ESTIMATING HUMAN ACTIVITY PATTERNS IN DYNAMIC ENVIRONMENTS BASED ON SMART, WEARABLE SENSORS: A FEASIBILITY STUDY

by

CHARIKLEIA P.CHATZAKI

BSc. Technological Educational Institute of Crete, 2011

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

DEPARTMENT OF INFORMATICS ENGINNERING
SCHOOL OF ENGINEERING
TECHNOLOGICAL EDUCATIONAL INSTITUTE OF CRETE

MSC IN INFORMATICS & MULTIMEDIA

Heraklion
2016

Approved by:

Major Professor
Associate Professor Manolis Tsiknakis

Copyright

CHARIKLEIA CHATZAKI

2016

Abstract

The demands for understanding human activities have steadily grown in recent years in the health-care domain, especially in elder care support, rehabilitation assistance, diabetes, cognitive disorders, assisting living and wellness management. A significant amount of resources can be saved if sensors can help caretakers record and monitor elders or patients continuously and report automatically when any abnormal behavior is detected. The recognition of various activities of daily living (ADLs) can reveal valuable information about a person's activity patterns.

Many studies have successfully identified activities using wearable sensors with very low error rate, but the majority of the previous works are done in very constrained settings. Readings from multiple body-attached sensors achieve low error-rate, but the complicated setting is not feasible in practice. On the other hand, smartphones have been accepted from the research community as a powerful solution for sensing applications due to the increasing number of smartphone users and due to the vast capabilities of modern smartphones. This project uses low-cost and commercially available smartphones as sensors to identify human activities. The growing popularity and computational power of smartphone make it an ideal candidate for non-intrusive body-attached sensors. Unlike many previous reported works, we relaxed the constraints of attaching sensors to fixed body positions with fixed device orientation. In our design, the phone can be placed at any position around waist such as jacket pocket and pants pocket, with arbitrary orientation.

In this work a feasibility study has been contacted to investigate whether a smartphone based recognition system can be used for estimating activity patterns in dynamic environments. To this end, different combinations of computational approaches have been taken. The computational pipeline was applied on separate activities of daily living and on complete sequences of activities, which describe a common scenario of daily living. The results showed that, using a 1 second-window with 80% overlap, the suggested feature sets and the k-NN classifier, the ADLs and the scenarios can be recognized with accuracy of 99% and 96 % respectively, when the 10-fold cross-validation evaluation method is applied. A further investigation using the aforementioned combinations and the evaluation method of Leave-One-subject-Out for the recognition of scenarios achieved accuracy of 79%.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to express my sincere gratitude to my supervisor Prof. Manolis Tsiknakis for the continuous support, motivation, engagement and for his patience, during the development of this thesis. I could not have imagined having a better supervisor for my master thesis.

My sincere thanks also go to my collaborators, Dr. Matthew Pediaditis and George Vavoulas for their continuous support and guidance through the learning process of this master thesis.

Finally, I would also like to thank the all volunteers for their contribution in the generation of the dataset.

Dedication

Dedicated to my beloved family and to the love of my life…….

1. Introduction

Human activity recognition aims to identify and recognize the activities of a person over an observed period of time and via a series of actions. A high number of scientific studies advocating the emerging task of activity recognition and modeling have been published. The first step in developing an activity recognition system is the sensing of the activities. The involved sensors play an important role in the performance of the recognition system. Furthermore, human behavior modelling, a domain which includes the recognition of human activity patterns is attracting high interest from the research community. Human behavior can be defined as the set of mental, physical and social activities experienced during the phases of human life. The activities of daily living of a human can reveal valuable information about his activity patterns and subsequently, with the association of other behavior characterizing elements, they can lead to the development of behavioral models.

The main objective of this study is to investigate the techniques and the processing steps that can be used to develop a recognition system for human activity patterns in dynamically changing environments. To achieve human behavior modelling or estimation of activity patterns it is essential to first recognize the activities of daily living effectively and accurately. In this study, techniques and methods for activity recognition are deployed and compared. The findings are applied to scenarios of activities of daily living, within the boundaries of real life conditions.

In the context of this thesis we will focus on the following: 1) Implementation of a thorough literature review; 2) Critical evaluation of existing approaches and methods in relation to recognition of complex activities; 3) Generation and annotation of an experimental dataset, including a range of elementary and complex human daily activities; 4) Implementation of a computational pipeline for recognition of complex human behaviors in dynamic environments, exploring the above dataset. In the process a variety of feature extraction, selection and classification methods will be tested.

This study is structured as follows. Section 2 presents a literature review of the most commonly used approaches for activity recognition and behavioral modeling. The methodology that is deployed in such systems is analyzed in Section 3. Section 4 presents the developed computational pipeline and Section 5 presents the respective results. Finally, the conclusions and suggestions of further work are reported in Section 6.

1.1.Human Activities

Humans who are physically active and healthy can perform various activities. Thus, it is essential to categorize those activities into sets in order to develop a human activity recognition system. In general, there are two main classes of human activities, the activities of daily living and the instrumental activities of daily living.

Activities of daily living (ADLs) are basic- fundamental tasks that people perform in their daily routine without needing assistance. The ability or inability to perform ADLs can be used as a very practical measure of ability/disability in many disorders. Based on the most common measure of functional ability namely, *the Katz Activities of Daily Living Scale*, the most representative ADLs include: bathing, dressing, transferring, using the toilet, continence, and eating [1]. Over the years, some additional measures of mobility have been included, such as walking, getting around inside, and getting around outside. On the other hand, instrumental activities of daily living (IADLs) are more complex tasks that require a certain amount of physical dexterity, sound judgment and organizational skills. The ability or inability to perform IADLs can be used as a measure of a person's ability/disability to live safely and independently [1]. Common paradigms of IADLs include: handling personal finances, meal preparation, shopping, traveling, doing housework, using the telephone, and taking medications.

Focusing on human activity recognition systems, Reyes-Ortiz [2] categorized the activities based on two classification schemes, a) with respect to their duration and complexity and b) with respect to the type of activity, as shown in Table 1.1 and Table 1.2 respectively.

Table 1.1 Classification of activities based on their duration and complexity, fully adopted by Reyes-Ortiz [2]

Duration/complexity	**Activity type**	**Examples**
Short events	Gestures	Waving hands, nodding head, laughing
	Transitions	Stand-to-sit, lie-to-sit
Basic activities	Static	Standing, sitting, reading
	Dynamic	Walking, running, cycling
Complex activities	Multi-activity	Cooking, assembling furniture, weight training
	Multi-user	Talking, ballroom dancing, hugging

In the framework of this thesis the recognition of ADLs/Short events and basic activities will to be investigated, based on information obtained from smart, wearable sensors. The rational for this is the fact that that these activities can be defined as prerequisite tasks in the "hierarchical" order of the most complex. For example, if a human performs a complex activity such as cooking this could include a sequence of basic activities with short events such as, walking, standing, moving hands etc.

Table 1.2 Classification of activities based on their type, fully adopted by Reyes-Ortiz [2]

Application	Examples
Daily living	Waving hands, nodding head, laughing
Locomotion	Walking, riding, standing, laying down, falling
Sports/fitness	Jumping, weight lifting, climbing, swimming
Communication/connectivity	Phone calling, texting, talking, signing
Complex activities (CAs)	Cooking, assembling furniture, weight training
Security/surveillance	Loitering, chasing, supervising, stalking

1.2. Input Sensors

As the first step in the deployment of an activity recognition system is the sensing of the activities, the choice of the appropriate sensors plays a crucial role in the effectiveness and performance of the system [3]. With respect to sensor placement and interaction with the user there are two main categories of sensors: the ambient sensors (also called external or environmental) and the wearable sensors [2]. Based on the remarkable growth of smartphone users, latest research focuses on the use of smartphone devices as a wearable sensor for activity recognition [4], [5].

1.2.1. Ambient sensors

Ambient sensors are sensors capable of jointly sensing multiple physical phenomena in various surroundings [6]. The use of ambient sensing techniques is commonly used in the context of smart spaces, homes and buildings. In accordance with the desired outcome, there is a wide range of sensors which can be used such as microphones, video cameras, presence sensors, radio frequency identification (RFID) tags, thermometers and depth sensors [2].

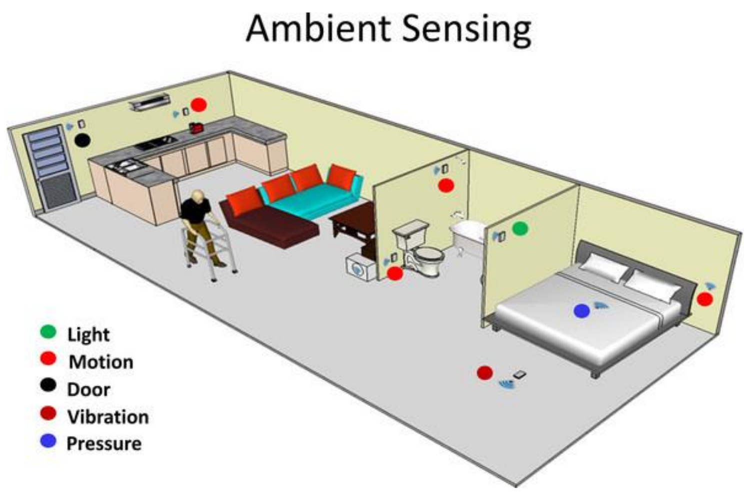

Figure 1.1 An example of the usage of ambient sensors in a smart home environment [7]

1.2.2. Wearable sensors

Wearable sensors are attached to different places of the person's body such as waist, chest, legs, and hands, depending on the relevant attributes of interest to be measured [2]. More recently, such sensors are also fitted to clothes or embedded in accessories [8], [9] . Depending on the application domain, wearable systems may require the development of a particular design and location, i.e. wrist bracelets [7]. One of the challenges using wearable sensors is the controlled affix position that must be chosen without hindering body movements [6] or without "loosing" body movements in case of activity recognition.

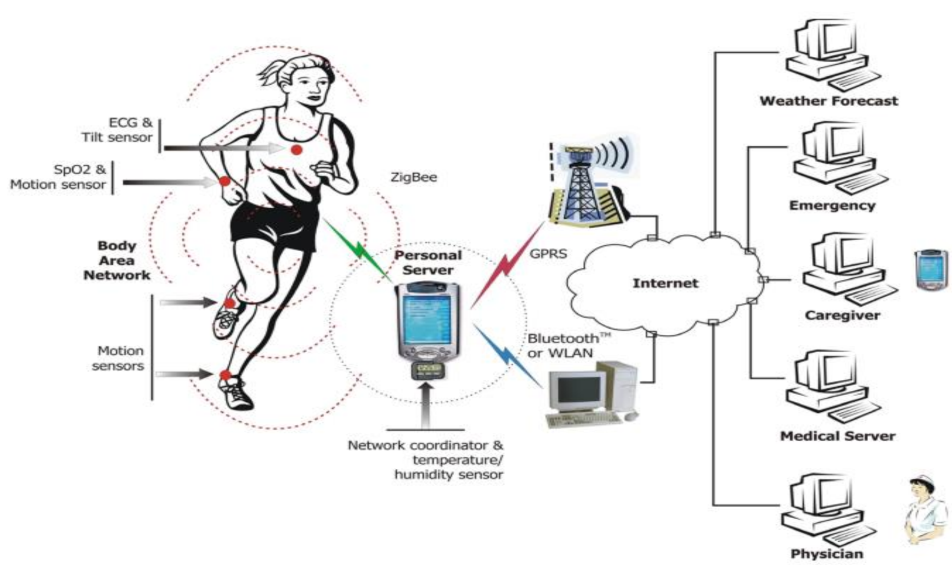

Figure 1.2 Multiple wearable sensors for monitoring user [8]

1.2.3. *Smartphones*

Smartphones have been accepted from the research community as a powerful solution for sensing applications since it encompass a number of advantages that can overcome existing issues in the field [2], [5]. Smartphone based sensing systems have been developed in the area of health monitoring, environmental monitoring, traffic monitoring, human behavioral monitoring and social networking [10]. The tremendous number of smartphone users, the standardized-universal architecture, the wireless network, the size of the device and finally the variety of the embedded sensors [5] makes the use of smartphone a valuable sensing device for activity recognition. There are three broad categories of sensors included in the majority of smartphone devices namely, motion sensors, position sensors and environmental sensors.

Table 1.3 Overview of smartphone's embedded sensors as described in the official site of android developers guide[1]

Sensor	Type	Description	Category
Accelerometer	Hardware	Measures the acceleration force in m/s^2 that is applied to a device on all three physical axes (x, y, and z), including the force of gravity.	Motion
Linear acceleration	Software or hardware	Measures the acceleration force in m/s^2 that is applied to a device on all three physical axes (x, y, and z), excluding the force of gravity.	Motion
Gyroscope	Hardware	Measures a device's rate of rotation in rad/s around each of the three physical axes (x, y, and z).	Motion
Gravity	Software or hardware	Measures the force of gravity in m/s^2 that is applied to a device on all three physical axes (x, y, z).	Motion
Rotation vector	Software or	Measures the orientation of a device by	Motion,

[1] http://developer.android.com/guide/topics/sensors/sensors_overview.html

	hardware	providing the three elements of the device's rotation vector.	position
Orientation	Software	Measures degrees of rotation that a device makes around all three physical axes (x, y, z). The inclination matrix and rotation matrix can be obtained for a device by using the gravity sensor and the geomagnetic field sensor	Position
Magnetic field	Hardware	Measures the ambient geomagnetic field for all three physical axes (x, y, z) in μT.	Position
Proximity	Hardware	Measures the proximity of an object in cm relative to the view screen of a device. This sensor is typically used to determine whether a handset is being held up to a person's ear.	Position
Pressure	Hardware	Measures the ambient air pressure in hPa or mbar.	Environment
Light	Hardware	Measures the ambient light level (illumination) in lx.	Environment
Ambient temperature	Hardware	Measures the ambient room temperature in degrees Celsius (°C)	Environment
Relative humidity	Hardware	Measures the relative ambient humidity in percent (%)	Environment

1.2.3.1. Accelerometer

The accelerometer measures the acceleration force that is applied to a device on all three physical axes (x, y, and z), including the force of gravity, in other words the proper acceleration ("g-force") [11]. The aforementioned forces could be static, for instance due to the constant gravitational force, or they could be dynamic, caused from movements or vibrations. The measurement of static acceleration due to gravity can be used for the estimation of device's tilt angle with respect to the earth. Moreover by analyzing the dynamic acceleration forces, information about the device’s motion patterns can be obtained. Hence an accelerometer can be used for measuring the changes in velocity and the changes in position.

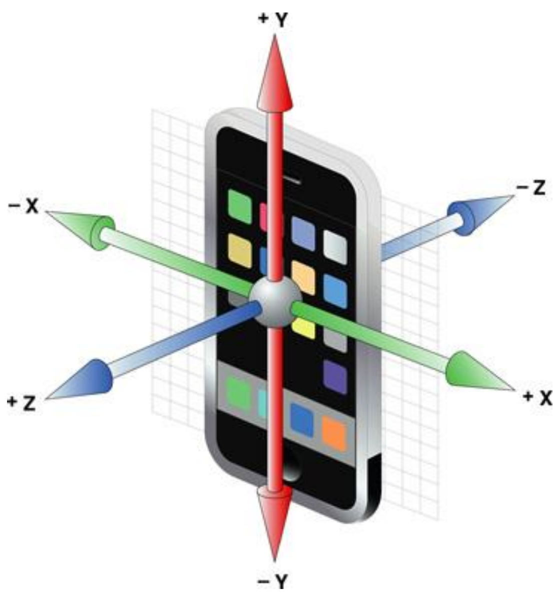

Figure 1.3 Localization of the three physical axes in a smartphone device (figure taken from Apple Inc.)

1.2.3.2. Gyroscope

The Gyroscope effect is based on the *Coriolis* effect and measures the rate of rotation around a particular axis [5]. Consequently, a gyroscope gives an indication of the angular rate. Thus a gyroscope can be used for measuring the changes in orientation and the changes in rotational velocity.

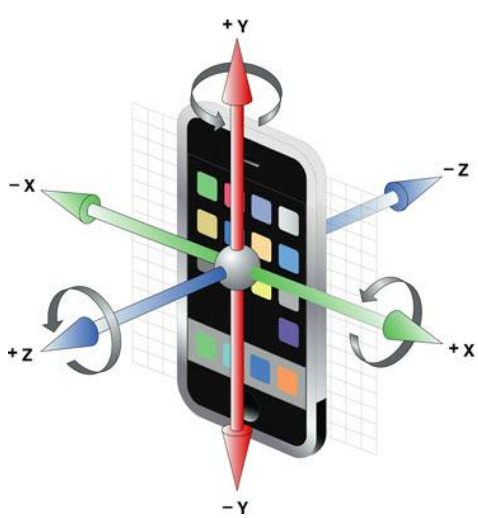

Figure 1.4 Gyroscope effect in a smartphone device (figure taken from Apple Inc.)

2. Literature review

2.1. Activity recognition

Human activity recognition is the process of identifying and recognizing the activities and goals of one or more humans from an observed series of actions [4]. A number of studies have been published advocating the emerge task of activity recognition, using several approaches based on vision sensors, inertial sensors, smartphone sensors and a combination of the aforementioned. Activity recognition can be applied in several domains such as surveillance, human-computer interaction and proactive computing [3] [12]. The scope of this study is to focus in activity recognition based in smartphone sensors, as a cost effective and power solution to monitor motion data in real life scenario [13]. Nevertheless, a brief review for comparison purposes has been contracted for activity recognition approaches using external sensors.

2.1.1. Recognition of Activities of Daily Living using Smartphones

A recognition system based on the accelerometer sensor of a smartphone for simple (biking, stairs up, driving, lying, running, sitting, standing and walking) and complex (cooking, cleaning etc.) activities was reported in [14]. Ten participants performed the activities while it was left at each user's will to decide about the placement of the smartphone, in terms of position and orientation. For the simple activities the user was also able to choose the action, the location, and the duration (manual start and stop) while for the complex activities these parameters were predefined. The sampling frequency was set at 80 Hz maximum although variations in the sampling rate were reported. Multiple windows sizes of 1, 2, 4, 8 and 16 seconds with 50% overlap were used. Although complex activities were classified with an accuracy of 50%, simple activities were classified with 93% accuracy with a Multilayer Perceptron and a window size of 2 seconds.

Siirtola and Roning [15], used accelerometer data from a smartphone with a sampling frequency of 40Hz while seven volunteers were performing five different activities: walking, running, cycling, driving a car, and sitting/standing. For each recording of the above activities, four smartphones were placed in different positions: trousers' front pocket, jacket's pocket, at backpack, at brachium and at the ear only when it was physically allowed. For the feature extraction process a sliding window of 7.5 seconds with 25% overlap in an online application and one with 50% overlap in an offline application, were used. Classification was achieved using five classifiers based on quadratic discriminant analysis arranged in a three stage decision tree topology. Average recognition rate of almost 98.9% was reported in the offline and 90% in the online system.

A smartphone-based recognition system is proposed in [16], in which the use of a low-pass filter and a combination of Multilayer Perceptron, LogitBoost and Support Vector Machine (SVM) classifiers produced an accuracy of 91.15%. Four volunteers performed six activities: slow and fast running, walking, aerobic dance, stairs up and stairs down. Samples were recorded continuously with the smartphone held in the hand of the user. In order to discriminate properly the collected data, volunteers were instructed to stop and wait a few seconds every time the willed to change an activity. The sampling rate was set at 100 Hz while a window of 1.28 seconds with 50% overlap was used for feature extraction.

A similar approach was introduced in [17] by Anjum and Ilyas, in which ten users performed seven different activities namely: walking, running, stairs up, stairs down, cycling, driving and remaining inactive, by carrying the smartphone in various positions. The collection of the data was performed for each activity separately and each user manually started and stopped the recoding. A sampling rate of 15 Hz and time window of 5 seconds were used. Based on the ranking of the information gain, nine features were selected from the auto correlation function. For the classification process Naïve Bayes, C4.5 Decision Tree, K-Nearest Neighbor and SVM classifiers were tested. The C4.5 Decision Tree performed better than the other classifiers with an accuracy of 95.2%.

Based on the accelerometer data of a smartphone Fan et al. [18] studied three different decision tree models based on: a) the activity performed by the user and the position of the smartphone (vector: activity-position), b) only the position and c) only the activity. Fifteen users performed five kinds of activities: stationary, walking, running, stairs up and stairs down with the smartphone placed into a carrying bag, a trouser pocket or in the hand. Ten-second samples of accelerometer data were recorded for each different kind of activity and position of the smartphone. The authors concluded that the model based only on the activity outperformed the other two with an accuracy of 88.32%.

Buber and Guvensan [19] developed a recognition system based on tri-axial accelerometer data of a smartphone's sensor. Eight activities were tested for this purpose, namely: walking, jogging, jumping, stairs up, stairs down, sitting, standing and biking. Five volunteers performed those activities with the smartphone placed in the front pocket of their trousers in a fixed orientation. The recordings were started and stopped by the user for each different activity. The sampling rate was set at 20 Hz and a 10 second moving window was used for feature extraction. The evaluation was performed with two feature selection algorithms (OneRAttributeEval and ReliefFAttributeEval) and six classification algorithms (J48, K-Star, Bayes Net, Naïve Bayes, Random Forest, and k-NN) using 10-fold cross-validation. The authors resulted in a combination of 15 features with k-NN to perform best at a recognition rate of 94%.

Using the accelerometer sensor of a smartphone Zheng et al. [20] proposed a two-phase method to achieve activity recognition. Five subjects performed four different types of activities (sitting, standing, walking and running) with the phone placed loosely in a pocket. For the training phase, two minute recordings were used, which included all the described activities, while for the testing phase, data from continuous records of several days were used. A sampling rate of 100 Hz was used. Noise reduction was achieved deploying Independent Components Analysis, specifically the fastICA algorithm, in combination with the wavelet transform for feature extraction. For the classification, a Support Vector Machine was employed using the WEKA toolkit. A maximum accuracy of 98.78% was reported for a leave-one-out validation.

Saputri et al. [21] proposed a system for activity recognition, which with the use of an Artificial Neural Network produced 93% accuracy. Twenty-seven subjects performed six types of activities, namely, walking, jogging, running, stairs up, stairs down and hopping. The smartphone was placed in the front trouser pocket, not clearly defined if the orientation was fixed, using a sampling rate of 50 Hz. In the feature extraction process, the window size was set at 2 seconds, while feature selection was performed using a self-devised three-staged genetic algorithm.

An activity recognition system along with the functionality of getting user's feedback for a refinement of the prediction model is proposed by Su et al. [22] Acceleration and compass (for the orientation) data of a smartphone were used to recognize different types of activities including: walking, sitting, stairs up and down. The position of the phone and the exact data collection method were not described in detail. A sampling frequency of 64 Hz was used, while time and frequency domain features were extracted. A comparison study with the reported results from Kwapisz et al. [23] (the comparison process is not reported clearly) was conducted and resulted in 98.7% accuracy with the use of a Multilayer Perceptron. Hidden Markov model (HMM) was applied on the top of the classification model for identifying temporal dependencies of the activities.

Using the accelerometer, gyroscope and proximity sensors along with the GPS module of a smartphone Han et al. [24] proposed a hierarchical activity recognition framework which extends the Naïve Bayes approach. Based on four different physical movements (walking, sitting, jogging, and standing) they produced ten different activity models according to predefined location points of the GPS namely, home, office and outdoor. Furthermore, when the undertaken activity was identified as outdoor, four activity models (Waiting for bus at bus stop; Having a meal at cafeteria; Exercising at gym; Visiting a park) were tested based only on the location (GPS) and one (driving a car) based on the moving speed of the user. Details on the collection data process, smartphone position and type of extracted features were not mentioned. The reported average classification accuracy was 92.96%.

Another activity recognition system based on smartphone sensors is proposed by Hung et al. [25] using an open dataset [26], [27], which includes six activities (Standing, Walking, Running, Upstairs, Downstairs, Laying) performed by thirty volunteers with the smartphone positioned at the waist. In the referred dataset, data was collected with a sampling rate of 50 Hz and pre-processing included a sliding window of 2.56 sec in duration with 50% overlap. Forty-five features were extracted and three different classifies were tested, namely, Decision Tree (J48), Support Vector Machine and Logistic Regression, with the last one outperforming the others with an overall accuracy of 96%.

With the use of kernel discriminant analysis for the feature set and ANNs based on the feed-forward backpropagation algorithm Kahn et al. [28] reached an accuracy of 96% for activity recognition. Six volunteers performed five activities (sitting, walking, stairs up, stairs down and running) with the smartphone placed in five different positions, namely, the shirt's top pocket, the jeans' front-left pocket, the jeans' front-right pocket and a jeans' rear pocket. A sampling rate of 45 Hz and a sliding window of two seconds without overlapping were applied.

Sang et al. [29] proposed an approach for recognizing activities of daily living using the accelerometer and gyroscope sensors' data of a smartphone. Five different activities were tested with the smartphone placed on user’s pocket, namely: driving a motorbike, stairs up and down, sitting and putting the smartphone on a table. Authors have not reported details for the data collection protocol. Classification process included the test of kNN and ANN with overall accuracy of 74% and 75.3% respectively.

Table 2.1 Overview of the methodology and results followed by the related studies which make use smartphone's sensors

Study	No of subjects	Activities[1]	Recordings	Sampling Frequency	Window size/overlap	No of Features	Smartphone position	Algorithms[2]	Performance
[14]	10	BIK, STU, DRI, LAY, RUN, SIT, STD WAL.	Simple activities: user starts & stops recs Complex: predefined duration	80Hz	1,2,4,8,16/ 50%	6	user's choice (position & orientation)	MLP, NB, BN, DT, B-FT, K-star	MLP: 93% 2s window
[15]	7	WAL, RUN, BIK, DRI, SIT/STD	-	40Hz	online: 7.5s/25% offline: 7.5s/50%	76	5 smartphones : various position	DC & QDA	90% online 98.9% offline
[16]	4	RUN, SWL FWL, ADN, STU, STN	User stops and wait inactive between activities	100Hz	1.28s/50%	18	hand of the user	J48, K-Star, BN, NB, RF, kNN	MLP & LB &SVM: 91,15% Accuracy
[17]	10	RUN, STN, STU, BIK, STC, DRI, INA	User starts & stops recs	15Hz	5s	9*	various positions	NB, C4.5, KNN, SVM	C4.5: 95.2%.
[18]	15	STC, WAL, RUN, STU,STN	-	-	10s	10*	bag, trouser pocket & hands	ID3 DC	80.29%
[19]	5	WAL, JOG, STN, STU, SIT, JUM, BIK	User starts & stops recs	20Hz	10s	15	front pocket fixed orientation	J48, K-Star, BN, NB, RF, kNN	k-NN: 94%
[20]	5	SIT, STD, WAL, RUN	Training: continuous two minute recordings, Testing:	100Hz	-	feature vector	freely in pocket	SVM	98.78%

			continuous records of several days						
[21]	27	WAL, RUN, STN,STU, HOP	-	50Hz	2s	21	front pocket	ANN	93%
[22]	4	8 activities (not all specified) WAL, STU, STN,SIT	Testing: not specified Training: user feedback on three rated recognized activities	64Hz	4s	*8	-	HMM and J48, Logistic regression, MLP	MLP 98.7%
[24]	-	WAL, JOG, STD,SIT 15 models based on the location	-	50Hz	-	-	-	Adaptive Naïve Bayes	92.96%
[25]	30	STD, WAL, RUN, STN, STU, LAY	Dataset used: [26]	50Hz	2.56S/50%	45	Waist	J48,Logistic Regression, SVM	LR 96%
[28]	6	SIT,WAL, STU,STD, RUN	not -	45Hz	2 sec/no overlappin g	-	shirt's top pocket, jeans' front-left, right, rear pocket, coat's inner pocket	ANN based on the feed-forward backpropagati on algorithm	96%
[29]	-	SIT, DRVm, STU, STN, PPT	-	-	-	*4	On pocket	KNN, ANN	ANN 75.3%

* Feature set includes that number of features but is not limited to.

2.1.2. Recognition of Activities of Daily Living using on-body sensors

A comparison study of classification methods for human activity recognition based on acceleration data has been conducted by Mannini and Sabatini in [30]. A dataset described in detail in [31] was used, which includes signals from five on body accelerometers located at the hip, wrist, arm, ankle, and thigh of the user, while performing the following activities: sitting, walking, lying, stairs up, standing, running and cycling. A feature vector of thirty components was built from 50%-overlapping sliding windows of 6.7 seconds in duration. The following single frame classifiers were tested: Naive Bayesian (NB), Support Vector machine (SVM), Binary decision tree (C4.5), Gaussian Mixture Model (GMM), Nearest mean (NM), Logistic classifier, k-NN, Parzen classifier and ANN (multilayer perceptron) along with the cHMM-based sequential classification algorithm. The results indicated an accuracy of 98.5% for the Nearest mean and 98.4% for the cHMM classifier.

Guirya et al. [32] introduced an activity recognition system based on an accelerometer sensor mounted at the chest of the user and a smartphone placed in a pocket, acting as a supporting device. Twenty four volunteers performed the following activities: sitting, standing, lying, walking, walking on a treadmill at 5, 6, 8 km/h and cycling on an indoor bike. Authors tested a custom classifier and the following machine learning classifiers: Naïve Bayes, C4.5, CART, MLP and SVM. The best overall accuracy of 98% was reported with the use of C4.5 classifier.

An activity recognition system which utilizes body worn wireless accelerometers is proposed by Gaputa and Dallas in [33] . Three accelerometer sensors, mounted at the waist of the user, were used to obtain signals when two volunteers performed the following activities: walking, jumping, running, sit-to-stand/stand-to-sit, stand-to-kneel-to-stand, and being stationary (sitting or standing at one place). A sampling rate of 126Hz and a sliding window of 6 seconds with 50% overlap were deployed. Thirty-one features were selected using Relief-F, and Sequential Forward Floating Search (SFFS) while k-NN (10 neighbors) and Naïve-Bayes classifiers were utilized for the classification with overall accuracy of 98% and 95% respectively.

Table 2.2 Overview of the methodology and results followed by the related studies using various sensors

Study	**No of subjects**	**Activities[1]**	**Sampling Frequency**	**Window size/overlap**	**No of Features**	**Sensors**	**Algorithms[2]**	**Performance**
[30]	13	SIT, WAL, LYI, STU, STN, RUN, STD, CYC	76.25 Hz	6.7s/50%	30	5 bi-axial accelerometers, located at the hip, wrist, arm, ankle, and thigh.	NB, GMM, Logistic, Parzen, SVM, NM, k-NN, ANN, C4.5 & cHMM-based sequential classifier	NM 98.5% cHMM 98.4%
[32]	24	SIT, STD, LYI, WAL, WAL at 5, 6, 8 km/h, CYC indoor bike	120Hz	-	5*	Accelerometer sensor on chest & Smartphone on pocket (as server)	Naïve Bayes ,Custom, C4.5 ,CART, MLP, SVM	C4.5 98%
[33]	2	WAL, JUM, RUN, SIT, STD, sit-to-stand/stand-to-sit, stand-to-kneel-to-stand, one place)	126Hz	6s/50%	31	3 Accelerometers on waist.	k-NN, Naïve Bayes	

* Feature set includes that number of features but is not limited to

2.1.3. Public Datasets

The nature of the problem of activity recognition requires the collection of a large amount of data for training the classification system. To this end, researchers have created datasets in accordance with the processing approach. As already mentioned, the approaches include vision-based i.e. video recording of activities, or body sensors i.e. accelerometers on straps and smartphone sensors.

A dataset which uses a wide selection of sensors (accelerometers, video and audio sensors) to track and collect perceptual data from the user's perspective during the activity of cooking has been made available from the Carnegie Mellon University [34]. On the other hand, Chaquet et al. [35] focused on a vision based approach and have conducted a survey of available video datasets for human activity recognition. A different approach for collecting data describing a user's activities is introduced by Sarkar [36], which consists of a web-based activity data collection tool with a series of web-based interfaces allowing the user to configure and provide his activity experience. Abdallah et al. [37] introduced an adaptive model that learns incrementally from the evolving data streaming on an online recognition system. A comprehensive report on public datasets based on smartphone acceleration data is described below [2.1.3.1].

2.1.3.1. Public datasets based on smartphone's acceleration data

Kwapisz et al. [23] introduced an activity recognition platform based tri-axial accelerometer data of a smartphone. Twenty-nine users performed six different types of activities, namely: walking, jogging, ascending stairs, descending stairs, sitting, and standing in a predefined period of time. During the recordings the phone was placed into a user's pants pocket regardless of orientation. A widow of 10 seconds in size was applied to the data prior to feature extraction (43 features) for the classification process, where decision trees (J48), logistic regression and multilayer neural networks were tested. The system achieved an overall accuracy of 91.7% with the use of Multilayer Perceptron.

Using an IOS smartphone, McCall et al. [38] created a dataset for the recognition of daily activities. Ten subjects performed nine different activities (biking, running, climbing, standing, descending, treadmill walking, biking, walking and jump roping). For the recordings the smartphone was placed in the right side of user's belt. The sampling rate was set to 60 Hz. The overall classification accuracy using the hierarchical KNN model was 76%.

Miao et al. [39] used the built-in accelerometer, gyroscope, proximity sensor, light sensor and the magnetic sensor of a smartphone to collect data while seven volunteers performed the following activities: standing still/sitting on a sofa/sitting at a desk, walking, running, stairs up and down. For the recordings the smartphone was placed on six body positions without standard orientation, the positions were the two front and back pockets on the trousers and the two front pockets on the coat. A low pass filter with 0.25 Hz cutoff frequency was employed to separate acceleration due to gravity (GA) and linear acceleration (LA). Six features resulting from a sliding window of 1.6 seconds with 50 % overlap were computed. The Sequential Minimal Optimization (SMO), J48 και Naïve Bayes (NB) classifiers were tested with the last one to reach an overall accuracy of 89.6%.

Another dataset which includes six activities: standing, sitting, lying down, walking, stairs up and down has been created and tested in [27]. The accelerometer and gyroscope sensors of an android smartphone were used for signal acquisition with a sampling frequency of 50 Hz. Thirty volunteers performed two trials of the described activities. For the first trial the smartphone was placed in the left side of user's belt while for the second trial the position was left at the user's will. For feature extraction a moving window of 2.56 seconds with 50% overlap was used. A total of 561 features were extracted from the time and the frequency domain. A multiclass Support Vector Machine (SVM) classifier used for recognizing activities with an overall sensitivity 96%.

Data from the accelerometer, gyroscope and magnetic sensors of five smartphones placed on different body positions while ten users were performing seven activities were recorded in [40]. A sampling frequency of 50Hz was applied. The activities performed were walking, stairs up and down, running, sitting on a chair and standing. A moving window of 2 seconds with a 50% overlap was used for feature extraction. Four feature sets were created in the time and frequency domain for the evaluation. The following algorithms were tested: Bayes Network (BN), NB, LibSVM, LR, k-NN, PART, NNGE (rule-based classifier based on k-NN), RF and J48. An important outcome produced by the authors is that the use of the gyroscope together with the accelerometer is more efficient in most cases, although it is difficult to make a general remark because the role of each sensor depends on the location of the mobile.

Table 2.3 Overview of pubic datasets based on smartphone’s sensors

Study	No of subjects	Activities	Sensors used	Sampling Frequency	Window size/overlap	No of Features	Algorithms	Performance
Kwapsiz et al. 2011	29	WAL, JOG, STN, STU, SIT, STD	Acc	20Hz	10s/0%	43	J48, LR, MLP	MLP: 91,7% Accuracy
McCall 2012	10	BIK,EBK, STU, STN, JRO, RUN, STD, WAL, TWL	Acc	60Hz	8.33s/0%	105	hierarchical model with kNN and feature selection algorithms	76% Accuracy
Miao 2015	7	STC(STD/SIT), RUN, STN, STU	Acc, Gyro, Magn, proximity, light	25Hz	1.6s/50%	30	SMO, J48, NB	NB: 89,6% Accuracy
Anquita 2013	30	STD, SIT, LAY, WAL, STU, STN	Acc, Gyro	50Hz	2.56s/50%	561	Multi class SVM	96% Accuracy
Shoaib 2014	10	WAL, RUN, STN, STU, SIT, STD, BIK	Acc, Gyro, Magn	50Hz	2s/50%	36	BN, NB, LibSVM, LR, kNN, PART, NNGE, RF και J48	Evaluation of different sensors and phone positions for each activity

2.2.Behavioral modeling based on human physical activity patterns

Modeling human behavior is extremely useful in a variety of domains, such as surveillance-based security and ambient assistive living [3] , [12]. Human behavior can be defined as the set of mental, physical and social activities experienced during the phases of human life. Thus, modeling human behavior is a multidimensional problem that can be studied only by targeting specific aspects of each application in the field of interest. The activities of daily living of a human can reveal valuable information about his activity patterns and subsequently, with the association of other behavioral characteristic elements, they can lead in the development of behavioral models. The behavioral characteristic elements (based on the studies that are analyzed in sections 2.2.1 and 2.2.2), except from motion, are:

- Location (indoor and outdoor)
- Interaction with humans, devices, objects
- Time
- Environment

Based on the above, in order to achieve human behavior modeling using activity patterns and to create the profile of the user it is essential to firstly recognize the activities of daily living effectively and accurately. A user's activity profile can be defined as the specific way that a person performs an activity [41]. The majority of the published studies handle the activity recognition task as the recognition of separate activities and not as a sequence of activities. An activity sequence can be defined as the act of several activities in chronological order.

The need of testing sequential dataset of daily activities for identifying abnormal activity patterns of elderly was shown in the study of Mendoza et al. [42], where data was collected over a six month period from a group of older adults in a geriatric center. The dataset consists of timestamped values and flags which characterize the action undertaken in a sequential manner. Authors applied clustering and mining sequential patterns for the analysis.

Yu and Korkmaz in [43] transformed an open sequential dataset [44], which contains binary temporal data from a number of sensing nodes monitoring the ADLs performed in a home setting by a single inhabitant, into a sequence graph and formulated the problem as k-hop longest path problem before the application of a heuristic algorithm.

An important number of studies have been published for deriving activity patterns and/or behavior models based on physical activities in the context of smart environments. Location and motion are two fundamental elements for analyzing human behavior [45], which can be addressed in a smart environment with the use of sensors and tags on objects [46], such as Radio Frequency Identification (RFID) tags [47]. In the context of a smart environment an activity is usually considered as a sequence of events (such as the activation of a RFID tag, or a WiFi fingerprint) that are generated continuously from the sensors [48].

2.2.1. Ontology based approaches for activity modeling

An unsupervised approach based on ontology of activities of daily living and on semantic models of the environment and the deployed devices in combination with probabilistic reasoning methods is proposed by Dimitrov et al. in [46]. Data, including basic and complex activities, were recorded in two ambient intelligence labs. Thus for each instance characteristic information based on the ontologies of activities, devices and environment were combined for explaining the observed event.

Chen et al. [41] introduced an ontology based hybrid approach to activity modeling. The proposed system consists of three main phases: a) creation of initial basic ADL models through ontological engineering; b) the ADL models are used in the application for activity recognition and produce the respective classification results; c) classification results are used for finding new activities and user profiles and to also update the ADL models. Activity monitoring achieved through the interaction of the user with objects with attached sensors, namely, dense sensing technique. Moreover information about the timestamp values and the location were considered. To maintain quality of the model a user needs to validate and finalize the position and the label of an activity model.

Another approach, which combines ontological and pattern clustering techniques is introduced by Gayathri et al. [48] with the objective of identifying activity patterns. The authors used an open public dataset [49], where the sensor data is received in the form of events and contain information such as date, time, sensor identification and status. In the data-driven approach of their methodology, data segmentation based on events and hierarchical event pattern clustering were applied; whereas in the knowledge-driven approach ontology construction and activity modeling based on the occupants, activation of sensors were used for the prediction of the activities.

2.2.2. Activity modeling based on multiple sensors

Rashidi et al. [49] proposed a method which combines sequence mining and clustering algorithms to identify frequent activities and cluster similar activity patterns together. Models are learned to recognize these particular activities with the use of Hidden Markov Models. For the creation of the dataset (publicly available) twenty volunteers performed basic and complex activities namely: telephone use, hand washing, meal preparation, eating and medication use, and cleaning, in a smart apartment. A sensor network was designed to capture data from the following sensors: motion sensors positioned on the ceiling, sensors for ambient temperature readings, custom-built analog sensors to provide readings for hot water, cold water, and stove burner use, Voice over IP for phone usage, contact switch sensors, and pressure sensors. Each sensor reading is tagged with the date and time of the event, the ID of the sensor that generated the event, and the sensor value.

Another approach for activity modelling in the environment of a smart home with embedded sensors and absolute positions of RFID tags is introduced by Amirjavid et al. [50]. Fuzzy clustering methods were applied for constructing activity models, which consider uncertainty as a property of the activities. Uncertainties are characterized as temporal and could arise from: sensors; lack of knowledge; different ways of executing the same scenario; ignorance of world data or variables to avoid from process complexity. Two case studies were undertaken, the first one focused on the recognition of the activity "making coffee" and the second one focused on "drinking water'', ''making tea'' and ''making coffee''.

Pei et al. [45] focused on the development of a system based on smartphone sensors for human behaviour modelling, the LoMoCo model, which combines location information and motion states. For the activity recognition module, four users carried the smartphone in their pants' front pocket in a fixed orientation while they performed, in a closed environment, six scenarios of activities (fetching coffee, fetching water, taking a break, having lunch, working, unknown class of scenario). These activities included the following motion states: sitting, normal walking, fast walking, standing, sharp turning and gradient turning. By analysing the signals obtained from the smartphone's sensors, thirteen features were extracted in both the time and frequency domain. For the feature selection step, the sequential forward selection algorithm was used, while for the classification step the Least Square-Support Vector Machine algorithm was chosen. For the location module, the GPS sensor of the smartphone was used mainly for outdoor and the fingerprinting approach of WiFi positioning (43 reference points) for the indoor location. The authors concluded that although they achieved classification accuracy of approximately 90%, further investigation is required regarding the optimal feature selection method and a study for the detection of abnormal behaviors is also suggested.

An activity sequence-based indoor pedestrian localization approach using smartphones is introduced by Zhou et al. [51]. The proposed approach consists of two main modules: the map matching and the activity detection module. The activity sequence includes several consecutive activities, i.e. turning at a corner, turning around, taking the elevator or the escalator and stairs up and down, which occur when the user walks through the special location points on the building that mark an activity change namely, a corner, an elevator, an escalator, and a stair. Four participants performed the activities while holding the smartphone in their hand. The sampling frequency was set at 100 Hz. Activity detection was achieved based on the accelerometer and the gyroscope sensors of the smartphone, with the use of a decision tree algorithm. After the detection of an activity, a Hidden Markov Model was used to match the activities in the activity sequence to the corresponding nodes, which were labelled with coordinates, of the indoor road network.

3. Methodology

A great challenge in activity recognition is to accurately recognize over a continuous record of sensor data the part of a signal that represents a particular activity. The activity recognition chain/process includes several stages, which include a) data acquisition b) preprocessing c) segmentation d) feature extraction and e) classification, as it can be seen in Figure 3.1. Noise filters can be applied prior or/and post the data segmentation stage. The decision of the methodology followed in each processing stage plays a crucial role in the final outcome of the recognition system.

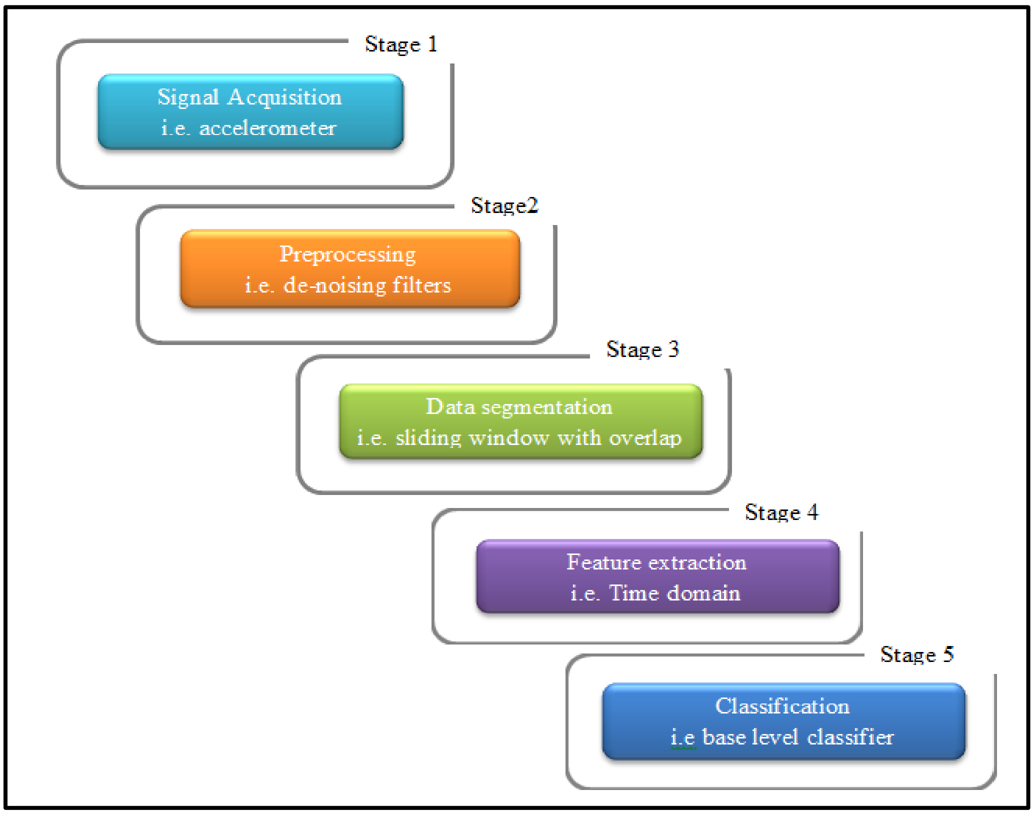

Figure 3.1 Overview of activity recognition process

3.1. Preprocessing

To the direction of increasing the detection accuracy of the activity recognition systems, pre-processing techniques are used. As all process, these techniques have a computational cost that must be taken in to account with respect to the produced improvement of classification. The preprocessing stage includes signal filtering techniques and signal segmentation techniques that are further analyzed in the following sections 3.1.1and 3.1.2.

3.1.1. Signal filtering

The main purpose of signal filtering, in the preprocessing stage of a recognition system, is to eliminate or to minimize noise caused either from the user or the sensor. Another objective is to increase the differences in the contribution of each spectral component to the total signal and in this way make certain wavelengths more selective. According to the signal of interest and the goal of preprocessing different types of filters are applied. General purpose methodologies are smoothing and differentiation. The smoothing technique is applied for reducing the random noise in the signal while differentiation can be used to enhance spectral differences.

In activity recognition with accelerometer signals the most popular filters, according to [52], [12], are a) the band pass filter, which is used to eliminate low frequency acceleration correlated with the orientation of the sensor and high frequency components of noise; b) the low pass filter which is used to eliminate noise caused by dynamic motions of the user; c) an average smoothing method.

Although the filtering technique is applied for noise reduction and thus ultimately for a better classification accuracy, it does increase computational cost, which is crucial when dealing with a real-time system. After a comprehensive study for the impact of preprocessing procedures and in particular for filtering, authors in [52] concluded that the results of their classification were not significantly affected by the presence or absence of noise in the signal. It should also be noted that, when dealing with such a complex signal such as the one obtained from a smartphone's sensors, filtering of the signal could cause unintended removal of important relevant information.

3.1.2. Data segmentation

Data segmentation aims to divide the continuous raw or filtered signal into small signal segments, which represent a respective activity and can be used for feature extraction. The choice of the segmentation technique is very important, due to its impact on the feature-set to be extracted and subsequently used for classification. Once more the computational cost of the technique must be taken in to account. The segmentation process analysis that follows has been conducted based on the findings of a comprehensive study on the processing of sensor data by

Krishnan and Cook [53]. Different approaches on how the sensor data can be handled are presented in Figure 3.2.

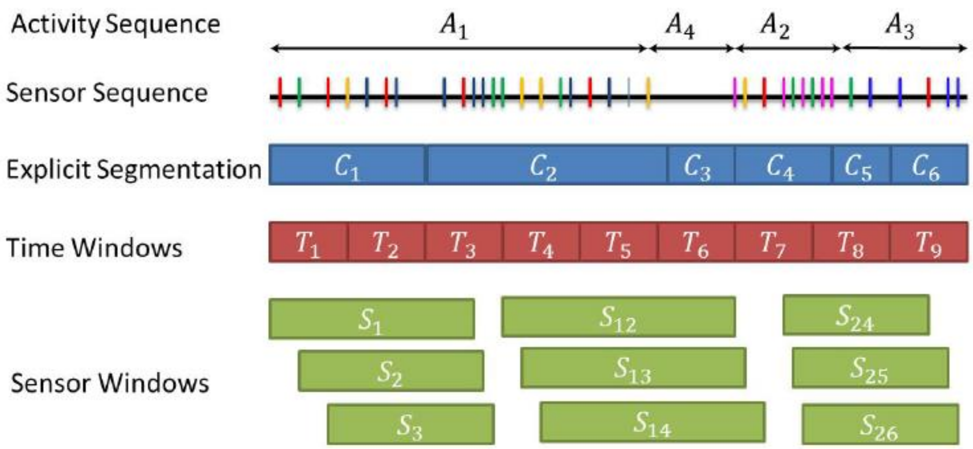

Figure 3.2 Different approaches to processing sensor data (Source [53])

3.1.2.1. Event segmentation

The input sequence of sensor events is partitioned into a number of non-overlapping subsequences [53]. These subsequences represent activity segments and are defined by their corresponding starting and ending time points, as shown in Figure 3.2. It must be noted that the segments may not always align exactly with the aforementioned activity boundaries. Furthermore there are two different approaches commonly used for event segmentation: a) supervised machine learning and b) unsupervised machine learning where features are utilized to identify the boundaries of each activity.

3.1.2.2. Window segmentation

Window segmentation can be further analyzed in to window segmentation with overlapping (also called shift) and window segmentation without overlapping. The size of the window can be either fixed (time or size based) or dynamically adjusted [53]. With respect to the technique, the size of the window has a strong effect on the performance of the recognition system [12], [52], [53], [54], [55]. The most common ranges of window sizes are: a) 1 to 6 seconds for time-based windows and b) 5 to 30 seconds for size-based windows, as described in the subsequent subsections.

Time based windows

In a timestamp-based sliding (also called moving) window, the continuous sensor data is equally divided in time intervals. As a result, with this approach every window includes a fixed amount of data. This approach is commonly used when the system's sensors transmit data at a constant time and rate, such as the accelerometers and gyroscopes [12], [53]. Time-based sliding windows can be used with or without data overlap as it can be seen in Figure 3.3 and Figure 3.4. Depending on the percentage overlap, more or less data overlaps from window N into $N + 1$. This is also referred to as a window shift [56].

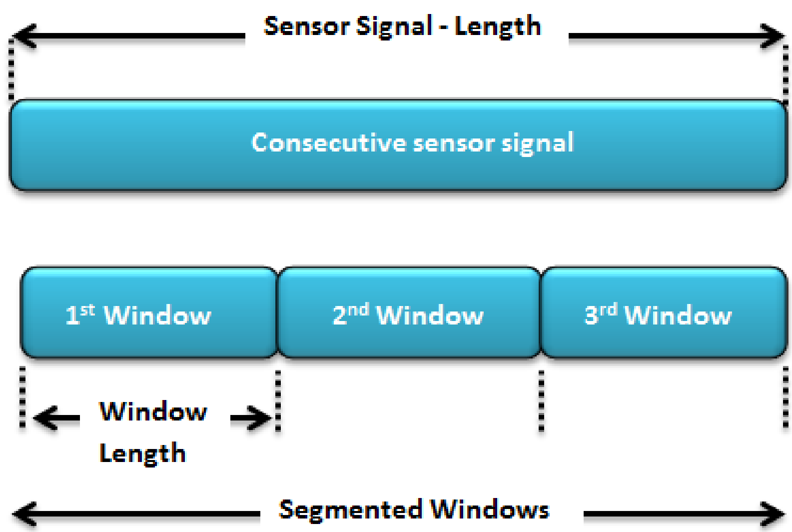

Figure 3.3 Time based sliding window with no overlap

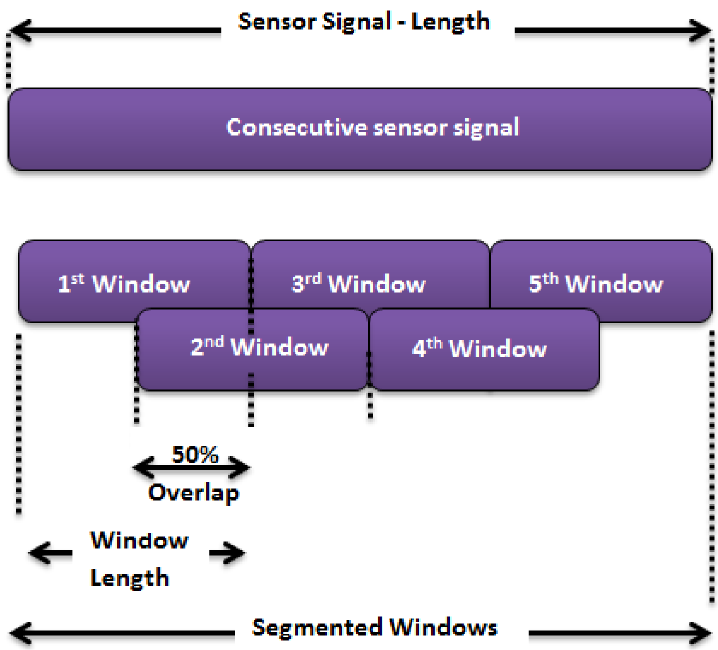

Figure 3.4 Time based sliding window with 50% overlap

Size based windows

In a size-based window approach, also referred to as bursty, fixed-size and sequence-based sliding window, the sensor data are sent to the system in an asynchronous way, not constant in time, and thus the size of the window is defined by the number of sensor events and can vary in time durations. This approach is commonly used when multiple sensors are used [12], [53].

Weighting events within a window

A time-based weighting window is proposed as a solution to the problem of fixed size windows containing sensor events widely spread apart in time. To overcome this, a time-based weighting factor can be applied to each event in the window based on its relative time to the last event in the window [53].

Dynamic window sizes

In a dynamic window size approach, the size of the window is defined dynamically by rules or processes that have been set as prerequisites. Kozina et al. [57] applied dynamic segmentation in acceleration data by deploying methods for searching a significant change point.

The significant change is defined as the difference of values between the maximum and the minimum element in a sequence of consecutive data samples. A combination of a sliding window with overlap and a dynamic sliding window is proposed by Noor et al. [58]. This method categorizes the activities in static, dynamic and traditional. When a traditional activity window is detected, a dynamic sliding window mechanism is executed to expand the window size and capture the activity.

3.2.Feature extraction

Human activity recognition from sensor data is generally preceded by a feature extraction step, in order to obtain the most relevant signal features and to remove the redundant. The outcome of the feature extraction process is a feature set (also called feature vector) that contains all relevant and valuable information that was included in the initial/ non-reduced dataset. Several approaches from different domains of representation have been used in the activity recognition systems based on accelerometer data. The most commonly used and representative extracted features in the time, frequency and discrete domains, based on the analysis of previous studies [4] [13] [59] [60] [61] [62] [63] [64] [65] [66] are presented in the following sections 3.2.1, 3.2.2, and 3.2.3.

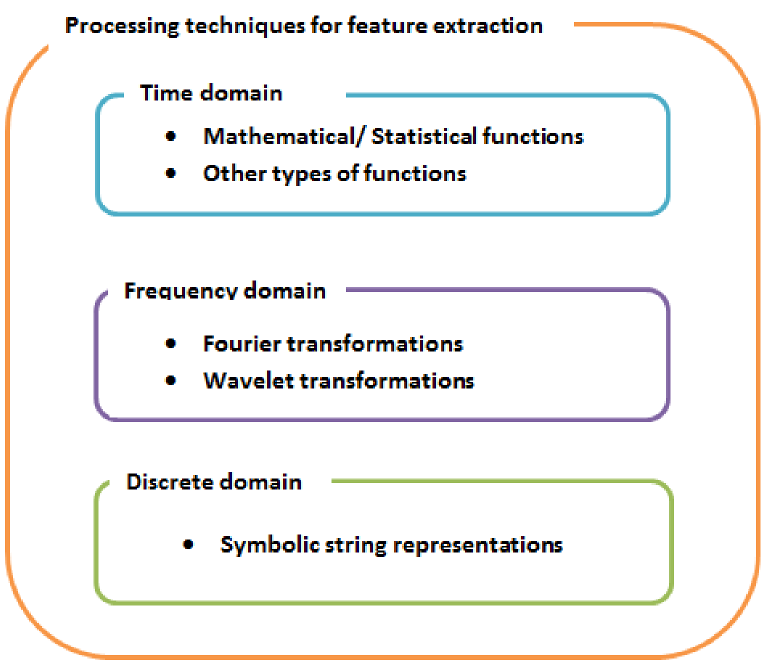

Figure 3.5 Overview of processing techniques for feature extraction, based on [61].

3.2.1. Time domain features

3.2.1.1. Statistical/Mathematical functions

Time-domain features include mean, median, variance, skewness, kurtosis, range, etc. These features are widely used in the field of human activity recognition [67], [68]. Based on the review of the related work conducted, an elaborate list of time-domain features commonly used can be found below.

- **Mean** : Average value
- **Median** : Median signal value
- **Standard Deviation (STD)**: Mean deviation of the signal compared to the average
- **Variance**: Square of the standard deviation
- **Min, Max** : Minimum and maximum of the signal
- **Range:** The difference between the largest and smallest values of the signal
- **Root Mean Square (RMS):** Quadratic mean value of the signal
- **Correlation**: The amount of correlation/similarity that exists between signals
- **Cross-Correlation**: The amount of correlation/similarity of two signals as a function of the lag of one relative to the other
- **Autocorrelation**: cross-correlation of a signal with itself, there will always be a peak at a lag of zero, and its size will be the signal power
- **Linear correlation coefficient**: The amount of correlation/similarity of a linear relationship between two signals
- **Integration**: measures the signal area under the data curve
- **Kurtosis:** The degree of peakedness of the sensor signal distribution toward the mean
- **Skewness:** The degree of asymmetry of the sensor signal distribution. A symmetric dataset will have skewness near 0.
- **Mode**: The number that appears most often in the signal
- **TrimMean**: Trimmed mean of the signal in the window

- **Asymmetry coefficient**: The first moment of the data in the window divided by STD over the window

3.2.1.2. Other functions

- **Zero Crossing Rate (ZCR):** Total number of times the signal changes from positive to negative or back, normalized by the window length
- **Mean Crossing Rate (MCR)**: Total number of times the signal changes from below average to above average, normalized by the window length
- **Differences**
- **Angular velocity**: is the rate of change of angular displacement, specifies the angular speed (rotational speed) of an object and the axis about which the object is rotating
- **Tilt angle**: relates the angle of tilt in the xy-plane, and the angle of inclination from the gravity vector, to the measured acceleration in each axis
- **Signal Magnitude Vector (SMV):** Sum of the Euclidean norm over the three axis over the entire window normalized by the window length
- **Normalized Signal Magnitude Area (SMA):** Acceleration magnitude summed over three axes normalized by the window length

Bouten et al. [67], applied the integral method to offer estimation of energy expenditure using an inertial sensor. The authors used the total Integral of Modulus of Accelerations (IMA). This metric is referred to the time integrals of the module of accelerometer signals:

$$IMA_{tot} = \int_{t=1}^{N} |a_x|\, dt + \int_{t=0}^{N} |a_y|\, dt + \int_{t=0}^{N} |a_z|\, dt$$

Where a_x, a_y, a_z denote the orthogonal components of accelerations, t denotes time and N represents the window length.

Other time-domain features such as Zero-Crossings Correlation-Coefficient root mean square, etc. are also used in [61].

3.2.2. Frequency domain features

3.2.2.1. Fourier Transformations

- **DC (direct current) component**: is the average value of the signal
- **Coefficients sum**: summation of spectral coefficients
- **Dominant frequency**: the one frequency with the largest power in the power spectral density
- **Spectral Energy**: The equivalent to the energy (refers to the area between the signal curve and the time axis) of the signal
- **Spectral Entropy**: Measure of the distribution of frequency components, describes the complexity of a system. And helps to differentiate signals with similar energy values. To calculate entropy -> Power Spectral Density(PSD) then normalize PSD and then calculate entropy
- **Spectral roll-off**: measurement of the skewness of the spectral shape
- **Spectral centroid**: corresponds to the relative location of the "center of gravity" of the spectral power distribution.
- **Spectral flux**: measurement of how quickly the power spectrum of a signal is changing
- **Binned distribution**: represents the fraction of values that fall within equally-spaced bins that span the entire range of sensor values, it is actually the histogram of FFT

3.2.2.2. Wavelet transforms

For a wavelet transform a scalable modulated window is shifted along the signal while the spectrum is calculated for every position. Then, the aforementioned process is repeated for every new cycle. The outcome of this process will be a collection of time-frequency representations of the signal, all with different resolutions. In the wavelet transform only the computation of spectral coefficients are used.

- **Coefficients sum:** summation of spectral coefficients

3.2.3. Discrete domain features

- **Euclidean-based Distances:** corresponds to the numeric distance between two signals
- **Levenshtein Edit Distance:** measurement of the difference between two sequences and thus determines which set of representative samples is more similar to the given input sample.
- **Dynamic Time Warping:** measurement of similarity between two sequences which may vary in time or speed.

3.3.Feature Selection

Feature selection (also called variable selection, attribute selection) is the process of selecting a subset of relevant features for use in the classification scheme of the constructed system. The objective of feature selection is to improve the accuracy, the speed and the cost-effectiveness of the classification system. There are three general classes of feature selection methods, shown in Figure 3.6: filter methods, wrapper methods and embedded methods, [69], which are analyzed in the following sections 3.3.1, 3.3.2 and 3.3.3.

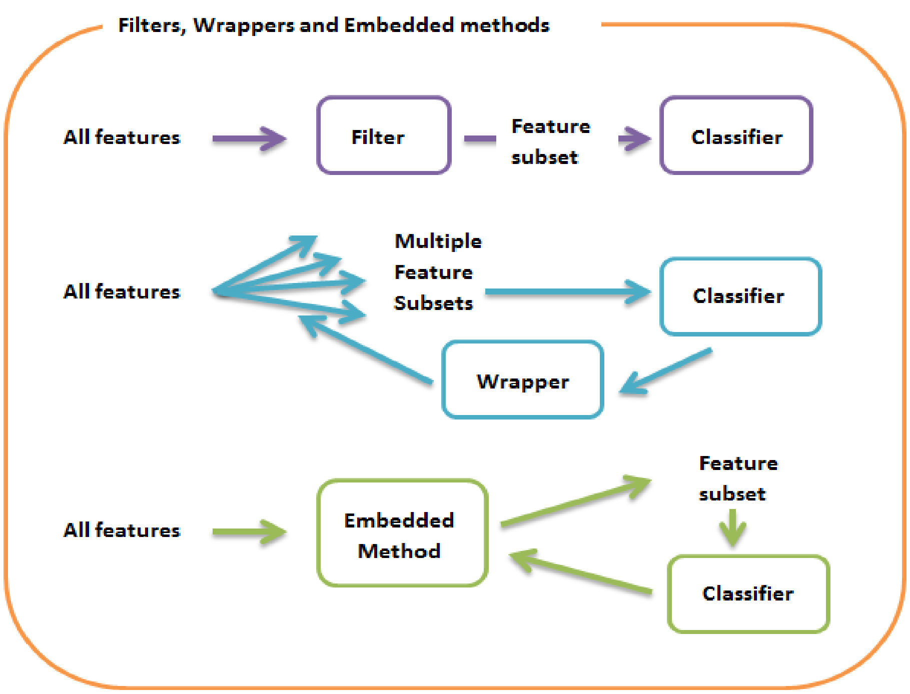

Figure 3.6 The three classes of feature selection methods

Feature extraction and selection techniques are used to face the dimensionality problem, Figure 3.7. In the feature extraction approach the existing features are transformed in to a lower dimensional space while in selection, a subset of the existing features is selected without any further transformation.

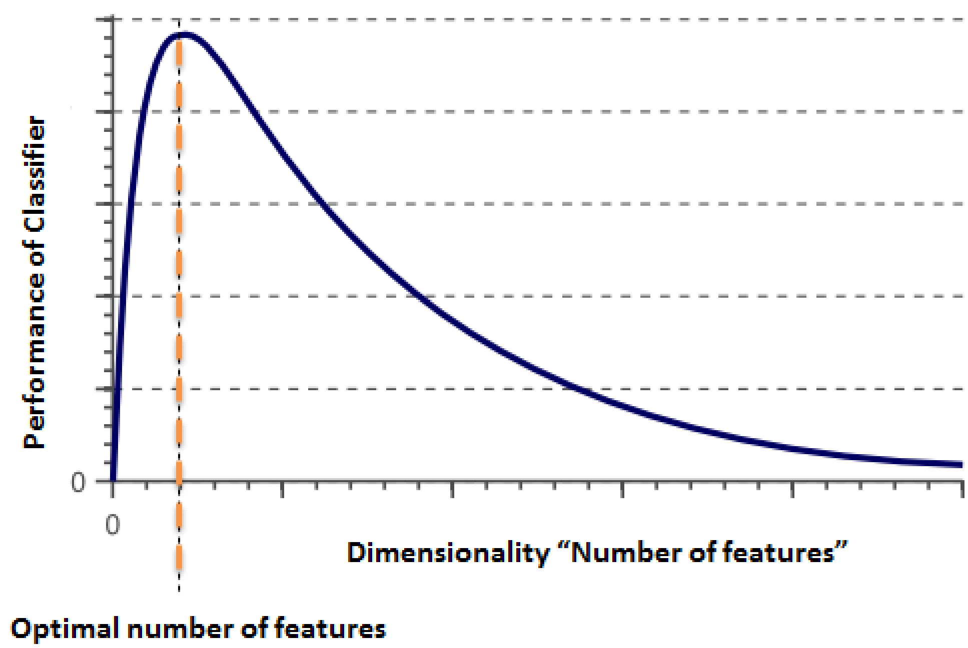

Figure 3.7 The curse of dimensionality. The performance of the classifier increases until the reach of the optimal number of features, further increase of the feature number results in a decrease of the performance.

3.3.1. Filter methods

Filter methods use variable ranking techniques. The features are ranked by the score and a threshold is applied to remove respective features with variables below it, or the cut-off point in the ranking is chosen via cross-validation. Ranking methods are considered as filter methods since they are applied before classification [69]. Because of the method's low complexity, the filter methods are characterized with a low computational cost. The disadvantage of a filter approach is that it does not consider the effect of the selected feature subset on the performance of the applied classification algorithm [70]. Filter methods can also be used as a preprocessing step for wrapper methods, allowing a wrapper to be used on larger problems.

Examples of some filter methods include:

- The Chi squared test
- Information gain
- Correlation coefficient scores.
- Mutual Information

Examples of specific algorithms [63], [71] are:

- ReliefF, [72]
- Correlation-Based Feature Selection (CFS), [73]
- Fast Correlated Based Filter (FCBF), [74]
- INTERACT, [75]

3.3.2. Wrapper methods

Wrapper methods consider the selection of a set of features as a search problem, where different combinations are prepared, evaluated and compared to other combinations. Evaluation of the combined features is performed by a predictive model and a score based on the produced accuracy is assigned. Since wrapper methods apply exhaustive search, they have high computational cost. To overcome this issue, greedy search strategies such as *forward selection* and *backward elimination* are proposed [76]. According to Tang et al. [70], having a predefined classifier, a typical wrapper model will perform the following steps:

Step 1: searching a subset of features,

Step 2: evaluating the selected subset of features by the performance of the classifier,

Step 3: repeating Step 1 and Step 2 until the desired quality is reached.

Examples of some wrapper methods include the following algorithms:

- Best-first search, [77]
- Random hill-climbing algorithm (stochastic), [77]
- Forward and backward passes (heuristics), [69]
- Genetic algorithms (heuristics), [69]
- Recursive feature elimination algorithm, [76]

3.3.3. Embedded methods

Embedded methods incorporate the feature selection as part of training process [76], by learning which features contribute best to the accuracy while the model is being created. With this approach we achieve reduction of the computation time since there is no need of reclassifying different subsets [69]. Embedded methods can be further categorized into three classes: a) the pruning methods, such as Support Vector Machine Recursive Feature Elimination (SVM-RFE) b) models with a build-in mechanism, such as C4.5 decision tree and c) the regularization or penalization methods, which are the most common, such as the LASSO, Elastic Net and Ridge Regression [70].

3.4. Classification

In the classification stage, the task of the classifier is to use the feature vector, provided by the feature extraction stage, in order to assign the signal to a category according to established criteria. The performance of the classifier depends strongly on the variability of the feature values for signals in the same category relative to the variability of feature values for signals in different categories [78]. In respect with the complexity of the described actions and/or signal noise, it can be noted variability in feature values of signals in the same category.

3.4.1. Types of machine learning algorithms

There are two broad categories proposed for the taxonomy of machine learning algorithms, the first one is based on the learning style [79] and the second one is based in the similarity [80]. Based on the learning style, machine learning algorithms can be categorized, in respect of the type of the input in the training procedure and the respective outcome, in to following six groups: supervised learning, unsupervised learning (also called clustering), semi-supervised learning, reinforcement learning, transduction and developmental learning [79], [81].

In general there are no clear borderlines in the taxonomy of classifiers. Manini et al. [30] have categorized the classifiers as follows: Firstly, a differentiation is performed between sequential and single-frame classifiers. A sequential classifier takes into account the previous classifications in order to decide for the current feature vector. On the other hand a single frame classifier assigns labels to the input data regardless of previous assignments. Moreover, single frame classifiers can be further divided in to the following three approaches: 1) probabilistic, where given a sample input, the classifier can predict a probability distribution over a set of classes and not only the most probable class for the sample; 2) geometric, where classification is performed based on decision boundaries, constructed during the training phase, which divide the feature space in order to indicate the region that the input sample belongs to; and 3) binary decision, where the classifier divides the given sample into two classes based on the classification rule that has been set.

3.4.1.1. Supervised learning algorithms

In supervised learning, the model is trained with a set of labeled data that have a corresponding relevance to the testing data. Based on the training set, the algorithm builds a model that can make predictions when new/unknown data is given, thus generalizing the problem. In most cases, the use of a large training set yields models with higher predictive power. Furthermore, based on their similarity the algorithms in supervised learning can be categorized in classification algorithms and regression algorithms. Some of the most popular algorithms for activity recognition with the use of smartphone's sensors [4], [12], [30], [82] , [83], [84], [85] are described below.

Support Vector Machine (SVM)

A support vector machine is a binary classifier which constructs dimensional hyperplanes that separate all data points of one feature class from those of the other feature class [83]. The data points that are closest to the separating "line" of the hyperplanes are called feature vectors. The number of feature vectors is equal to the number of hyperplanes plus one.

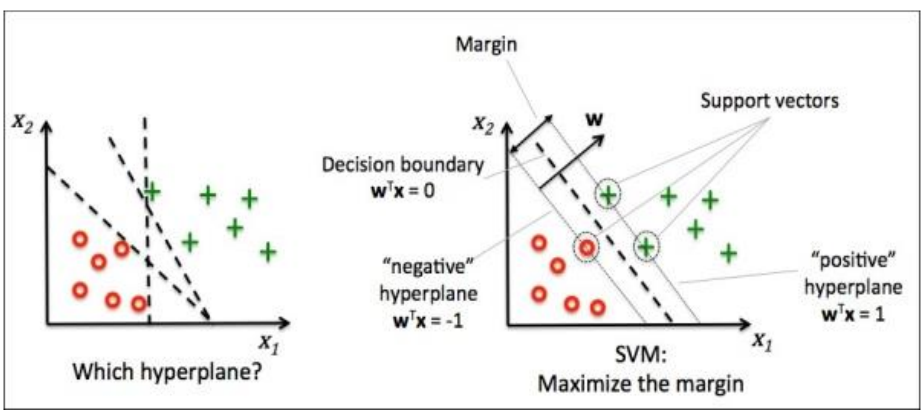

Figure 3.8 An overview of Support Vector Machine (Img. safaribooksonline.com)

A good separation is achieved by the hyperplane that has the largest distance to the nearest training-data point of any class (margin) [79]. The relationship between the margin and the generalization is inversely proportional, the larger the margin the lower the generalization error of the classifier.

Naïve Bayes:

The Naïve Bayes is a simple probabilistic technique based in the Bayesian theorem. The classification of a new input is produced by combining prior probabilities and measures of likelihood in order to form a posterior probability using the Bayes' rule. Furthermore there is a strong independence between features [83]. The Naïve Bayes algorithm can be used for classification and regression.

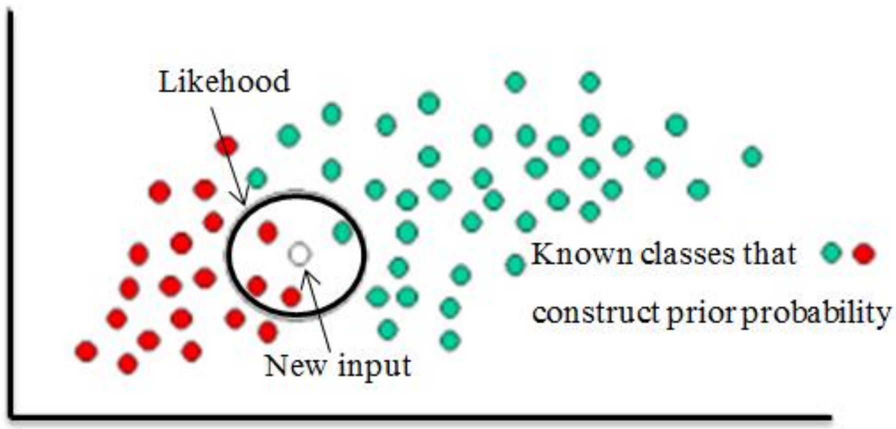

Figure 3.9 An overview of Naïve Bayes

Nearest neighbors (kNN)

The k-Nearest Neighbors algorithm is a non-parametric, instance (or lazy) based algorithm that can be used for classification and regression. The kNN algorithm uses similarity measures (i.e. Euclidian distance) to find the closest samples in the feature space [83]. The value of k determines the number of classes included in the voting scheme. If $k=1$ the input is assigned to the class of that single nearest neighbor. The 1-NN rule provides acceptable classification performance in most applications [86].

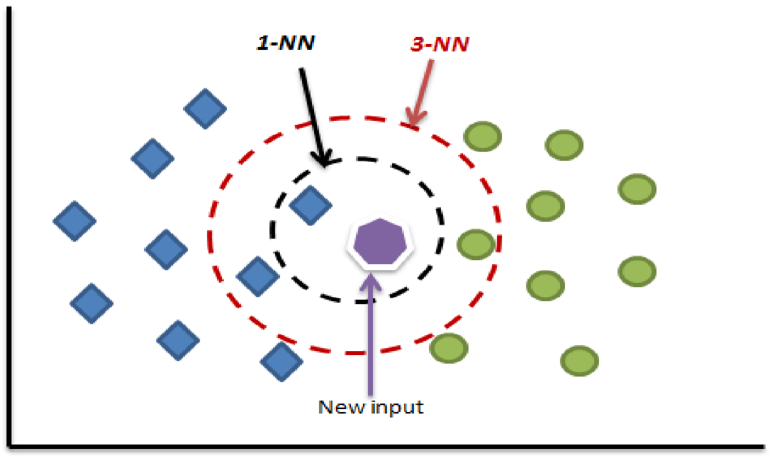

Figure 3.10 An overview of kNN

Neural Network

A neural network consists of layers of neurons (nodes) and weighted interconnections. Generally, it can be defined by the following parameters: a) the interconnection pattern between the different layers of neurons b) the learning process for updating the weights of the interconnections and c) the activation function that converts a neuron's weighted input to its output activation.

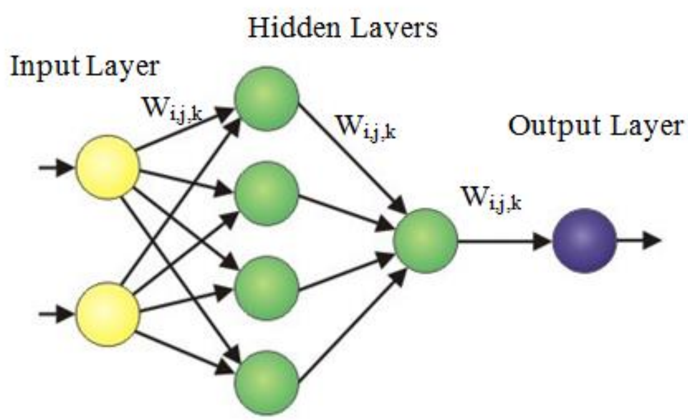

Figure 3.11 An overview of an ANN

Decision trees

Decision trees use a tree-like graph to undertake the decision process of classifying an instance based on the feature values. Starting from the root, each node is a feature from the instance to be classified, and each branch shows the value that the node can receive [87].

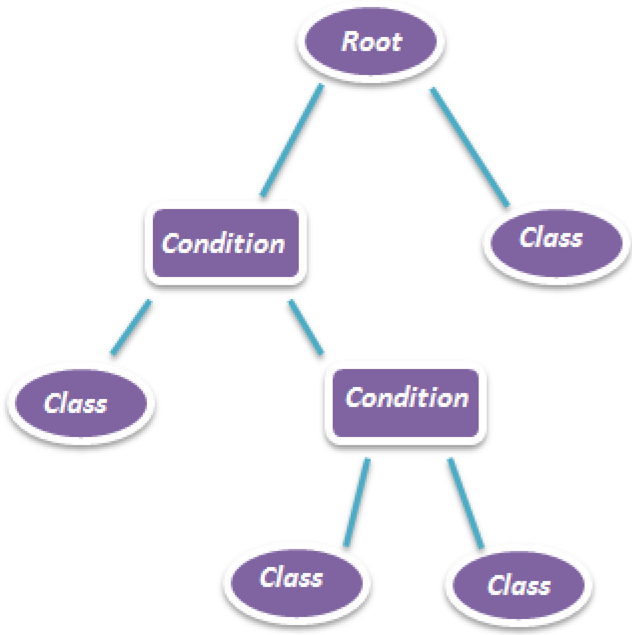

Figure 3.12 An overview of a Decision tree

3.4.1.2. Unsupervised learning

In the case of unsupervised learning the input vector of an algorithm consists of unlabeled data. The inputs are organized into groups based on similarities and differences among the input patterns. Furthermore there is no feedback of evaluation for a potential classification. The dimensionality of the feature space and the complexity of the properties of interest could be higher than in the case of supervised learning [88]. Unsupervised learning includes the following algorithms [89]:

- Clustering techniques (k-Means, K-Median, Hierarchical clustering)
- Self-Organizing Maps
- Principal Component Analysis (Kernel Principal Components, Sparse Principal Components)
- Non-negative Matrix Factorization

3.4.2. Evaluation methods

The most commonly used error estimation method for the evaluation of the classifier of the activity recognition system is the n-fold cross validation method [85]. Another important error estimation method that has been used less frequently but has a high value is the leave-one-out method [31].

3.4.2.1. n-fold cross validation

Cross-Validation is a statistical method of evaluating and comparing learning algorithms. The basic form of cross-validation is *n*-fold cross validation where data is segmented into *n* equal sized partitions, named folds. Subsequently, one fold is kept for testing and the remaining folds are used for the training of the model [90]. This procedure is repeated until all folds have been tested. In machine learning the 10-fold cross validation is most commonly used [91].

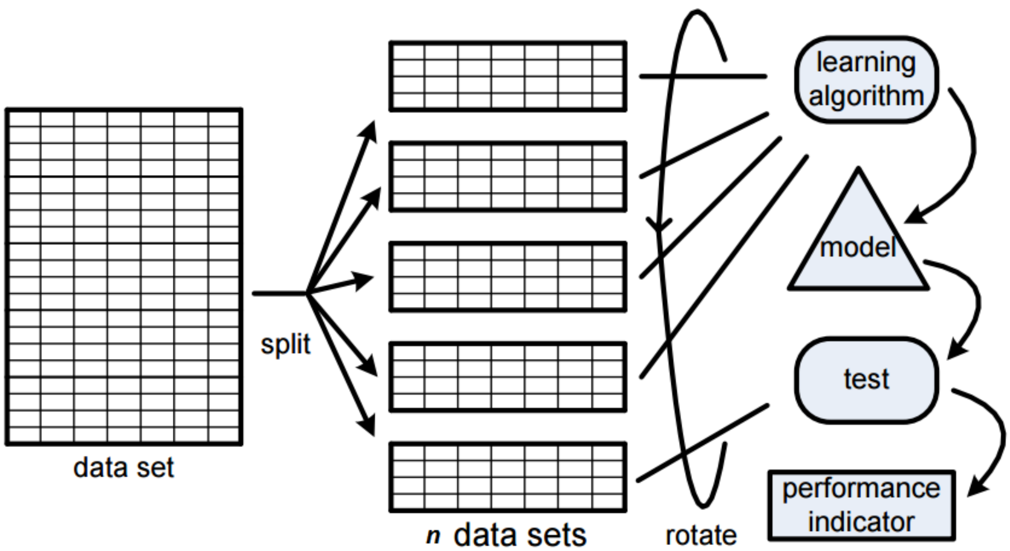

Figure 3.13 n-fold cross validation [92]

3.4.2.2. Leave-One-Out Cross Validation method

The leave-one-out-cross-validation (LOOCV) method is the degenerate case of *n*-fold cross validation, where *n* is equal to the total number of data samples [92]. Thus, the model is tested on a single sample while all the remaining samples are used for the training. Although this method has large computational requirements its estimation is unbiased [86].

3.4.2.3. Leave-One-subject-Out method

Similarly to the LOOCV method, in the Leave-One-subject-Out method the model is tested on a single's subject samples while all the remaining samples from the rest subjects are used for the training. This method has less computational requirements than the LOOCV and retains the unbiased estimation.

3.4.3. Evaluation measures

The evaluation of the performance of an algorithm, in supervised learning, can be measured with the use of a confusion matrix [93]. The confusion matrix is a visualization of the combinations between the predicted instances and the actual instances, as shown in Table 3.1.

Table 3.1 Confusion matrix

		Predicted class		**Total Instances**
		0	**1**	
Actual Class	0	True Positive (TP)	False Negative (FN)	P
	1	False Positive (FP)	True Negative (TN)	N

- True Positives (TP): The number of instances that have correctly classified as “0”
- True Negatives (TN): The number of instances that have correctly classified as “1”
- False Positives (FP), also called Type 1 Errors: The number of instances that are “1” but have classified as “0”
- False Negatives (FN), also called Type 2 Errors: The number of instances that are “0” but have classified as “1”

Using the descriptors of the confusion matrix several other metrics can be computed. The choice of metric depends on the developed algorithm, the application of the system and the desired findings [85]. Nevertheless, the accuracy of the classifier is the most common metric in activity recognition systems [94].

Accuracy describes the number of instances that have been correctly classified:

$$\mathbf{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

Sensitivity (also True Positive Rate or Recall) measures the proportion of positives that are correctly identified as positives.

$$\mathbf{Sensitivity} = \frac{\mathrm{TP}}{\mathrm{P}} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

Specificity (also True Negative Rate) measures the proportion of negatives that are correctly identified as negatives.

$$\textbf{Specificity} = \frac{\mathrm{TN}}{\mathrm{N}} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}}$$

Precision (also Positive Predictive Value) describes the degree to which the predicted positive values is expressed

$$\textbf{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

F-score or F_1score or F-measure is the harmonic mean of precision and sensitivity.

$$\textbf{F} - \textbf{score} = 2 * \frac{\mathrm{TP}}{(2 * \mathrm{TP} + \mathrm{FP} + \mathrm{FN})}$$

The Error Rate measures the deficiency of the algorithm.

$$\textbf{Error Rate} = \frac{\mathrm{FP} + \mathrm{FN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

4. Implementation

4.1.Dataset description

For the purpose of this study a publicly available[2] benchmark dataset namely, the *MobiAct* dataset [4], was used. Furthermore an extended version of the aforementioned dataset was implemented to fit the needs of the study.

The *MobiAct* dataset incorporates signal data from the accelerometer, gyroscope and orientation sensors of a smartphone while participants performed different types of activities and falls. In particular, it encompass four different types of falls and nine different ADLs, as shown in Table 4.1 and Table 4.2, from a total of 57 participants with more than 2500 trials. These types of falls were selected because they can occur from different body positions [13]. The choice of the ADLs were based on bibliography and the three following criteria: a) fall-like activities; b) sudden or rapid activities which produce signals similar to falls; and c) common everyday activities [4].

The extended version of the *MobiAct* dataset includes two extra types of ADLs (Chair up and Sitting), as shown in Table 4.2, and five different scenarios of daily living which include all the different types of the separate ADLs, as shown in Table 4.3, Table 4.4, Table 4.5, Table 4.6 and Table 4.7 . The sequences of the activities are based on a scenario of daily living where the subject leaves his home, takes his car to get to his working place (although real driving was not recorded), reaches his office, sits on the chair and starts working. Once he gets off his work, he takes his car and goes in an open area to perform physical exercise. Once more he gets in the car and finally returns back home to rest. The initial scenario was split into five sub-scenarios, which are connected with idle ADLs (standing and sitting), in order to avoid recording issues that would lead to several repetitions and the frustration of the participants. As it can be seen the scenarios evokes all the ADLs included in the *MobiAct* dataset, a number of which have been recognized with accuracy of 99% in a previous study [4]. The main purpose for the construction of scenarios is to investigate how the recognition of different activities with natural transitions between them in continuous recordings, will affect the performance of the system, since in a real life scenario there is no clear separation from one activity to another. Although the durations of

---

[2] The *MobiAct* dataset is available for download from www.bmi.teicrete.gr

the scenarios are short, one can assume that if the included activities can be accurately recognized, then this can also be achieved in a wider time frame as well. This could subsequently lead to the generation of activity patterns if the transition points between activities, the activities themselves and as the parameter of time will be combined together. The precise recognition of scenarios of daily living, in a base level, is crucial for a recognition system applicable on real life. As far as we know there is no study in the domain of activity recognition using smartphones that addresses the issue of recognizing sequences of activities or scenarios of activities. As shown in section 2.1.1 of this work, the studies that evoke continuous recordings request from the participant to remain inactive for a period of time before the beginning of a new activity.

In summary the extended version of the *MobiAct* dataset includes:

- Four different types of falls performed by 66 participants
- Eleven different types of ADLs performed by 19 participants and nine types of ADLs performed by 59 participants
- Five sub-scenarios which construct one scenario of daily living, which consists of a sequence of 50 activities and performed by 19 participants.

Table 4.1 Falls recorded in the *MobiAct* dataset

Label	Activity	Trials	Duration	Description
FOL	Forward-lying	3	10s	Fall Forward from standing, use of hands to dampen fall
FKL	Front-knees-lying	3	10s	Fall forward from standing, first impact on knees
SDL	Sideward-lying	3	10s	Fall sideward from standing, bending legs
BSC	Back-sitting-chair	3	10s	Fall backward while trying to sit on a chair

Table 4.2 Activities of Daily Living recorded in the *MobiAct* dataset

	Label	Activity	Trials	Duration	Description
Initial Version of MobiAct	STD	Standing	1	5m	Standing with subtle movements
	WAL	Walking	1	5m	Normal walking
	JOG	Jogging	3	30s	Jogging
	JUM	Jumping	3	30s	Continuous jumping
	STU	Stairs up	6	10s	Stairs up (10 stairs)
	STN	Stairs down	6	10s	Stairs down (10 stairs)
	SCH	Sit chair	6	6s	Sitting on a chair
	CSI	Car step in	6	6s	Step in a car
	CSO	Car step out	6	6s	Step out of a car
New	CHU	Chair up	6	6s	Getting up from a chair
	SIT	Sitting	1	60s	Sitting with subtle movements

Table 4.3 1st Scenario of Leaving the Home (SLH)

1st Scenario of Leaving the Home (SLH), Total duration 125’’				
	No.	Label	Activity	Description
New Scenarios of MobiAct dataset	1	STD	Standing	The recording starts with the participant standing outside the door and locking the door. Then walks and descent stairs to leave his home. Following, he riches the parking area where he stands in front of the car, unlocks the lock of the car, opens the door and gets in the car. He remains sited for some seconds, then he gets out of the car, closes the door and stands in front of the door to lock the car.
	2	WAL	Walking	
	3	STN	Stairs down	
	4	WAL	Walking	
	5	STD	Standing	
	6	CSI	Car-step in	
	7	SIT	Sitting	
	8	CSO	Car-step out	
	9	STD	Standing	

Table 4.4 2nd Scenario of Being at Work (SBW),

<table>
<tr><th colspan="5">2nd Scenario of Being at work (SBW), Total duration 185’’</th></tr>
<tr><td rowspan="9">New Scenarios of MobiAct dataset</td><td>No.</td><td>Label</td><td>Activity</td><td>Description</td></tr>
<tr><td>1</td><td>STD</td><td>Standing</td><td rowspan="8">The recording starts with the participant standing outside the cars door. Then walks from the parking area to his work building. He walks and ascent stairs till he riches his office where he stops in front of the door. Once he finds the keys he opens the door, gets in his office and walks to his chair, where he sits.</td></tr>
<tr><td>2</td><td>WAL</td><td>Walking</td></tr>
<tr><td>3</td><td>STU</td><td>Stairs up</td></tr>
<tr><td>4</td><td>WAL</td><td>Walking</td></tr>
<tr><td>5</td><td>STD</td><td>Standing</td></tr>
<tr><td>6</td><td>WAL</td><td>Walking</td></tr>
<tr><td>7</td><td>SCH</td><td>Sit chair</td></tr>
<tr><td>8</td><td>SIT</td><td>Sitting</td></tr>
</table>

Table 4.5 3rd Scenario of Leaving work (SLW)

<table>
<tr><th colspan="5">3rd Scenario of Leaving work (SLW), Total duration 185’’</th></tr>
<tr><td rowspan="13">New Scenarios of MobiAct dataset</td><td>No.</td><td>Label</td><td>Activity</td><td>Description</td></tr>
<tr><td>1</td><td>SIT</td><td>Sitting</td><td rowspan="12">The recording starts with the participant sitting in the chair in his office area. Then he gets up from the chair, walks to the door and stands outside the office door. Once he find the keys, he lock the door and walks and descent stairs till he riches the parking area. He walks to his car and stands in front of the car, unlocks the lock of the car, opens the door and gets in the car. He remains sited for some seconds, then he gets out of the car, closes the door and stands in front of the door to lock the car.</td></tr>
<tr><td>2</td><td>CHU</td><td>Chair up</td></tr>
<tr><td>3</td><td>WAL</td><td>Walking</td></tr>
<tr><td>4</td><td>STD</td><td>Standing</td></tr>
<tr><td>5</td><td>WAL</td><td>Walking</td></tr>
<tr><td>6</td><td>STN</td><td>Stairs down</td></tr>
<tr><td>7</td><td>WAL</td><td>Walking</td></tr>
<tr><td>8</td><td>STD</td><td>Standing</td></tr>
<tr><td>9</td><td>CSI</td><td>Car-step in</td></tr>
<tr><td>10</td><td>SIT</td><td>Sitting</td></tr>
<tr><td>11</td><td>CSO</td><td>Car-step out</td></tr>
<tr><td>12</td><td>STD</td><td>Standing</td></tr>
</table>

Table 4.6 4th Scenario of Being Exercise (SBE)

<table>
<tr><th colspan="5">4th Scenario of Being Exercise (SBE), Total duration 125’’</th></tr>
<tr><td rowspan="10">New Scenarios of MobiAct dataset</td><td>No.</td><td>Label</td><td>Activity</td><td>Description</td></tr>
<tr><td>1</td><td>STD</td><td>Standing</td><td rowspan="9">The recording starts with the participant standing in front of the car. He starts his exercise by walking, then starts jogging from some seconds and once again walking. Then he stops for some seconds to get a breath and he starts jumping and once more he standing to relax a little. Finally he walks till his car and stands outside the door.</td></tr>
<tr><td>2</td><td>WAL</td><td>Walking</td></tr>
<tr><td>3</td><td>JOG</td><td>Jogging</td></tr>
<tr><td>4</td><td>WAL</td><td>Walking</td></tr>
<tr><td>5</td><td>STD</td><td>Standing</td></tr>
<tr><td>6</td><td>JUM</td><td>Jumping</td></tr>
<tr><td>7</td><td>STD</td><td>Standing</td></tr>
<tr><td>8</td><td>WAL</td><td>Walking</td></tr>
<tr><td>9</td><td>STD</td><td>Standing</td></tr>
</table>

Table 4.7 5th Scenario of Returning at Home (SRH)

<table>
<tr><th colspan="5">5th Scenario of Returning at Home (SRH), Total duration 155’’</th></tr>
<tr><td rowspan="13">New Scenarios of MobiAct dataset</td><td>No.</td><td>Label</td><td>Activity</td><td>Description</td></tr>
<tr><td>1</td><td>STD</td><td>Standing</td><td rowspan="12">The recording starts with the participant standing outside the cars door. He unlocks the lock of the car, opens the door and gets in the car. He remains sited for some seconds, then he gets out of the car, closes the door and stands in front of the door to lock the car. Then walks from the parking area to his home. He walks and ascent stairs till riches his home door, where he stands to finds the keys. Then he opens the door, gets in his home, walks till a chair and sits.</td></tr>
<tr><td>2</td><td>CSI</td><td>Car-step in</td></tr>
<tr><td>3</td><td>SIT</td><td>Sitting</td></tr>
<tr><td>4</td><td>CSO</td><td>Car-step out</td></tr>
<tr><td>5</td><td>STD</td><td>Standing</td></tr>
<tr><td>6</td><td>WAL</td><td>Walking</td></tr>
<tr><td>7</td><td>STU</td><td>Stairs up</td></tr>
<tr><td>8</td><td>WAL</td><td>Walking</td></tr>
<tr><td>9</td><td>STD</td><td>Standing</td></tr>
<tr><td>10</td><td>WAL</td><td>Walking</td></tr>
<tr><td>11</td><td>SCH</td><td>Sit chair</td></tr>
<tr><td>12</td><td>SIT</td><td>Sitting</td></tr>
</table>

4.1.1. Data acquisition

The extended version of the *MobiAct* dataset was developed keeping the exact same acquisition protocol as for the *MobiAct* dataset. All the recorded activities and scenarios were performed at the Technological Educational Institute of Crete.

The main objective is to simulate real life conditions, which is a key focus of research in the field [2]. Based on this objective, the smartphone was placed loosely in any random orientation into the trousers' pocket by the participant. The accelerometer, gyroscope and orientation sensors of a Samsung Galaxy S3 smartphone with the LSM330DLC inertial module (3D accelerometer and gyroscope) were used. As the orientation sensor is software-based, it derives its data from the accelerometer and the geomagnetic field sensor. A calibration of the gyroscope was performed using the device's integrated tool, prior to the recordings. For the data acquisition, an Android application was used for the recording of raw data for the acceleration, the angular velocity and orientation [95]. In order to achieve the highest sampling rate possible the parameter "SENSOR_DELAY_FASTEST" was enabled. Finally, each sample was stored along with its timestamp in nanoseconds. For each recording the application produces three files, one for each sensor, in a txt format.

Each activity recording was performed in the unique way of each participant, while at the same time guidance from the instructor was provided in order to ensure the reliability of data. For example, the instructor informs the participant to sit on a particular chair, with his own way and rhythm. The participants placed the smartphone in the pocket; after hearing a sound signal notifying the start, she/he sits on the chair and waits until the stop signal sounds. The whole process is monitored by the instructor. When recording complex scenarios, the instructor also performed the scenario activities along with the participant in order to hold a timestamp for each change of activity with a second smartphone. In addition to the monitoring for activity change, participants were instructed to inform the instructor with voice commands each time they would transit into another activity that could not be obvious. For example, in case of the participant sitting inside the car and the next activity is to step out of the car, then the instructor asks the participant to step out and confirm back the command just before starting the action. It was crucial to make sure that all the sequences and transitions of activities that construct the scenarios were performed in a real, natural, manner.

4.1.2. Dataset characteristics or study participants

The extended version of the *MobiAct* dataset includes records from 66 participants, 51 men and 15 women. In particular, 66 subjects performed the falls described in Table 4.1, 59 subjects performed nine of the eleven ADLs described in Table 4.2 while 19 performed all the ADLs, and finally 19 subjects performed the scenarios presented in Tables 4.3-4.7. The subjects' age spanned between 20 and 47 years, the height ranged from 160 cm to 193 cm, and the weight varied from 50 kg to 120 kg. The average profile of the subject that occurs based on the described characteristics is 26 years old, 176 cm of height and 76 kg weight. All participants had different physical status, ranging from completely untrained up to athletes (minimum of cases). The challenge of the generalization [2] is addressed due to the high number of participants, the range of ages and the range of physical status included in the MobiAct dataset.

4.2.Methodology of signal processing

4.2.1. Software and tools used

For data collection an android application, which was developed previously [4], [13], [95], was used. Subsequently in the processing chain, the matrix laboratory (MATLAB) platform was used for the signal data annotation and for the feature extraction. Finally for the feature selection and classification process the Waikato Environment for Knowledge Analysis (WEKA) tool [96], a popular suite of machine learning algorithms developed at the University of Waikato, was used.

4.2.2. Pre-processing

Based on the findings of Fida et al. [52] described earlier in Section 3.1 and in a try to avoid removal of important signal information, filtering techniques did not applied in this work. The pre-processing of raw signal data consists of three main steps: 1) sorting and conversation of file's format, 2) synchronization and 3) data annotation.

The sorting of the files was made according to the type of each recording (type of activity, sub- scenarios). Subsequently, a conversion to editable format for further processing in matlab (.mat files) was applied.

Fort the synchronization of files the linear interpolation technique was used for handling the issues of the variability in the values of the timestamp and the sampling rate. The change on the sampling rate is caused by the Android function, onSensorChanged, which runs for each sensor when its value is changing and not simultaneously.

Since the developed recognition system was based on supervised learning algorithms the annotation of the data was mandatory. Labels were assigned to the data manually, following slightly different protocol for the two broad categories, ADLs and Scenarios. The labels and the corresponding activities and the labels of the sequence of activities which produce the scenarios are described in detail in Section 4.1 and in particular in Tables 4.1-4.7. Starting with the annotation of the separate ADLs, the signal is labeled from one to three labels, representing the initial activity state of the participant, the desired ADL, and the ending activity state. For example in case of the SCH (sit on chair) activity the initial state is “sitting”, following “sit on chair” and the ending state is “standing”. The annotation was made, empirically, by monitoring the signal change and recording the respective time values. Examples of signal annotation, randomly chosen recordings, for all the ADLs are shown in Figures 4.1- 4.11.

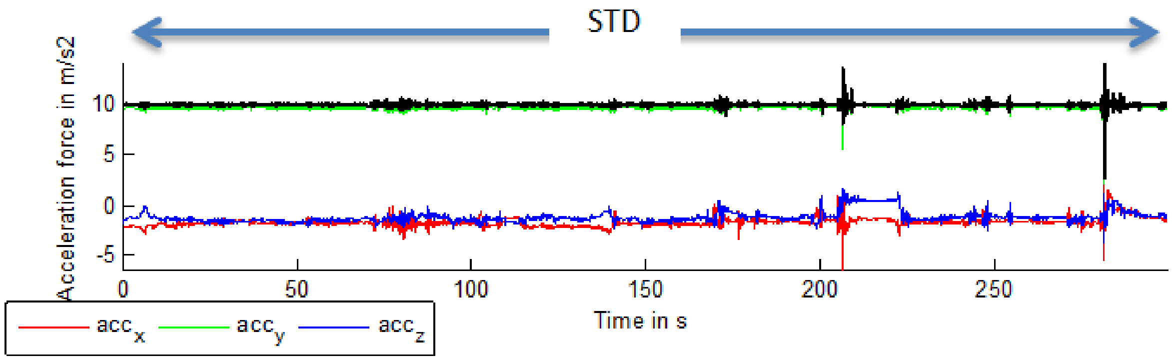

Figure 4.1 Annotation of the activity "Standing"

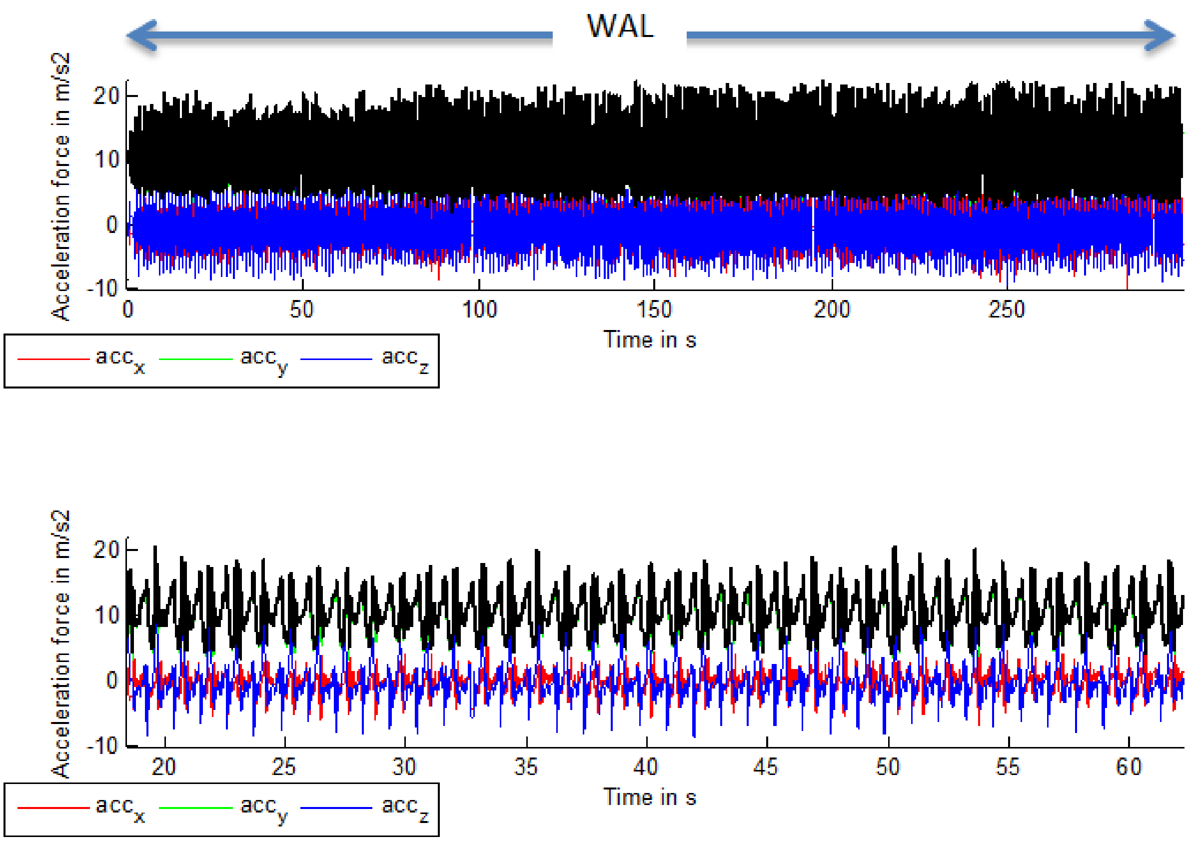

Figure 4.2 Annotation of the activity "Walking", original and zoomed in

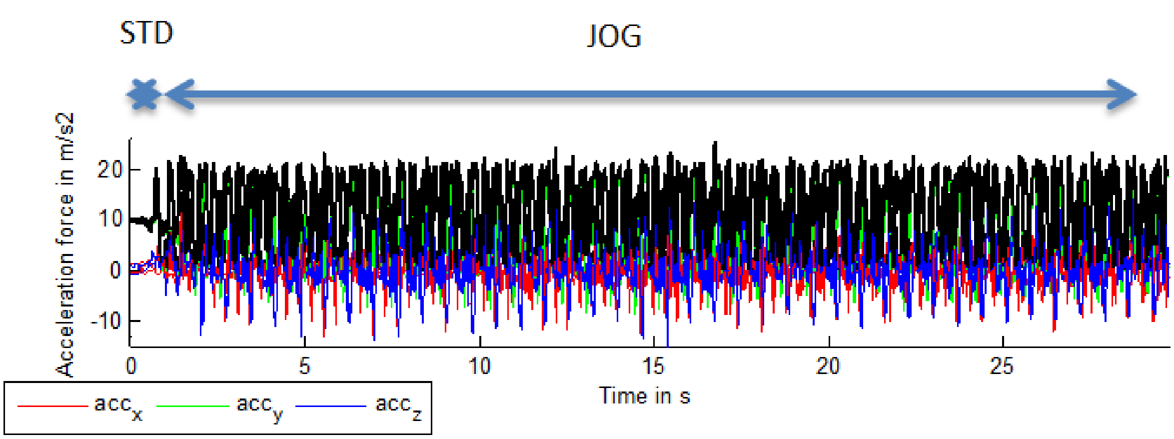

Figure 4.3 Annotation of the activity "Jogging"

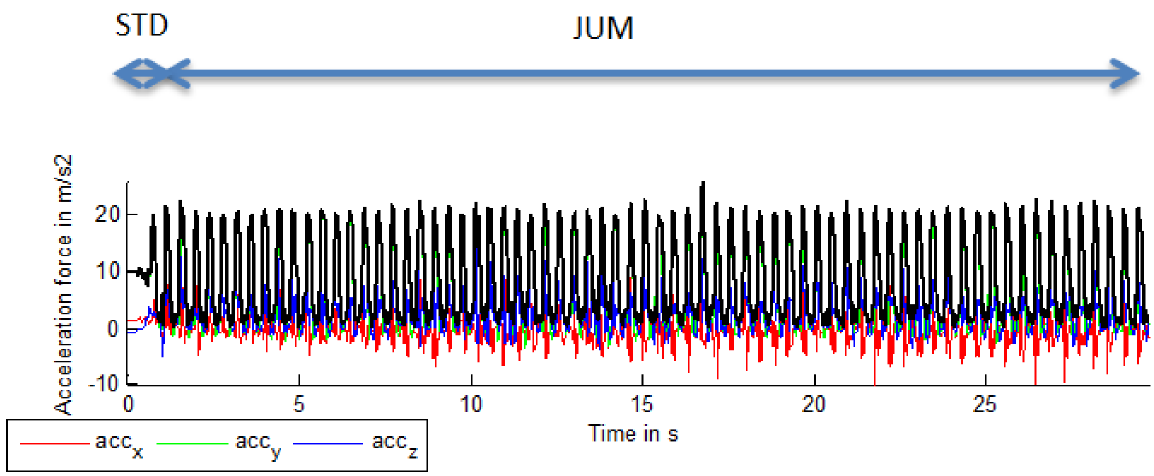

Figure 4.4 Annotation of the activity "Jumping"

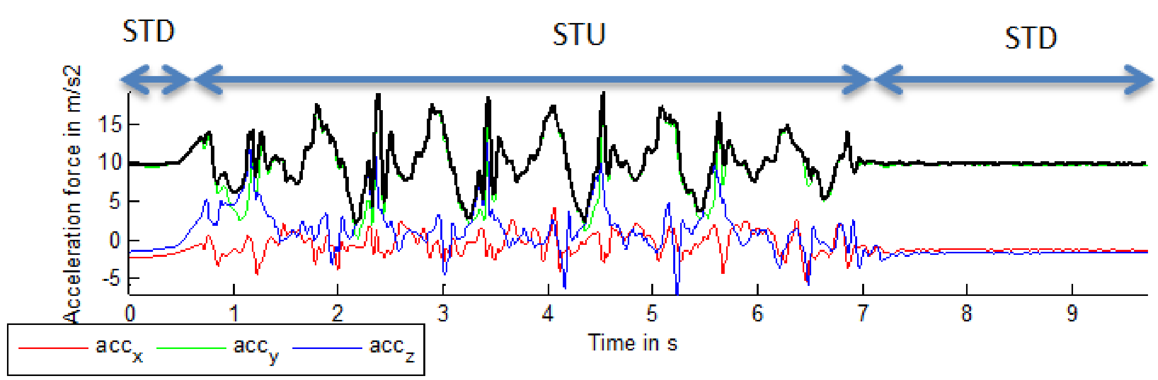

Figure 4.5 Annotation of the activity "Stairs up"

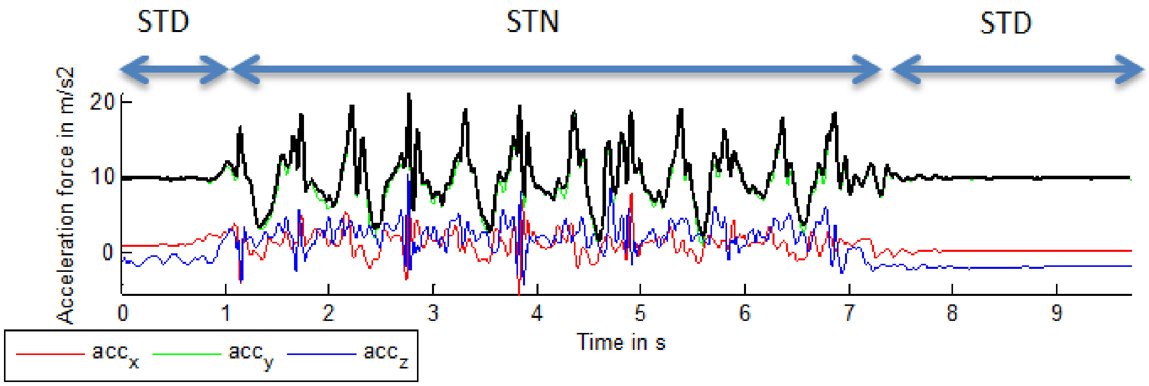

Figure 4.6 Annotation of the activity "Stairs down"

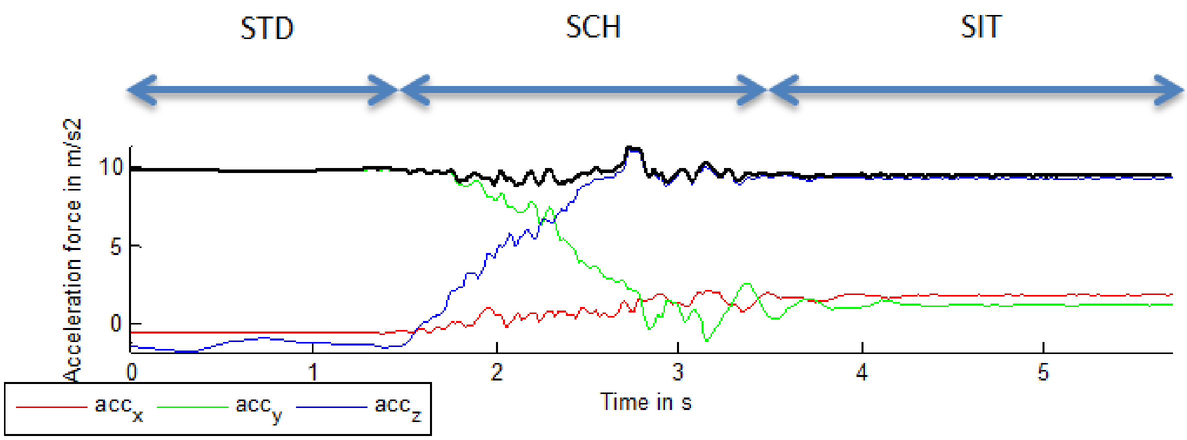

Figure 4.7 Annotation of the activity "Stairs down"

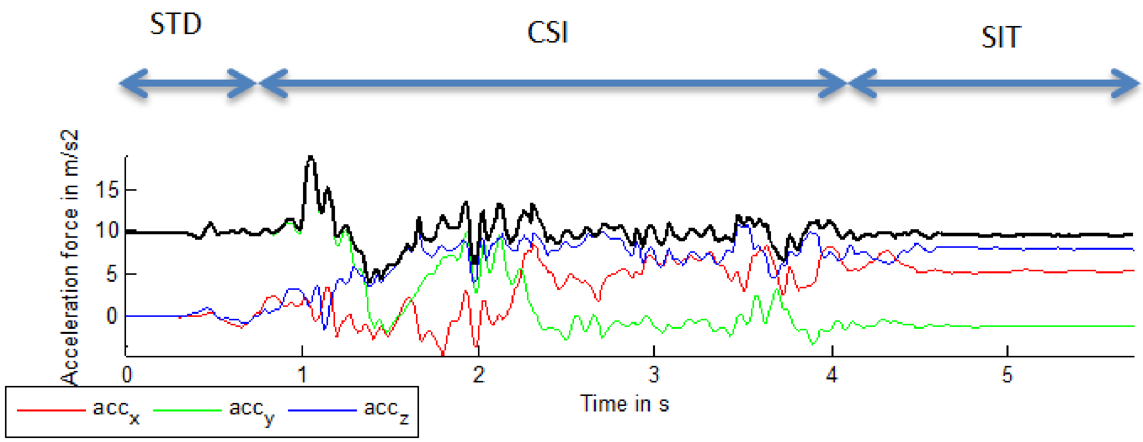

Figure 4.8 Annotation of the activity "Car step in"

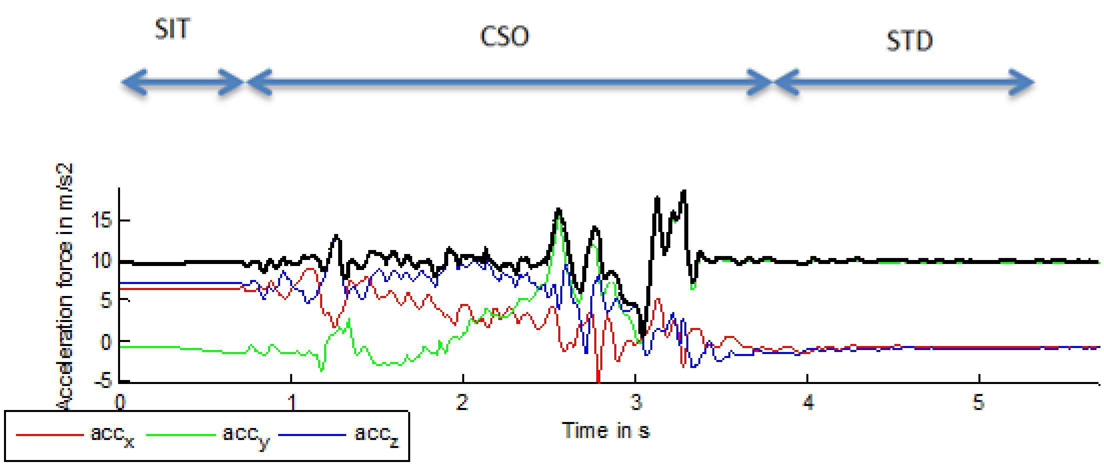

Figure 4.9 Annotation of the activity "Car step out"

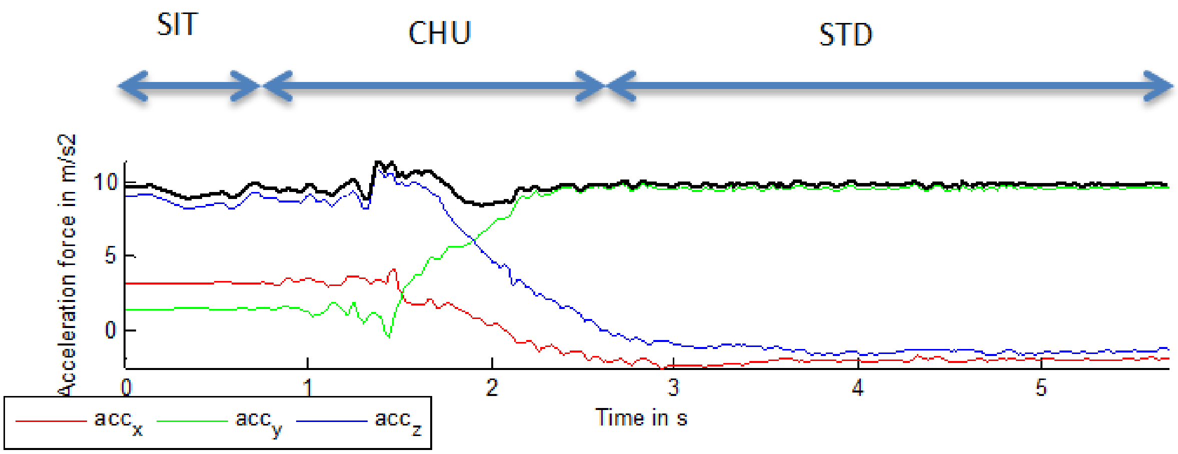

Figure 4.10 Annotation of the activity "Chair up"

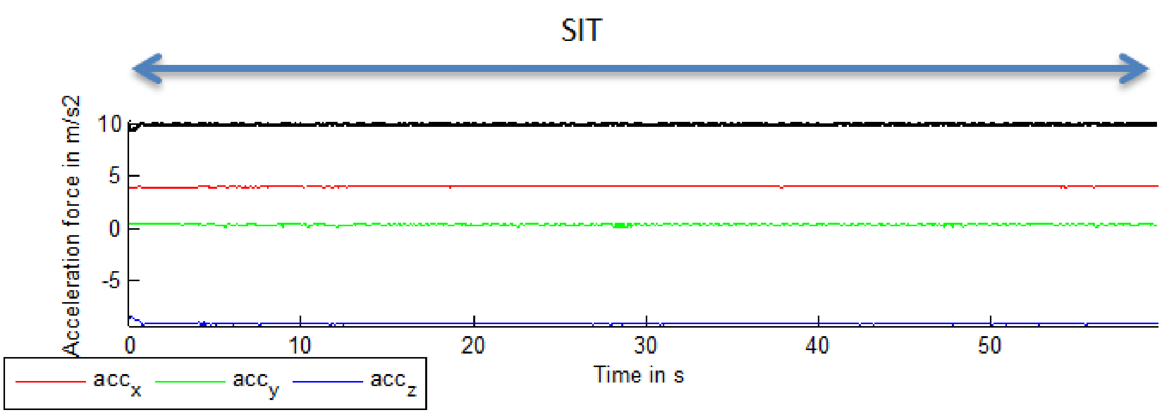

Figure 4.11 Annotation of the activity "Sitting"

Concerning the annotation process of the scenarios, there were three different sources to accurately record the time point where a transition of an activity was performed. As described in Section 4.1, the instructor of the recordings was saving a timestamp in another smartphone for each activity transition and for cross-check a handwritten copy was made. Thus, for the annotation of the scenarios the same protocol with the annotation of ADLs was followed and in advance, two checks were performed with the recorded timestamps. Examples of scenario annotations are shown in Figures 4.12 - 4.17.

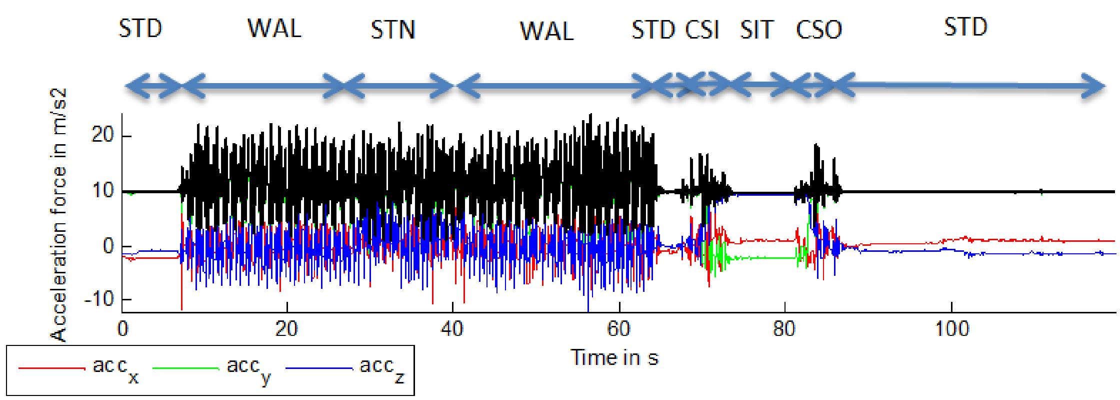

Figure 4.12 Annotation of the scenario "leaving home- SLH"

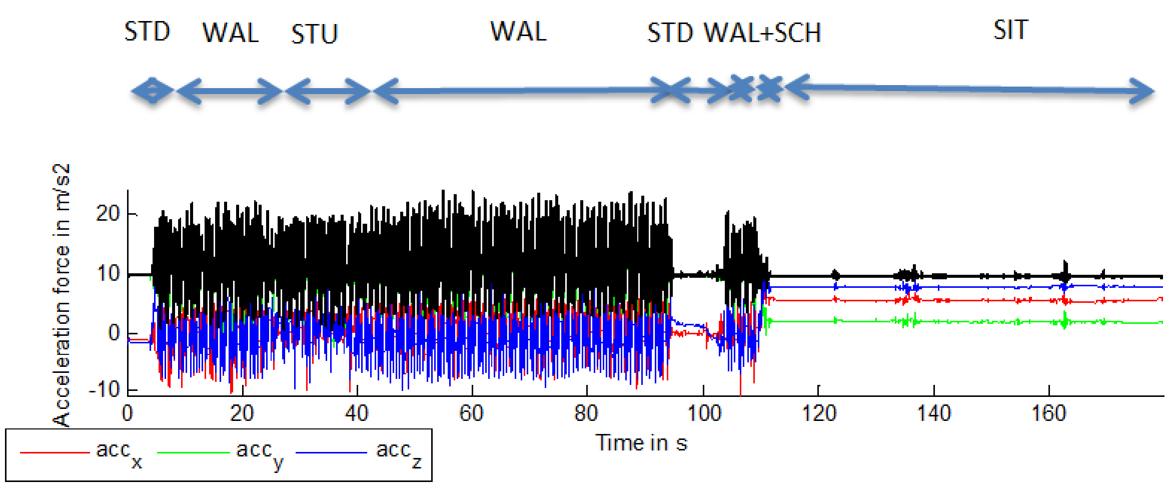

Figure 4.13 Annotation of the scenario "Being at work- SBW"

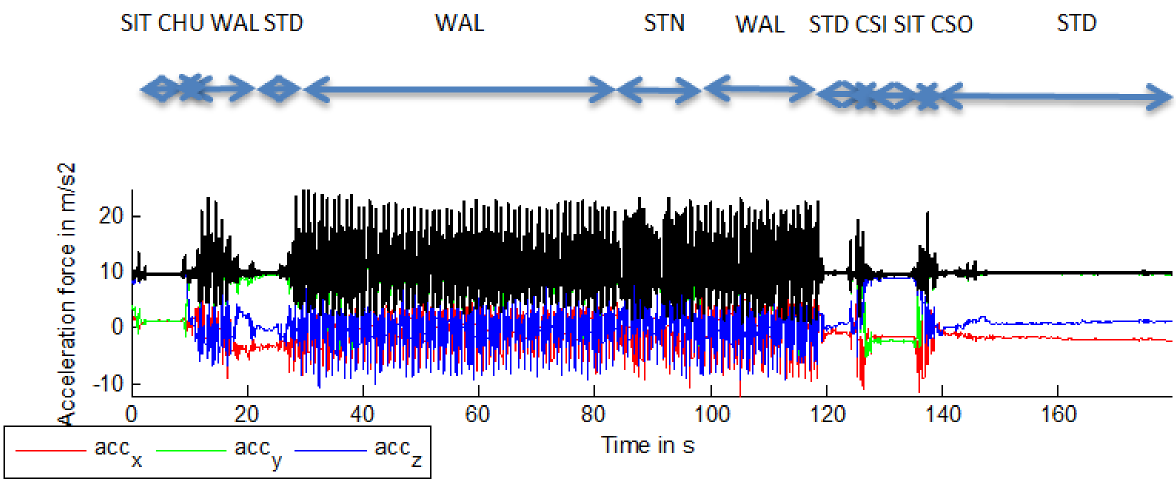

Figure 4.14 Annotation of the scenario "Leaving work- SLW"

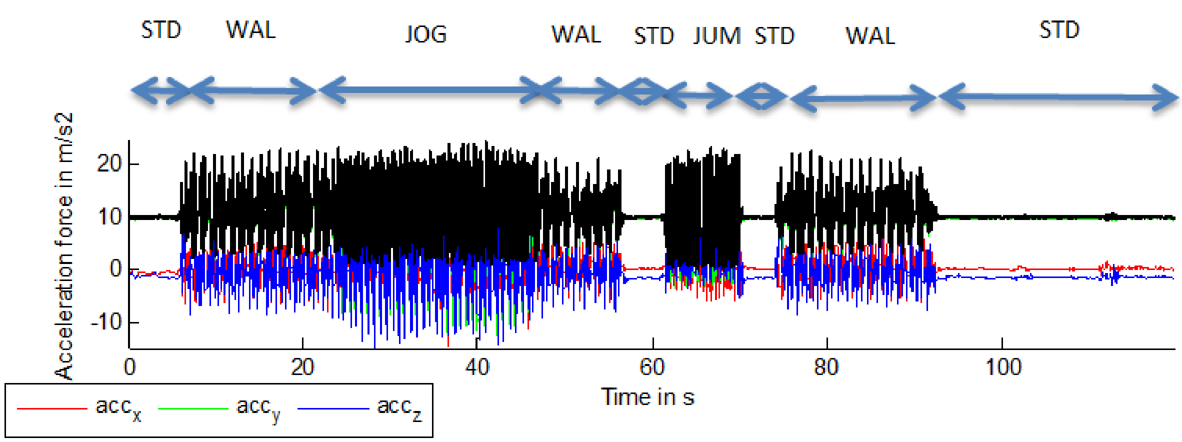

Figure 4.15 Annotation of the scenario "Being exercise - SBE"

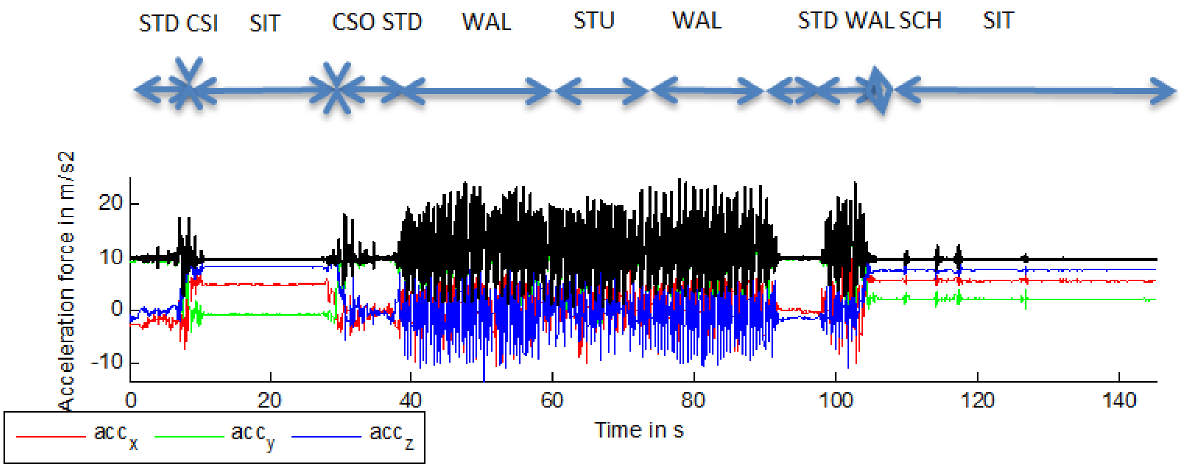

Figure 4.16: Annotation of the scenario "Return Home - SRH"

4.2.3. Feature extraction & selection

The first step before the feature extraction process is the determination of the window segmentation technique. The time-based sliding window is the most widely used technique for recognition systems based on accelerometer data. The window size has a high impact in the performance of the system [52]. The time length of the activities for recognition and the window size have a strong correlation, i.e. a big window size will overcome an activity of small duration. Taking in to consideration that the duration of the smallest investigated activities varies between 1,5 and 2 seconds (CHU, SCH) and based on previous studies [12] [53], [54], [55] three different sizes of windows, 1, 1.5 and 2 second, were tested. Furthermore an overlap of 80%, based on previous study [4] was combined with each window size.

The choice of features for extraction was based on the findings of a previously published study [4] while a further try to reduce the number of the feature set was performed, for reduction of the computational cost. Features were extracted from time and frequency domains, which are also accords to the most representative in the domain of activity recognition as described in Section 3.2. Subsequently, four features sets were created which were tested with all possible combinations for each window size.

Feature set A

In the first feature set 28 features in total have been excluded from the proposed set reported by Vavoulas et al. [4]. 27 features excluded from the absolute of signal and the spectral centroid as it has been reported to affect the results of activity recognition negatively. The first feature set consists of 40 features in total from time and frequency domain.

- 21 features: mean, median, standard deviation, skew, kurtosis, minimum and maximum of each axis (x, y, z) of the acceleration.
- 1 feature: the slope SL, which is defined as:

$$SL = \sqrt{(max_x - min_x)^2 + (max_y - min_y)^2 + (max_z - min_z)^2}$$

- 4 features: mean, standard deviation, skew and kurtosis of the tilt angle TA_i between the gravitational vector and the y-axis. The tilt angle is defined as:

$$TA_i = \sin^{-1}\left(\frac{y_i}{\sqrt{{x_i}^2 + {y_i}^2 + {z_i}^2}}\right)$$

where x, y and z is the acceleration in the respective axis.

- 10 features: mean, standard deviation, minimum, maximum, difference between maximum and minimum, entropy of the energy in 10 equal sized blocks, short time energy, spectral roll off, zero crossing rate and spectral flux from the magnitude of the acceleration vector.

- 4 features: the absolute of kurtosis of each axis (x, y, z) of the acceleration and the absolute of kurtosis of the tilt angle TA_i between the gravitational vector and the y-axis.

Feature set B

The second feature set consists of 37 features in total. It includes the features described in the feature set A minus the kurtosis of each axis (x, y, z) of the acceleration.

Feature set C

The third feature set consists of 33 features in total. It includes the features described in the feature set B minus the four features of the absolute.

Feature set D

The last feature set consists of 36 features in total. It includes the features described in the feature set B minus the absolute of kurtosis of the tilt angle TA_i.

4.3. Classification & evaluation

For the classification process the k-Nearest Neighbor (k-NN) classifier, with 1 nearest neighbor, and the C4.5 Decision Tree (DT) were used. Both of them achieved high performance in the specific application, as reported in [4]. The implementation of the algorithms was performed using the WEKA tool [96], where k-NN is referred to as IBk and the C4.5 DT as J48. The evaluation methods that applied were: 1) the n-fold-cross validation (with n equal to 10) which is the most commonly used approach and 2) the Leave One subject Out (LOO) validation, which is the most realistic scenario to simulate real life conditions.

5. Results

In the experimental process 24 combinations of: a) window size (1, 1.5, 2 sec with 80% overlap); b) feature set (A, B, C, D, described in Section 4.2.3) and c) classification algorithm (k-NN, C4.5 DT) using the 10-fold-cross validation method, were performed for the recognition of both ADLs (Table 5.1) and Scenarios (Table 5.2). Consequently, the two best combinations were further evaluated with the deployment of the Leave-One-subject-Out validation method for the recognition of scenarios. The classification results for the recognition of ADLs for all the deployed combinations show an accuracy exceeding 98%, as shown in Table 5.1.

Although the differences do not significantly vary, the best accuracy is shown in the combination of using a 2 seconds-window with 80% overlap, using the feature set A, B or C and the k-NN classifier. Despite the high accuracy, none of these combinations can consider as optimal since short activities have been overlooked, as shown in the confusion matrices (Table 5.3 and Table 5.4). This finding is encountered in both window sizes of 1.5 and 2 seconds. On the contrary, the confusion matrices of the combinations using 1 second-window with 80% overlap, the feature set A or B and the k-NN classifier suggest the best results despite the slightly lower accuracy. In particular when using window of size 2 seconds, 214 instances of the activity SCH and 4 instances of the activity CHU are obtained while with the use of 1 second window, 1992 and 447 instances of the activities SCH and CHU, respectively, are obtained. The recognition of short activities is crucial since they represent transitions of long activities, in terms of duration.

Subsequently, the recognition of scenarios is performed using the same protocol of the combined parameters. The results indicate a slight reduction in the accuracy, with all the deployed combinations exceeding 91% accuracy, as shown in Table 5.2. The combinations of 1 second-window with 80% overlap, the feature set A or B and the k-NN classifier achieved the best accuracy of 95.9553% and 95.9198%, respectively. These findings confirm the aforementioned rejection of the two best classification results in terms of accuracy for the ADLs recognition. Nevertheless, the accuracy of the recognition of scenarios has been decreased compared to the accuracy of recognition of ADLs, using the two best combinations. This could be caused because of the lower number of data or due to the nature of the transition actions between activities.

Table 5.1 Classification results (% accuracy) for recognition of ADLs

Recognition of ADLs	**Feature set A**		**Feature set B**		**Feature set C**		**Feature set D**	
	k-NN	**C4.5 DT**	**k-NN**	**C4.5 DT**	**k-NN**	**C4.5 DT**	**k-NN**	**C4.5 DT**
Window 1 sec 80% overlap	**99.1166%**	98.2168 %	**99.1233 %**	98.2215 %	99.1108 %	98.2329 %	99.1163 %	98.201 %
Window 1.5 sec 80% overlap	99.368 %	98.8013 %	99.3652 %	98.8327 %	99.3601 %	98.8024 %	99.3714 %	98.8024 %
Window 2 sec 80% overlap	**99.5395 %**	99.1065 %	**99.5403 %**	99.1027 %	**99.5403 %**	99.131 %	**99.5364 %**	99.0844 %

Table 5.2 Classification results (% accuracy) for recognition of Scenarios

Recognition of Scenarios	**Feature set A**		**Feature set B**		**Feature set C**		**Feature set D**	
	k-NN	**C4.5 DT**	**k-NN**	**C4.5 DT**	**k-NN**	**C4.5 DT**	**k-NN**	**C4.5 DT**
Window 1 sec 80% overlap	**95.9553%**	91.6467%	**95.9198%**	91.7105%	95.8858%	91.8%	95.9099%	91.6637%
Window 1.5 sec 80% overlap	95.3301 %	91.6339 %	95.2853 %	91.7642 %	95.2447 %	91.7919 %	95.2874 %	91.6702 %
Window 2 sec 80% overlap	**95.1197 %**	91.7185 %	**95.0968 %**	91.7871 %	95.0454 %	91.65 %	95.1054 %	91.8185 %

Table 5.3 Confusion Matrix of k-NN, with feature set A and 2s Window (99.5395 %)

=== Confusion Matrix ===

a	b	c	d	e	f	g	h	i	j	k	<-- classified as
2	2	0	0	0	2	0	0	0	0	0	a = CHU
0	1095	240	0	0	8	0	0	1	0	0	b = CSI
0	216	1317	0	0	1	0	0	1	1	0	c = CSO
0	0	0	12445	5	0	0	0	0	0	9	d = JOG
0	0	0	18	12510	0	0	0	0	0	1	e = JUM
4	7	9	0	0	194	0	0	0	0	0	f = SCH
0	0	0	0	0	0	2743	2	0	0	0	g = SIT
0	0	0	0	0	0	6	44630	0	0	0	h = STD
0	0	0	0	0	0	0	0	4690	34	0	i = STN
0	0	0	0	0	0	0	0	34	5256	0	j = STU
0	0	0	0	0	0	0	1	0	0	45244	k = WAL

Table 5.4 Confusion Matrix of k-NN, with feature set B and 2s Window (99.5403 %)

=== Confusion Matrix ===

a	b	c	d	e	f	g	h	i	j	k	<-- classified as
2	2	0	0	0	2	0	0	0	0	0	a = CHU
0	1094	241	0	0	8	0	0	1	0	0	b = CSI
0	218	1315	0	0	1	0	0	1	1	0	c = CSO
0	0	0	12445	5	0	0	0	0	0	9	d = JOG
0	0	0	19	12509	0	0	0	0	0	1	e = JUM
4	7	8	0	0	195	0	0	0	0	0	f = SCH
0	0	0	0	0	0	2743	2	0	0	0	g = SIT
0	0	0	0	0	0	6	44630	0	0	0	h = STD
0	0	0	0	0	0	0	0	4692	32	0	i = STN
0	0	0	0	0	0	0	0	33	5257	0	j = STU
0	0	0	0	0	0	0	0	0	0	45245	k = WAL

Table 5.5 Confusion Matrix of k-NN, with feature set A and 1s Window (99.1166%)

=== Confusion Matrix ===

a	b	c	d	e	f	g	h	i	j	k	<-- classified as
336	3	2	0	0	106	0	0	0	0	0	a = CHU
6	3760	479	0	0	27	0	0	3	10	0	b = CSI
1	410	4243	0	0	13	0	0	10	11	0	c = CSO
0	1	0	25533	56	0	0	0	6	5	145	d = JOG
0	0	0	112	25744	0	0	0	2	4	21	e = JUM
100	23	23	0	0	1846	0	0	0	0	0	f = SCH
0	0	0	0	0	0	5558	18	0	0	0	g = SIT
0	0	0	0	0	0	13	89527	0	0	1	h = STD
0	2	0	1	2	0	0	0	10698	381	12	i = STN
0	1	4	3	0	0	0	0	364	11853	9	j = STU
0	0	0	11	0	0	0	4	0	0	90753	k = WAL

Table 5.6 Confusion Matrix of k-NN, with feature set B and 1s Window (99.1233%)

=== Confusion Matrix ===

a	b	c	d	e	f	g	h	i	j	k	<-- classified as
335	4	2	0	0	106	0	0	0	0	0	a = CHU
6	3762	483	0	0	24	0	0	1	9	0	b = CSI
2	414	4239	0	0	13	0	0	9	11	0	c = CSO
0	1	0	25529	56	0	0	0	6	5	149	d = JOG
0	0	0	115	25739	0	0	0	2	4	23	e = JUM
97	22	23	0	0	1850	0	0	0	0	0	f = SCH
0	0	0	0	0	0	5563	13	0	0	0	g = SIT
0	0	0	0	0	0	11	89529	0	0	1	h = STD
0	1	0	1	1	0	0	0	10715	368	10	i = STN
0	1	2	2	0	0	0	0	366	11854	9	j = STU
0	0	0	10	0	0	0	4	0	0	90754	k = WAL

Given that the best overall results are obtained with the use of k-NN, feature set A or B, 1s window with 80% overlap and 10-fold cross-validation as an evaluation method, a further investigation was made using the aforementioned combinations and the evaluation method of Leave-One-subject-Out for the recognition of scenarios. The average accuracy using the feature set A reached 78.9987% while using the feature set B an accuracy of 79.5350% was achieved, as shown in Table 5.7. It is noticeable that the accuracy was reduced about 15% only by changing the evaluation method with a most realistic one, the LOO. Furthermore, a wide range of classification accuracy is observed with the lowest one to be at 41.1994 % and the highest one to reach 90.8401 %. The low accuracy findings might not have been observed if more data were available. Nevertheless, the findings suggest that more work should be made using methods that simulate real life conditions. In Figures 5.1-5.3, the actual and the classified activities of the scenarios are shown, using the LOO method for three indicative participants were the classification accuracy reached low, middle and high levels. As a general observation it can be stated that the activities that are misclassified mostly are the STN and the STU.

Table 5.7 Classification results (% accuracy) for recognition of Scenarios using LOO method

Leave-One-subject-Out	**k-NN, Feature set A, window 1s -80% overlap**	**k-NN, Feature set B, window 1s -80% overlap**
1st subject (sub1)	84.5677 %	84.6755 %
2nd subject (sub2)	84.7791 %	84.806 %
3rd subject (sub3)	74.8855 %	75.0741 %
4th subject (sub5)	59.6387 %	59.0725 %
5th subject (sub6)	41.1994 %	54.3143 %
6th subject (sub12)	90.8401 %	90.7046 %
7th subject (sub20)	66.5051 %	66.2359 %
8th subject (sub45)	73.3721 %	72.7209 %
9th subject (sub53)	79.3317 %	79.4395%
10th subject (sub58)	78.9547 %	78.1882%
11th subject (sub59)	88.1683 %	88.331%
12th subject (sub60)	84.0022 %	83.8944 %
13th subject (sub61)	87.4652 %	87.0474 %
14th subject (sub62)	71.3978 %	71.0746 %
15th subject (sub63)	88.3515 %	88.026%
16th subject (sub64)	90.5471 %	90.6284 %
17th subject (sub65)	87.0443 %	87.2069%
18th subject (sub66)	85.5295 %	85.26%
19th subject (sub67)	84.3953 %	84.4651%
Average values	**78.9987%**	**79.5350%**

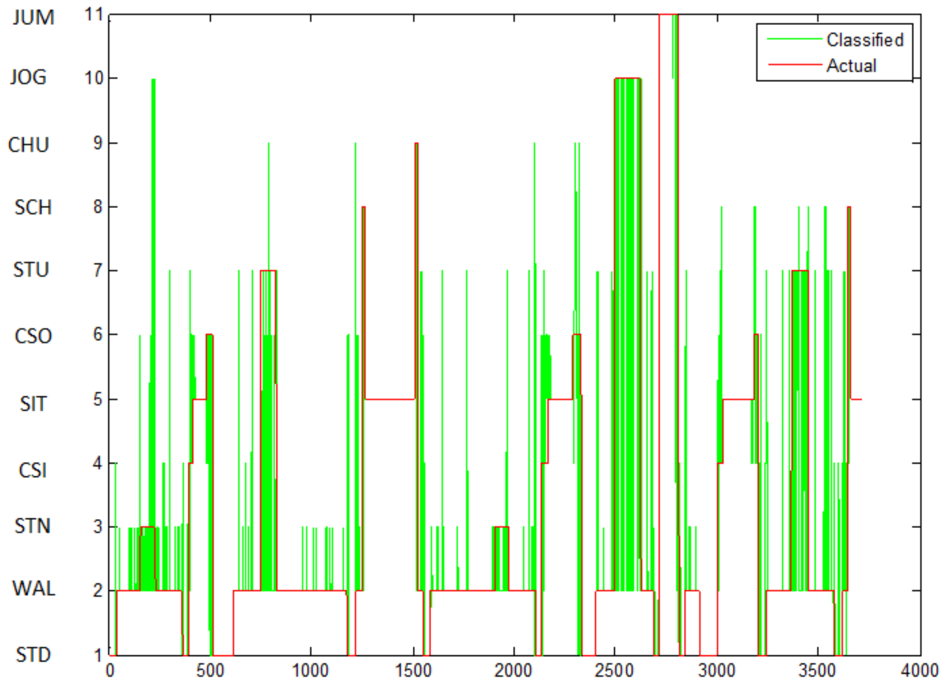

Figure 5.1 Visualization of the results obtained with the LOO method (sub1) and Feature set B

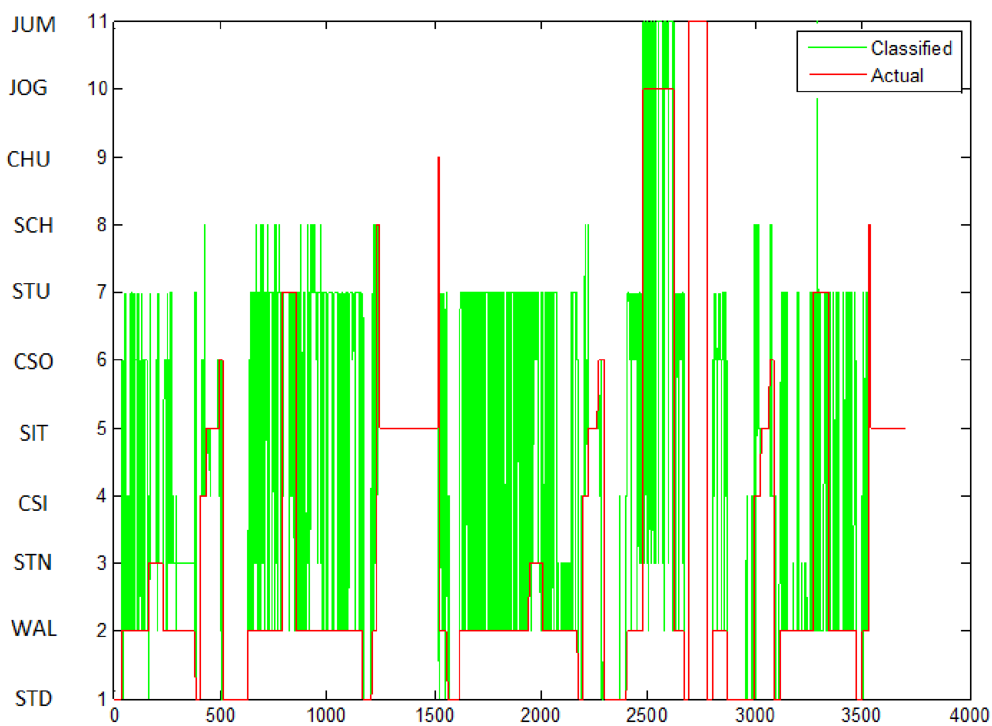

Figure 5.2 Visualization of the results obtained with the LOO method (sub6) and Feature set B

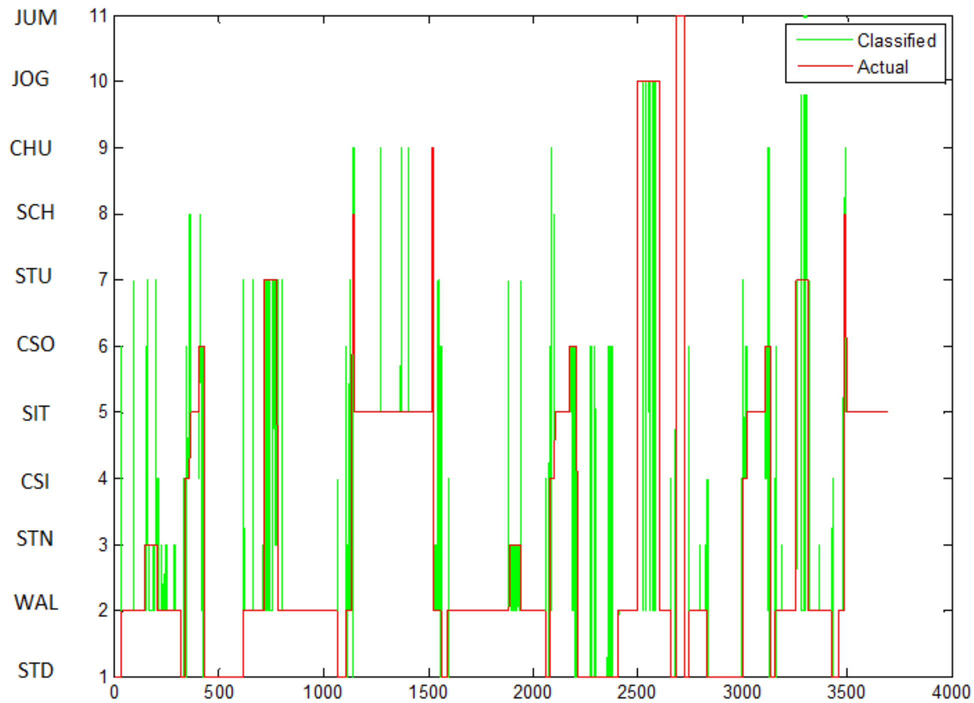

Figure 5.3 Visualization of the results obtained with the LOO method (sub64) and Feature set B

Finally, an unripe effort to check whether the developed computational method could be used to recognize the reordered scenarios (test set) based on the knowledge of the separate ADLs (train set), was performed. The accuracy of the recognition was further reduced, from the average results of the scenarios recognition using LOO, by 22%.

Table 5.8 Classification results (% accuracy) for Training: ADLs Testing: Scenarios

Train: ADLs Test: Scenarios	Combination with Feature set A	Combination with Feature set B
Train: All ADLs Test: All Scenarios	56.6315 %	56.8289%
Train: All ADLs Test: Scenario with highest accuracy (sub64)	65.2492%	64.9242%
Train: All ADLs Test: Scenario with middle accuracy (sub59)	59.0656%	58.3217%
Train: All ADLs Test: Scenario with lowest accuracy (sub6)	52.4209%	52.4479%

6. Conclusion

For recognizing-estimating consecutive human activity patterns, the first step is to accurately and effectively recognize separate activities of daily living. The activities of daily living can reveal valuable information about the person's activity patterns and subsequently, with the association of other behavioral characteristic elements, they can lead in the development of behavioral models. The use of smartphone as a sensor device is a powerful solution as it is widely used in everyday life and in dynamically changing environments.

The need of interfering more realistic conditions in the computational methodology is revealed via the differences of the accuracy when using the Leave-One-subject-Out method for the recognition of scenarios. The deployment of a hybrid approach using knowledge base techniques or event analysis techniques or additional sensors such as Radio Frequency Identification tags, for importing spatiotemporal aspects could be the lead way for the development of a system that can accurately estimate human activity patterns in real-life scenarios.

It has been obvious through this study that complex activities or behaviors are ultimately a sequence of events with specific spatiotemporal characteristics. It seems obvious to that exploring event-based processing methods could be a promising avenue, in our attempt to design a system, based on smartphone acquired data alone, to recognize complex, dynamic everyday behaviors tuned for the needs of elders and specific patient categories.

6.1.Future work - exploring event based processing methods

An event is simply something that happens, or contemplated as happening [97]. In particular any information that can carried along with time interval, regardless the source of information (sensor signal, video data, and GPS coordinate) can characterized as an event. A detailed study about the event recognition have been contacted by Skarlatidis [98], according to that study, the most important property of an event is that it occurs in a period of time, instantaneous or during some interval of time, independently of the carried information.

The single information represented by one event can be related with other events in various ways, e.g., temporally, spatially, and causally. Moreover, related events produce event patterns that can be combined with other unrelated events, in respect with the domain of application. This practice is commonly used in the domain of video surveillance, where the events representing that two or more people are undertaking physical activities at the same time (temporal relation), in specific distance and direction (spatial relations), could indicate the activity pattern of moving together (event pattern) [98].

Recognition systems that make use of multiple sensors for the deployment of events (event pattern detection systems) can be used to monitor an environment and respond to the occurrence of significant events. Examples of such application domains are health care monitoring, public transport management, telecommunication, network monitoring, credit card fraud detection and activity recognition [99], [100], [101].

The preceding short introduction to event-based processing methods, in our view, confirm their applicability in the problem domain discussed and verify our intuition regarding their potential, thus confirming our suggestion for future work along these lines of investigation.

Bibliography

[1] J. M. Wiener, R. J. Hanley, R. Clark and J. F. Van Nostrand, "Meauring the Activities of Daily Living: Comparisons Across National Surveys," *Journal of Gerontology: SOCIAL SCIENCES,* vol. 45, no. 6, pp. 229-237, 1990.

[2] J. L. R. Ortiz, Smartphone-Based Human Activity Recognition, Springer, 2015.

[3] L. Chen , J. Hoey, C.D. Nugent, D.J. Cook, and Yu Zhiwen, "Sensor-Based Activity Recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews,* vol. 42, no. 6, pp. 790-808, 2012.

[4] G. Vavoulas, C. Chatzaki, Th. Malliotakis, M. Pediaditis, and M. Tsiknakis, "The MobiAct Dataset: Recognition of Activities of Daily Living using Smartphones," in *2nd International Conference on Information and Communication Technologies for Ageing Well and e-Health*, Rome, Italy, 2016.

[5] C.-Y. Chen, Y.-H. Chen, C.-F. Lin,C.-J. Weng, and H.-C. Chien, "A Review of Ubiquitous Mobile Sensing Based on Smartphones," *International Journal of Automation and Smart Technology,* vol. 4, no. 1, 2012.

[6] B. Gerritsen and I. Horváth, "The Upcoming and Proliferation of Ubiquitous Technologies in Products and Processes," *Mechanical Engineering,* vol. 56, no. 11, pp. 765-783, 2010.

[7] S. Patel, H. Park, P. Bonato, L. Chan and M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation," *Journal of NeuroEngineering and Rehabilitation,* 2012.

[8] E. Jovanov, A. Milenkovic, C. Otto and P. C de Groen, "A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation," *Journal of NeuroEngineering and Rehabilitation,* 2005.

[9] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou and Y. Amirat, "Physical Human Activity Recognition Using Wearable Sensors," *Sensors,* vol. 15, p. 31314–31338, 2015.

[10] F. Xia, C.-H.Hsu, X. Liu, H. Liu, F. Ding, and W. Zhang, "The Power of Smartphones," *Multimedia Systems,* vol. 21, no. 1, pp. 87-101, 2015.

[11] M. Liu, "A Study of Mobile Sensing Using Smartphones," *International Journal of Distributed Sensor Networks,* vol. 2013, p. 11 pages, 2013.

[12] X. Su, H. Tong, and P. Ji, "Activity Recognition with Smartphone Sensors," *TSINGHUA SCIENCE AND TECHNOLOGY*, vol. 19, no. 3, pp. 235-249, 2014.

[13] G. Vavoulas, M. Pediaditis, C. Chatzaki, E. G. Spanakis and M. Tsiknakis, "The MobiFall Dataset: Fall Detection and Classification with a Smartphone," *International Journal of Monitoring and Surveillance Technologies Research (IJMSTR)*, vol. 2, no. 1, pp. 44-56, 2014.

[14] S. Dernbach, B. Das, N. C. Krishnan, B.L. Thomas and D.J. Cook , "Simple and Complex Activity Recognition through Smart Phones," *8th International Conference on Intelligent Environments (IE)*, pp. 214-221, 26-29 June 2012.

[15] P. Siirtola and J. Roning, "Ready-to-use activity recognition for smartphones," *IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, pp. 59-64, 16-19 April 2013.

[16] A. Bayat, M.Pomplun and D. A. Tran, "A Study on Human Activity Recognition Using Accelerometer Data from Smartphones," *11th International Conference on Mobile Systems and Pervasive Computing (MobiSPC'14)*, vol. 34, pp. 450-457, 2014.

[17] A. Anjum and M. U. Ilyas, "Activity recognition using smartphone sensors," *Consumer Communications and Networking Conference (CCNC)*, pp. 914-919, 11-14 January 2013.

[18] L. Fan, Z. Wang and H. Wang, " Human Activity Recognition Model Based on Decision Tree," *Proceedings of the 2013 International Conference on Advanced Cloud and Big Data (CBD '13)*, pp. 64-68, 2013.

[19] E. Buber and A.M. Guvensan, "Discriminative time-domain features for activity recognition on a mobile phone," *IEEE 9th International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, pp. 1-6, 21-24 April 2014.

[20] L. Zheng, Y. Cai, Z. Lin, W. Tang, H. Zheng, H.Shi, B. Liao, and J. Wang, "A Novel Activity Recognition Approach Based on Mobile Phone," *Multimedia and Ubiquitous Engineering*, pp. 59-65, 2014.

[21] T. R. D. Saputri, A. M. Khan, and S.-W. Lee, "User-Independent Activity Recognition via Three-Stage GA-Based Feature Selection," *International Journal of Distributed Sensor Networks*, p. 15, 2014.

[22] X.Su, H.Tong and P.Ji, "Accelerometer-based Activity Recognition on Smartphone," in *CIKM '14 Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, New York, NY, USA, 2014.

[23] J. R. Kwapisz, G. M. Weiss and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explorations Newsletter*, vol. 2, no. 2, pp. 74-82, 31 03 2011.

[24] M. Han, J. H. Bang, C. Nugent , S. McClean and S. Lee, "A Lightweight Hierarchical Activity Recognition Framework Using Smartphone Sensors," *Sensors*, vol. 14, no. 9, pp. 16181-16195, September 2014.

[25] W. Hung, F. Shen, Y.L. Wu, M.K. Hor and C.Y. Tang, "Activity Recognition with Sensors on Mobile Devices," in *Proceedings of the 2014 International Conference on Machine Learning and Cybernetics*, Lanzhou, 2014.

[26] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine," in *Proceedings 4th International Workshop, IWAAL*, Vitoria-Gasteiz, Spain, 2012.

[27] D. Anguita, A. Ghio, L. Oneto, X. Parra and J. L. Reyes-Ortiz, "A Public Domain Dataset for Human Activity Recognition Using Smartphones," in *ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence*, Bruges, Belgium, 2013.

[28] A.M. Khan, Y. K. Lee, S.Y.Lee and T. S. Kim, "2010 5th International Conference on Future Information Technology (FutureTech)," Busan, 2010.

[29] V. N. T.- Sang, N. D. Thang, V. V. Toi, N. D. Hoang and T. Q. D.- Khoa, "Human Activity Recognition and Monitoring Using Smartphones," in *5th International Conference on Biomedical Engineering,IFMBE Proceedings 46*, Vietnam, 2015.

[30] A. Mannini and A. M. Sabatini, "Machine Learning Methods for Classifying Human Physical Activity from On-Body Accelerometers," *Sensors*, vol. 10, no. 2, pp. 1154-1175, February 2010.

[31] L. Bao and S. S. Intille, "Activity Recognition from User-Annotated Acceleration Data," in *Proceedings Second International Conference, PERVASIVE*, Linz/Vienna, Austria, 2004.

[32] J. J. Guiry, P. Ven, J. Nelson, L. Warmerdam, and H. Riper, "Activity recognition with smartphone support," *Medical Engineering & Physics*, vol. 36, no. 6, pp. 670-675, June 2014.

[33] P.Gupta and T. Dallas, "Feature Selection and Activity Recognition System Using a Single Triaxial Accelerometer," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 6, pp. 1780 - 1786, May 2014.

[34] F. Torre, J. Hodgins, A. Bargteil, X. Martin, J. Macey, A. Collado, P. Beltran, "Guide to the Carnegie Mellon University Multimodal Activity (CMU-MMAC) Database," 2008.

[35] J. M. Chaquet, E. J. Carmona, A. F.-Caballero, "A survey of video datasets for human action and activity recognition," *Computer Vision and Image Understanding,* vol. 117, pp. 633-659, February 2013.

[36] A. M. J. Sarkar, "An Intelligent Tool for Activity Data Collection," *Sensors*, vol. 11, pp. 3988-4008, April 2011.

[37] Z. S. Abdallah, M. M. Gaber, B. Srinivasan and S.i Krishnaswamy, "Adaptive mobile activity recognition system," *Neurocomputing*, vol. 150, October 2014.

[38] C. McCall, K. Reddy and M. Shah, "Macro-Class Selection for Hierarchical K-NN Classification of Inertial Sensor Data," in *Second International Conference on Pervasive and Embedded Computing and Communication Systems, PECCS 2012*, Rome, Italy, 2012.

[39] F. Miao, Y. He, J. Liu, Y. Li and I. Ayoola, "Identifying typical physical activity on smartphone with varying positions and orientations," *BioMedical Engineering OnLine*, 2015.

[40] M. Shoaib, S. Bosch , O. D. Incel , H.Scholten and P. J. M. Havinga, "Fusion of Smartphone Motion Sensors for Physical Activity Recognition," *Sensors,* vol. 14, pp. 10146-10176, June 2014.

[41] L. Chen, C. Nugent, and G. Okeyo,, "An Ontology-based Hybrid Approach to Activity Modeling for Smart Homes," *IEEE Transactions on Human-Machine Systems,* vol. 44, no. 1, pp. 92-105, 2014.

[42] V. S.-Mendoza, J. Beltran, E. Chavez, J. Hernandez and J. A. G.-Macıas, "Abnormal Behavioral Patterns Detection from Activity Records of Institutionalized Older Adults," in *Proceedings, 6th International Workshop Human Behavior Understanding*, Osaka, Japan, 2015.

[43] X. Yu and T. Korkmaz, "Evident Behavior Analysis on ADLs," in *20th ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, NY, USA., 2014.

[44] F. J.Ordonez , P. Toledo and A. Sanchis, "Activity Recognition Using Hybrid Generative/Discriminative Models on Home Environments Using Binary Sensors," *Sensors* , vol. 13, no. 5, pp. 5460-5477, 2013.

[45] L. Pei, R. Guinness, R. Chen , J. Liu, H. Kuusniemi, Y. Chen, L. Chen and J. Kaistinen, "Human Behavior Cognition Using Smartphone Sensors," *Sensors*, vol. 13, pp. 1402-1424, 2013.

[46] T. Dimitrov, J. Pauli and E. Naroska, "Unsupervised Recognition of ADLs," in *Proceedings, 6th Hellenic Conference on Artificial Intelligence*, Athens, Greece, 2010.

[47] K. Finkenzeller, RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification, Second Edition, John Wiley & Sons, Ltd, 2013.

[48] K. S. Gayathri, S. Elias and S. Shivashankar, "An Ontology and Pattern Clustering Approach for Activity Recognition in Smart Environments," in *Proceedings of the Third International Conferenceo n Soft Computing for Problem Solving, Advances in Intelligent Systems and Computing*, India, 2014.

[49] P.Rashidi, D. J. Cook, L. B. Holder, and M. S.-Edgecombe, "Discovering Activities to Recognize and Track in a Smart Environment," *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING,* vol. 23, no. 4, 2011.

[50] F. Amirjavid , A. Bouzouane,and B. Bouchard, "Activity modeling under uncertainty by trace of objects in smart homes," *Journal of Ambient Intelligence and Humanized Computing,* vol. 5, no. 1, pp. 159-167, 2014.

[51] B. Zhou, Q. Li, Q.Mao, W. Tu and X. Zhang, "Activity Sequence-Based Indoor Pedestrian Localization Using Smartphones," *IEEE Transactions on Human-Machine Systems,* vol. 45, no. 5, pp. 562-574, 2015.

[52] B. Fida, I. Bernabucci, D. Bibbo, S. Conforto and M. Schmid, "Pre-Processing Effect on the Accuracy of Event-Based Activity Segmentation and Classification through Inertial Sensors," *Sensors*, vol. 15, pp. 23095-23109, 2015.

[53] N. C. Krishnan and D. J. Cook, "Activity Recognition on Streaming Sensor Data," *Pervasive and Mobile Computing*, vol. 10, pp. 138-154, 2014.

[54] O.i Banos, J.-M.l Galvez, M.l Damas, A. Guillen,L.-J. Herrera, H. Pomares, and I. Rojas, "Evaluating the effects of signal segmentation on activity recognition," in *Proceedings IWBBIO* , Granada, 2014.

[55] O. Banos, J.-M.l Galvez, M. Damas, H. Pomares and I. Rojas, "Window Size Impact in Human Activity Recognition," *Sensors*, vol. 14, pp. 6474-6499, 2014.

[56] S. D. Bersch, Djamel Azzi, R. Khusainov , I. E. Achumba and J. Ries, "Sensor Data Acquisition and Processing Parameters for Human Activity Classification,"

Sensors, vol. 14, pp. 4239-4270, 2014.

[57] S. Kozina, M. Luštrek, and M. Gams, "Dynamic Signal Segmentation for Activity Recognition," in *International Joint Conference on Artificial Intelligence (IJCAI)*, Barcelona, Spain, 2011.

[58] M. H. M. Noor, Z. Salcic, and K. I-K. Wang, "Dynamic Sliding Window Method for Physical Activity Recognition Using a Single Tri-axial Accelerometer," in *IEEE 10th Conference on Industrial Electronics and Applications (ICIEA)*, Auckland, 2015.

[59] W. Dargie, "Analysis of Time and Frequency Domain Features of Accelerometer Measurements," in *Proceedings of 18th Internatonal Conference on Computer Communications and Networks*, San Francisco, CA, 2009.

[60] S.J. Preece, J.Y Goulermas, L.P.J Kenney, .and D. Howard, "A Comparison of Feature Extraction Methods for the Classification of Dynamic Activities From Accelerometer Data," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 3, pp. 871-879, 2009.

[61] D.Figo, P. C. Diniz and D. R. Ferreira, "Preprocessing techniques for context recognition from accelerometer data," *Personal and Ubiquitous Computing*, vol. 14, no. 7, p. 645–662, 2010.

[62] D. Acharjee, A. Mukherjee, J. K. Mandal, and N.i Mukherjee, "Activity recognition system using inbuilt sensors of smart mobile phone and minimizing feature vectors," *Microsystem Technologies*, pp. 1-8, 2015.

[63] N.A. Capela, E.D. Lemaire, and N. Baddour, "Feature Selection for Wearable Smartphone-Based Human Activity Recognition with Able bodied, Elderly, and Stroke Patients," *PLoS ONE*, vol. 10, no. 4, 2015.

[64] M. Zhang and A. A. Sawchuk, "A Feature Selection-Based Framework for Human Activity Recognition Using Wearable Multimodal Sensors," in *Proceedings of the 6th International Conference on Body Area Networks*, Brussels, Belgium, 2011.

[65] E. Sejdic, K.A. Lowry, J. Bellanca, M.S. Redfern, and J.S. Brach, "A comprehensive assessment of gait accelerometry signals in time, frequency and time-frequency domains.," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 3, pp. 603 - 612, 2013.

[66] Z. He, "Activity recognition from accelerometer signals based on Wavelet-AR model," in *IEEE International Conference on Progress in Informatics and Computing (PIC)*, Shanghai, 2010.

[67] C.V. Bouten, K.T. Koekkoek, M.Verduin, R. Kodde,and J.D. Janssen, , "A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity," *IEEE Trans. Biomed. Eng.*, vol. 44, p. 136–147, 1997.

[68] K. Altun, B. Barshan,and O. Tunçel, "Comparative study on classifying human activities with miniature inertial and magnetic sensors," *Pattern Recognition*, vol. 43, p. 3605–3620., 2010.

[69] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Computers and Electrical Engineering*, vol. 40, pp. 16-28, 2014.

[70] J.Tang, S. Alelyani, and H. Liu, "Feature Selection for Classification: A Review," in *Data Classification Algorithms and Applications*, Charu C . Aggarwal, Chapman and Hall/CRC 2014, 2014, p. 37–64.

[71] N. Sánchez-Maroño, A. Alonso-Betanzos, and M. Tombilla-Sanromán, "Filter Methods for Feature Selection – A Comparative Study," in *Proceedings of the 8th international conference on Intelligent data engineering and automated learning*, 2007.

[72] Kira, K.,and Rendell, L., "A practical approach to feature selection," in *Proceedings of*, 1992.

[73] M. Hall, "Correlation-based Feature Selection for Machine Learning," in *PhD thesis, University of Waikato*, Hamilton, New Zealand, 1999.

[74] Yu, L.,and Liu, H, "Feature selection for high-dimensional data: A fast correlation based filter solution," in *Proceedings of The Twentieth International Conference on Machine Learning*, 2003.

[75] Zhao, Z., and Liu, H., "Searching for interacting features," in *Proceedings of International Joint Conference on Artificial Intelligence*, 2007.

[76] I. Guyon and A. Elisseeff, "An Introduction to Variable and Feature Selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157-1182, 2003.

[77] R. Kohavi, and G. H. Johnb,, "Wrappers for feature subset selection," *Artificial Intelligence*, vol. 97, no. 1-2, pp. 273-324, 1997.

[78] R. O. Duda, P.E. Hart and D.G.Stork, Pattern Classification, Canada: John Wiley & Sons, Inc., 2001.

[79] T. O. Ayodele, "Types of Machine Learning Algorithms," in *New Advances in Machine Learning*, INTECH, Science, Technology and Medicine open access publisher, 2010.

[80] C. C. Aggarwal, Data Classification: Algorithms and Applications, Chapman and Hall/CRC , 2014.

[81] R. O. Duda, P.E. Hart and D.G. Stork, Pattern Classification, second edition, New York: John Wiley & Sons, Inc, 2001.

[82] R. Caruana and A. Niculescu-Mizil , "An empirical comparison of supervised learning algorithms," in *ICML '06 Proceedings of the 23rd international conference on Machine learning*, New York, NY, USA, 2006.

[83] S.Saeedi and N. El-Sheimy, "Activity Recognition Using Fusion of Low-Cost Sensors on a Smartphone for Mobile Navigation Application," *Micromachines,* vol. 6, pp. 1100-1134, 2015.

[84] Media A. Ayua, S.i Aisyah Ismaila, A. F. Abdul Matina and T. Mantoro, "A Comparison Study of Classifier Algorithms for Mobile-phone's Accelerometer Based Activity Recognition," *International Symposium on Robotics and Intelligent Sensors 2012 (IRIS 2012),* vol. 41, pp. 224-229, 2012.

[85] Fikri Azli bin Abdullah, Mohd; Fahmi Perwira Negara, Ali; Shohel Sayeed, Md.; Deok-Jai Choi; Sonai Muthu, Kalaiaras, "Classification Algorithms in Human Activity Recognition using Smartphones," *World Academy of Science, Engineering & Technology,* no. 68, p. 422, 2012.

[86] A. K. Jain, R. P. W. Duin and J. Mao, "Statistical Pattern Recognition: A Review," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 22, no. 1, pp. 4-37, 2000.

[87] S. B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques," *Informatica,* vol. 31, pp. 249-268, 2007.

[88] T. Hastie, R. Tibshirani and J. Friedman, The Elements of Statistical Learning, Second Edition, New York: Springer, 2013.

[89] T. Hastie, R. Tibshirani and J. Friedman, "Unsupervised Learning," in *The Elements of Statistical Learning,* New York, Springer, 2009, pp. 485-585.

[90] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in *International Joint Conference on Artificial Intelligence,* 1995.

[91] P.Refaeilzadeh, L. Tang, and H. Liu, "Cross-Validation," in *Encyclopedia of Database Systems,* New York, Springer, 2009, pp. 532-538.

[92] Wil M. P. van der Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes, Berlin: Springer, 2011.

[93] M. Bramer, "Measuring the Performance of a Classifier," in *Principles of Data Mining*, Verlag London, Springer, 2013, pp. 175-187.

[94] M. B. del Rosario, S. J. Redmond and N. H. Lovell, "Tracking the Evolution of Smartphone Sensing for Monitoring Human Movement," *Sensors,* vol. 15, p. 18901–18933., 2015.

[95] G. Vavoulas, M. Pediaditis, E. G. Spanakis and M. Tsiknakis, "The MobiFall Dataset: An Initial Evaluation of Fall Detection Algorithms Using Smartphones," in *IEEE 13th International Conference on Bioinformatics and Bioengineering (BIBE)*, Chania, 2013.

[96] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. H. Witten, "The WEKA data mining software: an update,," *CM SIGKDD Explorations Newsletter,* vol. 11, no. 1, pp. 10-18, 2009.

[97] Luckham, D. and Schulte, R. , "EPTS Event Processing Glossary v2.0.," Technical report, 2011.

[98] A. SKARLATIDIS, "Event Recognition Under Uncertainty and Incomplete Data," University of Piraeus, 2014.

[99] Artikis, A., Skarlatidis, A., Portet, F., and Paliouras, G., "Logic-based event recognition," *Knowledge Engineering Review,* vol. 27, no. 4, p. 469–506, 2012b.

[100] Manfredotti, C., Hamilton, H., and Zilles, S., "Learning RDBNs for Activity Recognition," in *NIPS Workshop on Learning and Planning from Batch Time Series Data*, 2010.

[101] D. C. Luckham, "Event Processing for Business: Organizing the Real-Time Enterprise," in *Wiley*, 2011.