropy, xAcc wear Median frequency, yMag wear Spectral centroid, zMag wear Min
P007	yMag wear Interquartile range, yMag wear Mean, xMag wear Max, zMag wear Max, xMag wear Median frequency, zMag wear Standard deviation, zGyr wear Standard deviation, xAcc wear Min, xMag wear Min, yAcc wear Interquartile range, xGyr wear Variance
P008	yGyr wear Min, xGyr wear Spectral entropy, yAcc wear Median frequency, yAcc wear Spectral entropy, xMag wear Max, xMag wear Spectral centroid, zGyr wear Min, xAcc wear Standard deviation
P009	zGyr wear Standard deviation, zMag wear Max, zAcc wear Min, xAcc wear Spectral entropy, yGyr wear Spectral entropy, xMag wear Min, yMag wear Min, xGyr wear Spectral entropy, xMag wear Mean
P010	zGyr wear Interquartile range, zGyr wear Max, xMag wear Mean, zGyr wear Variance, zMag wear Max
P011	zGyr wear Spectral entropy, xGyr wear Standard deviation, zAcc wear Spectral entropy, zGyr wear Median, xAcc wear Median frequency, zMag wear Spectral centroid, yMag wear Spectral centroid, xGyr wear Spectral centroid, zAcc wear Standard deviation, zMag wear Max
P012	yAcc wear Spectral entropy, xAcc wear Min, xMag wear Min
P013	zMag wear Max, xMag wear Min, zGyr wear Interquartile range, yGyr wear Standard deviation
P014	xMag wear Min, xGyr wear Max, zGyr wear Variance
P015	xMag wear Min, xAcc wear Min, xMag wear Max, zMag wear Spectral entropy, yAcc wear Median frequency, zGyr wear Median frequency, xGyr wear Median frequency, yMag wear Min, yGyr wear Standard deviation, yAcc wear Spectral entropy
P016	yMag wear Median frequency, zGyr wear Min, zGyr wear Interquartile range, zAcc wear Standard deviation, xMag wear Max, yMag wear Standard deviation, zMag wear Max
P018	xMag wear Min, yAcc wear Standard deviation, zGyr wear Spectral entropy, yGyr wear Spectral entropy, xMag wear Mean, zMag wear Max, yGyr wear Interquartile range

Continued on next page

Table II.17 SSMs feature sets, GMM modality smartwatch for all activities

Subject	Feature set
P019	zGyr wear Interquartile range, yMag wear Max, yGyr wear Median frequency, zAcc wear Spectral entropy, xAcc wear Spectral entropy, xMag wear Spectral entropy
P020	yAcc wear Median frequency, xMag wear Max, zGyr wear Standard deviation, xMag wear Spectral entropy

Legend: **SSMs**- Subject specific models; **wear**- watch.

Table II.18: SSMs feature sets, GMM modality smartphone and -watch for all activities

Subject	Feature set
P001	zAcc wear Spectral entropy, yMag Max, zAcc Min, xGyr wear Spectral entropy, zAcc Spectral entropy, yGyr Median frequency, yGyr Interquartile range, xAcc wear Median frequency, yMag wear Mean, xMag wear Spectral entropy, yGyr Spectral centroid, zMag Min, xMag Min, xMag wear Mean, xMag Median frequency
P002	zMag Min, xGyr Max, yMag Mean, yAcc Max, yAcc Variance, xMag wear Min
P003	zGyr wear Spectral centroid, xGyr wear Median frequency, yAcc wear Spectral entropy, yAcc Max, xMag Min, yMag wear Min, zMag Max, yMag Spectral centroid, xMag Mean, xMag wear Mean, xMag wear Median frequency
P004	yAcc Min, yGyr wear Spectral centroid, xAcc Median frequency, xMag Max, yGyr Median frequency, zMag Min, xAcc Spectral entropy, yGyr Spectral entropy, xMag wear Min
P005	zGyr wear Standard deviation, yMag wear Spectral centroid, zAcc Spectral entropy, xAcc wear Median frequency, yMag Min, xGyr Spectral entropy, yAcc wear Spectral entropy, yAcc Min, xMag wear Mean, xMag Mean, zMag Max, yMag Mean
P006	xMag wear Spectral entropy, yAcc Standard deviation, yGyr Max, xMag Min, zMag wear Min, zMag wear Median frequency, xAcc wear Spectral centroid, yMag Max, zAcc wear Spectral entropy, zAcc Median frequency
P007	yGyr wear Variance, xMag wear Spectral entropy, yMag Min, yAcc Variance, zMag Min, xAcc Median frequency, zGyr Spectral entropy, yAcc Standard deviation, xAcc wear Median frequency, yGyr Min, xMag wear Median frequency

Continued on next page

Table II.18 SSMs feature sets, GMM modality smartphone and -watch for all activities

Subject	Feature set
P008	xAcc Max, xAcc wear Median frequency, yGyr Median frequency, xGyr Standard deviation, yAcc Median frequency, xMag wear Min, zAcc Median frequency, zMag Max, yAcc Spectral entropy, yAcc Spectral centroid, zGyr wear Variance, xMag Max
P009	xMag Spectral entropy, xGyr Spectral entropy, yMag wear Min, yMag wear Interquartile range, yAcc Max, xAcc Median frequency, zGyr Max, yGyr Spectral entropy, xMag wear Mean, yGyr wear Min, xAcc wear Spectral centroid, xMag Max, xMag wear Median frequency, xAcc wear Median frequency
P010	yAcc wear Spectral centroid, yGyr wear Spectral entropy, yAcc Interquartile range, zMag wear Min, yGyr Spectral centroid, xMag wear Max, xMag Max
P011	yAcc Min, zAcc wear Spectral centroid, yGyr Interquartile range, zAcc Median frequency, zGyr Median frequency, zMag Max, xGyr Spectral centroid, yAcc Median frequency, xAcc Spectral entropy, yMag Max, yGyr Spectral entropy, xAcc wear Median frequency
P012	yGyr Standard deviation, yMag wear Spectral entropy, xGyr Median frequency, xAcc wear Median frequency, yMag Max, xMag wear Mean, yAcc Standard deviation, yAcc wear Max, yAcc Median frequency, zMag Max, zAcc wear Max
P013	zAcc Standard deviation, yAcc Spectral entropy, xMag Spectral centroid, yGyr Standard deviation, yGyr Median frequency, zMag wear Max, xMag Mean, zMag Spectral centroid, yAcc wear Spectral entropy, xAcc wear Spectral entropy, zGyr wear Median frequency, xMag wear Max, yMag wear Max
P014	yGyr wear Median frequency, xGyr wear Max, xAcc Standard deviation, yGyr Interquartile range, xGyr Spectral entropy, xAcc Spectral entropy, yAcc Max, zAcc wear Spectral entropy, xGyr wear Spectral entropy, xMag wear Max, xMag Spectral centroid, xMag Min, yMag Min, yAcc wear Spectral entropy, yMag wear Mean
P015	zAcc Spectral entropy, zMag Min, yGyr Standard deviation, zMag wear Spectral entropy, zAcc wear Median frequency, xMag Max, zGyr Spectral entropy, zAcc Median frequency, zAcc Spectral centroid, xMag wear Mean, yAcc Variance
P016	zAcc Spectral entropy, xAcc Spectral entropy, zAcc wear Spectral entropy, xGyr Variance, yGyr wear Spectral centroid, xMag wear Min, yMag Max, zMag Min, yMag wear Min, yAcc Variance, zGyr wear Min, xGyr Standard deviation

Continued on next page

Table II.18 SSMs feature sets, GMM modality smartphone and -watch for all activities

Subject	Feature set
P018	xMag wear Mean, xGyr Median frequency, zGyr Variance, zMag Mean, yAcc Interquartile range, zMag wear Max, zAcc Median frequency, yMag wear Mean, xMag Median frequency, yAcc Max
P019	xGyr wear Spectral entropy, zGyr Min, xAcc wear Median frequency, xAcc Median frequency, zGyr Interquartile range, zMag Spectral entropy, yMag Max, xGyr Spectral entropy, zGyr wear Spectral entropy, xAcc Spectral centroid, xMag wear Max, zAcc wear Median frequency, zMag wear Min, yMag wear Mean
P020	yMag wear Min, yAcc Max, yAcc Spectral centroid, yMag Mean, xMag Max, yAcc Interquartile range, yGyr Spectral entropy, yMag wear Spectral centroid, xAcc wear Median frequency, yAcc Median frequency

Legend: **SSMs**- Subject specific models; **wear**- watch.

II.2 Feature sets for the one-stage general model (1GM)

II.2.1 Basic activities

Table II.19: 1GM feature sets for KMeans, three sensor modalities, and basic activities

Modality	Feature set
M_P	zAcc Variance, xAcc Interquartile range, yAcc Min, yGyr Standard deviation, zGyr Spectral entropy, yMag Min, yGyr Spectral entropy, xAcc Median frequency, xGyr Standard deviation, yMag Mean, xMag Spectral centroid, yMag Max
M_W	yAcc wear Median frequency, zGyr wear Spectral entropy, yMag wear Spectral centroid, zGyr wear Standard deviation, xMag wear Max, yGyr wear Spectral entropy, zMag wear Mean, xMag wear Min, xGyr wear Min, zGyr wear Max, zGyr wear Min
M_{P+W}	zGyr wear Max, yAcc Max, zGyr Standard deviation, xMag wear Min, zGyr Spectral entropy, xMag Spectral centroid, zMag wear Mean, yGyr Median frequency, yGyr wear Spectral entropy

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **wear**- watch;
1GM- One-stage general model

Table II.20: 1GM feature sets for AGG, three sensor modalities, and basic activities

Modality	Feature set

M_P	yGyr Min, zMag Spectral entropy, yAcc Standard deviation, yAcc Variance
M_W	zGyr wear Interquartile range, xGyr wear Spectral entropy
M_{P+W}	xMag wear Spectral entropy, xAcc Standard deviation, zAcc Median frequency, zGyr Variance, yAcc Standard deviation, yAcc Variance, xMag Spectral centroid

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **wear**- watch;
1GM- One-stage general model

Table II.21: 1GM feature sets for GMM, three sensor modalities, and basic activities

Modality	Feature set
M_P	zMag Spectral entropy, yAcc Spectral entropy, yGyr Standard deviation, yAcc Min, yAcc Standard deviation, xMag Spectral centroid, yAcc Variance, xMag Max
M_W	zMag wear Max, zGyr wear Max, yMag_wear_Spectral entropy
M_{P+W}	yAcc Max, zAcc wear Spectral centroid, xMag wear Mean, zGyr Spectral entropy

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **wear**- watch;
1GM- One-stage general model

II.2.2 All activities

Table II.22: 1GM feature sets for KMeans, all activities, and sensor modalities.

Modality	Feature set
M_P	zMag Mean, zAcc Median frequency, xMag Mean, xGyr Spectral centroid, yGyr Standard deviation, zGyr Standard deviation, xAcc Max, zGyr Spectral entropy, yAcc Spectral entropy, xGyr Standard deviation, xGyr Min, yGyr Spectral entropy, xAcc Median frequency, xMag Min, xMag Spectral centroid
M_W	yMag wear Spectral entropy, zGyr wear Min, yGyr wear Spectral centroid, zGyr wear Spectral entropy, zGyr wear Standard deviation, xMag wear Spectral centroid, xMag wear Spectral entropy
M_{P+W}	zMag Min, zGyr wear Spectral entropy, zGyr Max, zGyr wear Min, zGyr Interquartile range, yGyr Spectral entropy, xMag wear Spectral entropy, yAcc Spectral entropy, xAcc Min, zAcc Median frequency, zGyr Median frequency, xAcc Median frequency, xMag Spectral centroid, xMag Min, xAcc Standard deviation, zMag Spectral centroid

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **wear**- watch;
1GM- One-stage general model

Table II.23: 1GM feature sets for AGG, all activities, and sensor modalities.

Modality	Feature set
M_P	xGyr Standard deviation, yAcc Interquartile range, zMag Max, xGyr Spectral entropy, yGyr Min, xMag Spectral centroid
M_W	zGyr wear Min, xGyr wear Spectral entropy, yAcc wear Median frequency
M_{P+W}	yMag Spectral entropy, zGyr Spectral entropy, xMag Mean, zGyr Interquartile range, zAcc Median frequency, xAcc Median frequency, yAcc Spectral entropy, xAcc Spectral centroid

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **wear**- watch;
1GM- One-stage general model

Table II.24: 1GM feature sets for GMM, all activities, and sensor modalities.

Modality	Feature set
M_P	yGyr Standard deviation, zGyr Spectral entropy, xAcc Median frequency, xAcc Spectral entropy, xGyr Standard deviation, yAcc Interquartile range, xMag Spectral centroid
M_W	zGyr wear Max, yMag wear Spectral entropy, yGyr wear Spectral centroid
M_{P+W}	yMag wear Spectral entropy, zMag wear Spectral entropy, yAcc Median frequency, zGyr Median frequency, yGyr Standard deviation, yAcc wear Spectral centroid, zAcc wear Spectral entropy, xAcc Median frequency, zAcc Median frequency, yAcc wear Median frequency, xMag Max

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **wear**- watch;
1GM- One-stage general model

II.3 Feature sets for the two-stage general model (2GM)

II.3.1 Basic activities

Table II.25: 2GM feature sets for KMeans, three sensor modalities, and basic activities.

Modality	Feature set
M_P	xMag Max, yAcc Interquartile range, zMag Max, xAcc Min, yAcc Min
M_W	xMag wear Max, zGyr wear Interquartile range, yMag wear Min, xMag wear Min, zGyr wear Max
M_{P+W}	xMag Max, xMag wear Max, yAcc Max, yAcc Interquartile range, zMag Max

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **wear**- watch;
2GM- Two-stage general model

Table II.26: 2GM feature sets for AGG, three sensor modalities, and basic activities.

Modality	Feature set
M_P	yAcc Max, zAcc Interquartile range, zMag Max, yMag Max
M_W	xMag wear Max, xMag wear Min, xMag wear Mean, zGyr wear Standard deviation, xGyr wear Interquartile range
M_{P+W}	xMag wear Max, xGyr Standard deviation, yAcc Interquartile range, zMag Max, zGyr Interquartile range, yMag wear Max, yAcc Min

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **wear**- watch;
2GM- Two-stage general model

Table II.27: 2GM feature sets for GMM, three sensor modalities, and basic activities

Modality	Feature set
M_P	xMag Max, yAcc Max, zMag Max, yAcc Min, yGyr Standard deviation
M_W	xMag wear Max, yMag wear Min, xMag wear Min, zGyr wear Standard deviation, zGyr wear Interquartile range, zMag wear Max
M_{P+W}	xMag Max, zMag Max, xMag wear Max, yAcc Min, yAcc Interquartile range, yMag wear Min, xMag wear Min

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **wear**- watch;
2GM- Two-stage general model

II.3.2 All activities

Table II.28: 2GM feature sets for KMeans, all activities, and sensor modalities.

Modality	Feature set
M_P	zMag Max, xMag Max, yGyr Spectral entropy, zGyr Standard deviation, yAcc Max, yAcc Interquartile range, yGyr Standard deviation, zGyr Spectral entropy, xGyr Min, zMag Min
M_W	xMag wear Max, yGyr wear Spectral entropy, xMag wear Spectral entropy, xMag wear Mean, yMag wear Mean, xMag wear Min, xAcc wear Standard deviation, xGyr wear Interquartile range
M_{P+W}	xAcc Standard deviation, yAcc Max, yGyr Standard deviation, zMag Max, xMag wear Min, xMag wear Max, yMag wear Spectral centroid, xMag Max, zAcc Standard deviation, zMag Mean

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **wear**- watch;
2GM- Two-stage general model

Table II.29: 2GM feature sets for AGG, all activities, and sensor modalities.

Modality	Feature set
M_P	xMag Max, zGyr Spectral entropy, yGyr Standard deviation, zAcc Median frequency, zMag Max, yGyr Spectral entropy, yAcc Max, zGyr Standard deviation, xMag Spectral centroid, xGyr Min, zGyr Min, zMag Min, yMag Max, yGyr Max
M_W	xMag wear Max, xAcc wear Standard deviation, zAcc wear Max, xGyr wear Spectral entropy, xGyr wear Interquartile range, zGyr wear Interquartile range, yMag wear Min, xMag wear Mean
M_{P+W}	xMag wear Mean, zMag Mean, yAcc Median frequency, zGyr Max, yAcc wear Standard deviation, xMag Mean, xGyr Standard deviation, zMag wear Min

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **wear**- watch;
2GM- Two-stage general model

Table II.30: 2GM feature sets for GMM, all activities, and sensor modalities

Modality	Feature set
M_P	zMag Max, yAcc Variance, yGyr Median frequency, xMag Mean, yGyr Spectral entropy, xMag Spectral centroid, xMag Min, yAcc Min, yAcc Max, zGyr Median frequency
M_W	zGyr wear Interquartile range, xMag wear Max, xMag wear Mean, xAcc wear Median frequency, zMag wear Mean, yAcc wear Median frequency, yAcc wear Spectral entropy, zAcc wear Spectral entropy
M_{P+W}	xMag Mean, xMag wear Min, xMag wear Mean, xMag Max, zGyr Spectral entropy, yMag Max, yAcc Median frequency, zMag Mean, yAcc Standard deviation, yAcc Variance, yAcc Spectral entropy, xGyr Median frequency, yAcc Interquartile range, xGyr wear Median frequency

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch; **wear**- watch;
2GM- Two-stage general model

III

Random Forest Grid Search Results

Table III.1: Random Forest (RF) grid search results (number of trees and depth) using the one-stage general model datasets.

Activities	M_P		M_{P+W}	
	nr. trees	depth	nr. trees	depth
Basic	200	20	200	20
All	200	10	300	10

Legend: **M**- Modality; **P**- Smartphone; **W**- Smartwatch.

IV

Decision Tree for Interpretable Clustering

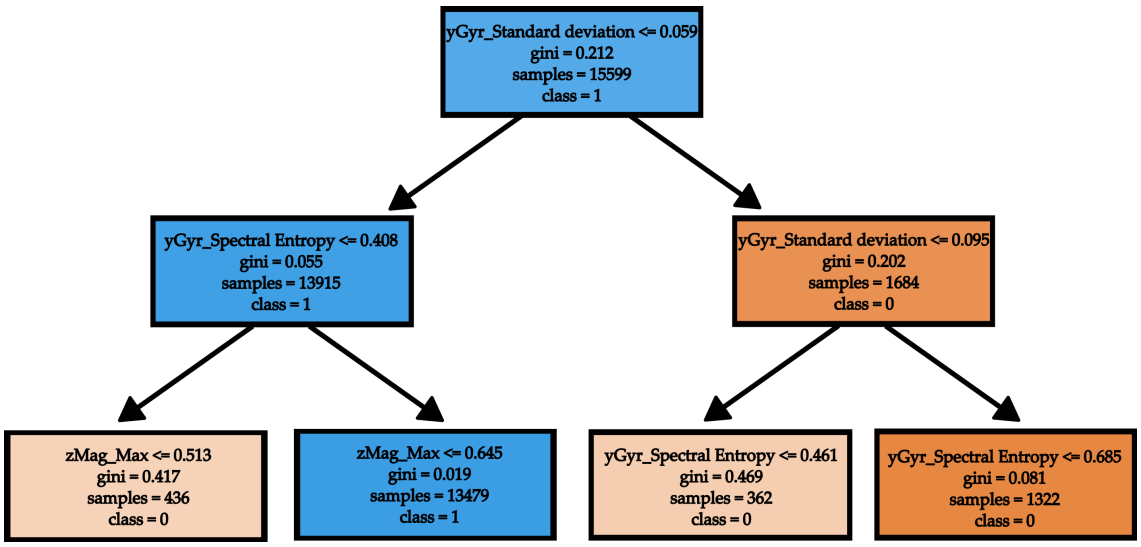


Figure IV.1: Decision tree, trained with the cluster labels as target class labels, up to a depth of two.