HUMAN ACTIVITY RECOGNITION FROM AN ACCELEROMETER ON THE CHEST: DATA TRANSFORMATION, FEATURE SELECTION, AND CLASSIFICATION PERFORMANCE

A Thesis

by

SEYED MOHAMMADREZA HOSSEINIAN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Mark Lawley
Co-Chair of Committee,	Farzan Sasangohar
Committee Members,	Ranjana Mehta
Head of Department,	Mark Lawley

December 2017

Major Subject: Industrial Engineering

ABSTRACT

The fourth risk factor for global mortality is lack of physical activity (PA). From the past to present, the relationship between public health and sedentary behavior or physical activity has been an interesting topic for scientists. In the past decade, use of accelerometers for recognizing PA has increased significantly. The aim of this thesis is build a new algorithm to recognize eight different static activities and seven dynamic activities from accelerometer data on the chest based on laboratory data.

To conduct this study, we used laboratory data which was collected from 30 healthy people. In order to extract required information for the analytical part, all activities were recorded in video files. After data collection, all activities were labeled. We used first order differencing to remove the effect of participant's characteristics. Median of angles and the area under the curve were considered as features and used as predictors in classifiers. We performed 81 different random-forest models to evaluate the effect of sample size and time window size in the accuracy of the model.

We achieved 98.2% accuracy in a random-forest model with 5000 sample size in 6 second time window. We found that there is a positive correlation between time window and sample size with accuracy of the random-forest model.

Also, we performed the Support Vector Machine (SVM) algorithm for same sample size and time window. The accuracy of the SVM model was 95.5%. Both models have reliable performance to recognize the activities in fifteen categories. In the next step, based on sedentary behavior and physical activities definitions, we combined some

categories and evaluated the performance of our models in the new categories. As a final result, we achieved 98.9% and 97.6% accuracy in seven different categories. The result of random-forest and SVM models demonstrate our features have provided well-separated data in each category. Future research is required to evaluate the performance of these models on the real-life data.

DEDICATION

To my best friend, and my wife, for all her unconditional love and support. I could not have accomplished this without you.

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my advisor Dr. Mark Lawley, my committee members Dr. Ranjana Mehta and Dr. Ferzan Sasangohar, and my mentor Dr. Madhav Erraguntla for their continuous guidance and support of my research, as well as their patience, motivation, enthusiasm, and immense knowledge.

Thanks also to my all my friends and colleagues and the department faculty and staff for making my time at Texas A&M University an amazing and memorable experience.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Professors Dr. Mark Lawley and Dr. Farzan Sasangohar of the Department of Industrial and Systems Engineering and Professor Dr. Ranjana Mehta of the Department of Environmental and Occupational Health.

All work for the thesis was completed by the student, under the advisement of Dr. Dr. Madhav Erraguntla of the Department of Industrial and Systems Engineering and the committee members.

Funding Sources

There are no outside funding contributions to acknowledge related to the research and compilation of this document.

NOMENCLATURE

PA	Physical Activity
SB	Sedentary Behavior
LD-R	Lying Down – Right Side of the Body
LD-L	Lying Down – Left Side of the Body
T-R	Tilting Right Side of the body (Without Movement)
T-L	Tilting Left Side of the body (Without Movement)
TRL	Tilting Right and Left
Walking1	Slow Walking
Walking2	Fast Walking
Bending-UD	Bending Up and Down
AUC	Area Under the Curve
FOD	First Order Differencing

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

The first chapter of this thesis provides background information concerning physical activity (PA), sedentary behavior (SB), the accelerometer device, and activity recognition. In the last two sections of this chapter, the motivation, goal, and methodology selected for this thesis are discussed.

I.1. Physical Activity and Sedentary Behavior

Physical activity is characterized as any bodily movement produced by muscles that require energy expenditure. The amount of energy required to perform an activity can be measured in kilocalories (Kcal) or kilojoules (kJ). One kilocalorie is equivalent to 4.184 kilojoules[1].

Physical activity can be classified differently. It can be measured in different segments of time, daily or weekly. The simple formulation can be performed to represent the caloric contribution of each level of activity. For instance, if we categorize the daily operations to sleep, occupation, and leisure the total energy expenditure due to physical activity can be computed by the formula below:

$$\text{Kcol}_{\text{(sleep)}} + \text{Kcal}_{\text{(occupation)}} + \text{Kcal}_{\text{(leisure)}} = \text{Kcal}_{\text{(Total daily physical activity)}}$$

One way to categorize the activities can be based on the intensity. We can categorize activities in light, moderate, and heavy intensity. In 1985, Caspersen et al. demonstrated that all these types of classification are acceptable to categorize the physical activity[1].

Physical activity is a very comprehensive concept which includes all operations that people perform in order to sustain life. People use a different level of energy to perform daily activities. For some activities, we need a high level of energy consumption, while for others, we need less. However, based on the definition of PA all operations of people are counted as physical activity, but there are other concepts related to a physical activity like sedentary behaviors which play an important role to understand the concept of PA.

Sedentary behavior (SB) is "any waking behavior characterized by an energy expenditure ≤1.5 METs while in a sitting or reclining posture" [2]. The measure of energy expenditure to perform activities is defined as Metabolic Equivalent of Task (MET)[3]. Figure 1 illustrates the relationship between daily activities and metabolic equivalent. We can see activities have been categorized in 4 different levels. Sedentary activity when the MET is less than 1.5, light activity when MET is between 1.5 and 3, moderate activity when the MET is between 3 and 6, and vigorous activity when the MET is greater than 6.

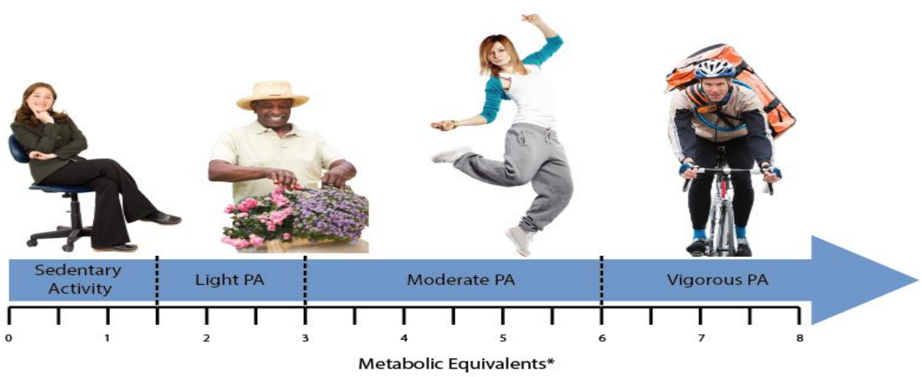


Figure 1. Relationship Between Physical Activity and Metabolic Equivalents Reprinted With Permission From [4]

The Harvard T.H. Chan School of Public Health categorized some activities in the same level which was explained. They calculated the METs for some daily activities and categorized the activities in four-levels: (1) sedentary, (2) light, (3) moderate, and (4) vigorous. Activities and METs are summarised in Table 1[5].

Table 1. Measuring Physical Activity

	Sedentary METs < 1.5	Light 1.5 < METs < 3.0	Moderate 3.0 < METs < 6.0	Vigorous METs > 1.5
1	Sitting (using computer) MET = 1.5	Walking (slowly) MET = 2.0	Walking (4mph) MET = 5.0	Hiking (4.5mph) MET = 7.0
2	Reclining (watching television) MET = 1.0	Standing (cooking, washing dishes) MET = 2-2.5	Bicycling (10–12 mph) MET = 6.0	Bicycling (14–16 mph) MET = 10.0
3	Lying down (not sleep) MET < 1	Playing most instruments MET = 2.0–2.5	Badminton (recreational) MET = 4.5	Carrying (heavy loads) MET = 7.5

From the table above we can see, there are some activities with the same name and different intensity which means we can not categorize the activities just based on their name. The ranges which are mentioned in Table 1 are related to healthy adults, and these ranges change if we want to categorize children or elderly activities. For instance, the metabolic equivalent of task threshold for sedentary behavior in children is considered as two (MET < 2) which is greater than adults[6]. This information guides researchers to choose best ranges based on their studies.

I.2. Accelerometer

Accelerometers are instruments that measure the acceleration which is the rate of velocity change of an object. The units for this measure is g-forces (g) or meters per second squared (m/s^2)[7]. Accelerometers can measure static and dynamic acceleration. Static acceleration means the constant force caused by gravity and dynamic acceleration means acceleration which is caused by moving or vibrating. By measuring static acceleration, we can find out the angle of the device with respect to the earth, and by measuring dynamic acceleration, we can find direction and intensity of movement.

The application of accelerometers has developed to multiple disciplines. Nowadays, we can find different types of accelerometers in various fields. Many electronic devices like smartphones, tablets, and cameras use the accelerometer to change the screen position based on phone direction, for example, if we want to watch movies or to read something on our phone/tablet, it is proper to see them in landscape view. The accelerometers can detect fall. Some companies use this property to protect the hard drive on their devices[8]. If we accidentally drop our device, the accelerometer detects the sudden fall and switch the hard disk off and prevent extra damage.

Accelerometers are used in cars to detect crashes and activate airbags[8]. Accelerometers can help to analyze the engine's problems using vibration testing. Applications of accelerometers are very vast, but the important question is which type of accelerometer is proper for which type of study? In other words, how should we select an accelerometer? Which feature of accelerometers should be considered when we want to conduct a study?

To answer these questions, we need to know the features which an accelerometer pakage presents. Some important features of the accelerometers are listed as[8]:

(1) Output: Some accelerometer provide analog outputs, and some generate digital outputs, and this will be specified by the hardware that we are interfacing the accelerometer with.

(2) Number of axes: Accelerometers can measure acceleration on one, two, or three axes.

(3) Output range (maximum swing): To gauge the acceleration of gravity for tilt sensing, we need low output range (±1.5 g), but if we want to use the accelerometer as impact sensor we need high output range (>±5 g). Impact sensors are designed to detect instances of sudden impact.

(4) Bandwidth: Means a number of times per second we can have acceleration reading. The unit of bandwidth is Hertz (Hz). For example, when the bandwidth of an accelerometer is 25 Hz it means it can record 25 estimates of acceleration during one second. For experiments which need to capture the motion or acceleration in a small bunches of time, we should select accelerators with high bandwidth.

Features mentioned above are the most important in choosing an accelerometer for a study. These features are common between all accelerometers but based on the nature of each study we should consider other factors to choose an accelerometer. For instance in high-temperature environments, we need to select some accelerometers which have been designed for that environment.

I.3. Activity Recognition

As stated by the World Health Organization (WHO), the fourth leading risk factor for global mortality is lack of physical activity. This risk factor is responsible for approximately 3.2 million deaths in the world[10]. Nowadays, it is well-known that lack of physical activity increases the chances of many adverse health outcomes such as diabetes, cardiovascular heart disease, and depression[10]. Furthermore, there is evidence that sedentary behavior (SB) increases the risk of chronic illnesses[11]. The Office of Disease Prevention and Health Promotion (ODPHP) has provided the Physical Activity Guidelines for Americans (PAG), which explains how children and adults could improve their health through physical activity[12]. Regular moderate intensity physical activity has considerable benefits for well-being.

From the past to present, the relationship between public health and sedentary behavior or physical activity has been an interesting topic for scientists. The Centers for Disease Control and Prevention (CDC) and the American College of Sports Medicine performed research in 1995 to evaluate the relationship between physical activity and public health. They found if Americans would adopt more active lifestyles instead of sedentary lives, public health would have fit enormously. An active lifestyle does not require a vigorous exercise program. They recommended that all US adults should perform at least thirty-minute moderate-intensity physical activity on most days (preferably all days) of the week[13].

There are some studies which considered the relationship between physical activities and a specific disease. Hu et al. compared the effect of vigorous physical activity vs. walking in risk of type 2 diabetes in women. In that study, all information about type and intensity of physical activities was collected by questionnaire. The first assessment was performed in 1986 and updated in 1988 and 1992. The result showed vigorous physical activity and walking are associated with a reduction in risk of type 2 diabetes in women[14].

In the past decade, use of accelerometers for recognizing PA has increased significantly [15,16,17,18,19,20]. Most of the studies used triaxial accelerometer which returns the estimation of acceleration along the x, y, and z-axis in units of gravity from which displacement and velocity can be estimated[11]. Many software and new devices have been developed for use in diverse fields. Most of these devices can classify activities by signal processing techniques and machine-learning algorithms in sedentary position (e.g., sitting), and movement position (e.g., moving slow, moving fast).

However, due to technology advancement, portable accelerometers have increasing potential for new applications for a wide variety of disease and conditions. For example, comparing the PA and SB levels in movement disorders patients (e.g., Parkinson disease) with healthy people in daily activities is an important topic. This kind of illness brings some limitation in people’s movement, and these situations cause more sedentary behavior in those patients’ life. As we mentioned before, evidence shows increasing sedentary behavior has a positive correlation with some chronic disease (e.g., diabetes, cardiovascular). Some studies have used the accelerometer data to identify sleep and waking patterns in infants[21]. Also, the correlation between PA and physical or mental fatigue can be measured if there is reliable algorithm to recognize activity through real-life data.

One important point in using the accelerometer to recognize the human activities is the position of the sensor on the body. Various studies have used accelerometers in different locations on the body to recognize activities.They put accelerometers on the ankle, wrist, and waist [16,17,22]. Some studies used more than one accelerometer to achieve their goals[23]. The point which should be considered is when we perform the activities some parts of the body have more movement and acceleration compared to other regions. For instance, in bicycling, we have more movement in the legs and we have more action on the downside of our body. Hence, if we put the accelerometer in some part of the body which has more movement, we will be able to capture more information.

This information makes it clear that, to recognize specific activities, the position of the sensor is one of the important features which should be considered. If we position a single accelerometer incorrectly, we should expect low performance of our model for certain kinds of activities.

In the current study, we used the single triaxial accelerometer in the chest position. More information about the aims of study and methodologies is provided in the next two sections.

I.4. Research Goals

The primary purpose of this study was to build an algorithm which can recognize eight different static activities and seven various dynamic activities from accelerometer data attached to the chest. This algorithm is constructed based on laboratory data on R software, an open source software, and programming language, for statistical computing and graphics.

The static activities we have considered include: (1) the supine position (lying face upward), (2) lying on the right side of the body (LD-R), (3) lying on the left side of the body (LD-L), (4) prone position, (5&6) tilting on the right (T-R) and tilting on the left (T-L) side of the body while standing, (7) bending forward, (8) standing position. Also, dynamic activities include: (9) twisting on the right and left, (10) tilting on the right and left (TRL), (11) bending forward and backward (Bending-UD), (12) squatting, (13) slow walking (Walking1), (14) fast walking (walking2), and (15) running.

The goals of this study are: (1) utilizing data transformation methods to remove the effect of individual characteristics from the model, (2) comparing the results of two different machine learning algorithms, SVM and random-forest which have reliable performance in nonlinear classes, (3) validating the best performing algorithm on a new data set to evaluate classification performance, (4) comparing the results of the model in the different time window and select the best time interval in order to conduct in final model, and (5) combining some activities as a general activity and evaluate the effect of this combination on accuracy of model.

I.5. Data Transformation, Feature Selection, and Classification Performance

In this section, we briefly explain the methodology and procedures which we used in the current thesis, and we will go into details of all process in Chapter III.

In this study, we used a First Order Differencing (FOD) transformation method which is very common in time series analysis. We used the concept of Area Under the Curve (AUC) to extract features from raw data. We selected the median of angles in each axis as another feature. To extract all features from raw data, we considered different time windows. As classifiers, we chose RandomForest and Support Vector Machine (SVM) which have reliable performance in non-linear classes. There are several studies which have selected these two classifiers to recognize different types of activities based on lab data and real-life data. Our models were validated by new data which have not been used in the training part, and the performance of the model in each category of activities was evaluated by the F-score formula. All results in details are given in different tables and figures in Chapter IV of this thesis.

CHAPTER II

LITERATURE REVIEW

A literature review was performed to examine previous research on Human Activity Recognition based on accelerometer data. Keywords such as "human activity recognition from accelerometer ", "detection of physical activity", " algorithm for activity recognition based on accelerometer data" were used in the NCBI databases and Google Scholar web search engine.

In 2009, Bonomi et al. conducted a study to develop a model for the detection of type and intensity of human activity using the single accelerometer[18]. They used the single accelerometer which had been mounted in the lower back. Twenty healthy people (13 men and 7 women) participated in their study for data collection.

All contributors performed different types of activities. They chose classes below as seven major classes of activities: (1) lying, (2) sitting, (3) standing, (4) dynamic standing (DS), (5) walking, (6) running, (7) cycling. The decision tree algorithm was used in this study to classify the activities. They used different segments of the acceleration signal to develop the trees in order to achieve the highest classification accuracy. Seven features were extracted in each segment for each axis (7×3) to perform the trees from the accelerometer data. The acceleration features in time domain were as follow: (1) average, (2) standard deviation, (3) pick-to-pick distance, and (4) cross-correlation between axes. Also, some features were computed in the frequency, (5) power spectral density, (6) amplitude of the spectral peak, and (7) frequency domain entropy. To compare the results of models in each step, F-scores, were computed. The highest classification accuracy that they have achieved was 93% by measuring acceleration features 6.4 or 12.8 second time window. The most important features which had a significant effect on the final model were the standard deviation in each axis and the cross-correlation between Y and Z axes.

In 2011, Gyllensten et al. performed the study to identify types of activity with a single accelerometer and evaluated the performance of the lab-based algorithm by real-life data[22]. They used two different devices in their study. The first instrument that they used was single triaxial accelerometer which was attached to the waist. The second device was IDEEA which is a multi-sensor activity recognition device.

The IDEEA consists of a data logger with five accelerometers: two mounted on the thighs, two on soles, and one on the upper sternum. Previous independent studies have shown that IDEEA has approximately 100% classification accuracy in laboratory-based data[24,25]. They considered the result of this device as a reference to classify the activities based on the first accelerometer. As static and dynamic activities, five classes below were considered: (1) lying down, (2) sitting/standing, (3) walking, (4) running, and (5) cycling.

Twenty people, ten men, and ten women contributed in this study. The sampling frequency for triaxial accelerometer was 20 Hz which means one sample per every 50 ms. The 6.4 second time window (128 sample) was chosen for feature extraction. They considered 113 different features which have been used in other studies. For instance, (1) mean, (2) standard deviation, (3) kurtosis, (4) skewness, (5) range, (6) cross-axis correlation, (7) accelerometer angle, (8) spectral energy in sub-bands (0–10 Hz in bands of 1.25 Hz), (9) spectral entropy, (10) peak frequencies, and (11) cross-spectral densities were computed as features. The authors performed four different algorithms: (1) decision tree, (2) neural network, (3) support vector machine (SVM), and (4) majority voting which consider the results of other three algorithms and classify the activities based on the majority of the predicted results. The F-score evaluated the performance of models. As a result, the best performing model was the majority voting model with 95.1% accuracy.

In 2014, Bayat et al. carried out the study to identify human activities by accelerometer data from smart phones[19]. They collected data from 4 subjects, two men, and two women who volunteered to participate in the research. The sampling rate of the accelerometer in this study was 100Hz; one sample per 10 ms. All participants accomplished six different tasks: (1) running, (2) slow walking, (3) fast walking, (4) aerobic dancing, (5) stairs-up, and (6) stairs-down. They used two different positions for smartphones: (1) in subjects hand, (2) in a pants pocket. They computed these features in 1.28 second time window for each axis: (1) average, (2) average of peak frequency (APF), (3) the variance of APF (VarAPF), (4) root mean square (RMS), (5) standard deviation, (6) the difference between maximum and minimum in each window (Minmax), (7) cross-correlation between axis. In total, they used 21(7×3) different features in their study. They performed 12 different types of algorithms as classifiers including SVM and Randomforest.

The overall accuracy for classification in their study was less than 90% in both positions of smartphones.

In most of literatures, scientist used the features which were extracted from the raw data in different type of classifiers. Also, they used accelerometer(s) in different positions on body and achived different level of accuracy based on posion of sensor(s)[26,27]. The most common machine learnnig algorithems which were used in these studies were support vector machine (SVM)[19,22,26] , random-forest[19,27,28], neural network (NN)[19,22], bayesian method[18], hidden markov models[29], and majority voting[22].

These information show both features and position of sensor have relationship with overal accuracy of models. In those studies, the effect of participants in the raw data was not considered while the subject’s characteristics like higth affect the row data.

CHAPTER III

METHODOLOGICAL DESIGN

This chapter describes the details of the data collection procedure, transformation method, features selection, and parameters used as predictors in classifiers. We will go through the details of each part in sections of this chapter. We will explain all steps used to build the model for activity recognition.

III.1. Subjects and Data Collection

To conduct this study, we used the laboratory data which was collected from 30 healthy people; 15 men and 15 women. Table 2 presents the participant's characteristics. Participants were recruited by convenience sampling, through sending an email to Texas A&M University society. All contributors gave written informed consent to participate in the study which was approved by Texas A&M University HUMAN SUBJECTS PROTECTION PROGRAM (IRB2017-0215D).

Table 2. Participant's Characteristics (Mean and Standard Deviation)

Parameters	Male (N= 15)	Female (N= 15)	All (N= 30)
Age (yr)	**25.41(4.84)**	**27.38 (7.59)**	**26.26 (6.15)**
Height (m)	**1.77 (0.08)**	**1.63 (0.05)**	**1.71 (0.10)**
Weight (LB)	**169.7 (23.26)**	**159.65 (43.29)**	**164.99 (33.12)**
BMI ($kg.m^{-2}$)	**24.46 (4.46)**	**27.05 (7.58)**	**25.58 (6.04)**

BMI: Body Mass Index

To perform the experiments, one research assistant helped the participants. In order to extract required information for the analytical part, all activities recorded in video files. Table 3 presents the static and dynamic activities that participants performed. All contributors performed at least 30 and 60 seconds for static and dynamic activities, respectively.

Table 3. Static and Dynamic Activities List

Static Activities	Dynamic Activities
1- Lying down in the supine position	9- Twisting on the right and left
2- Lying down on the left side of the body (LD-L)	10- Tilting on the right and left (TRL)
3- Lying down on the right side of the body (LD-R)	11- Bending forward and backward (Bending-UD)
4- Prone position	12- Squatting
5- Tilting on the right side of the body on standing position (T-R)	13- Slow walking
6- Tilting on the left side of the body on standing position (T-L)	14- Fast walking
7- Bending forward	15-Running
8- Standing	

As accelerometer devices, all participants wore the Equivital sensor which was placed on the specific belt on the left side of the chest. Equivital EQ02sensor (Hidalgo Ltd., Cambridge, UK) is a multi-parameter, ambulatory monitoring device. This device is small (78mm x 53mm x 10mm), lightweight (38gr), waterproof, with a battery life up to 24 hours. The sampling frequency for Equivital was confirmed to 25Hz; one sample on every 40 ms.

After data collection, all activity was labeled based on video file for all participants. During of this process, we found that (1) for one participant, the sensor has not collected the data and the file related to that person was empty, (2) two different people were not able to perform most of the activities correctly during the study. Those participants' data were excluded from dataset.

The data labeling was one of the most time-consuming parts of this study. In data labeling part, we considered the data for each activity and removed the data from transition part from one activity to other activity. Triaxial accelerometer records the acceleration in X, Y, and Z-axis. In order to better understand row data of the accelerometer, we plotted the data from some activities in a certain time window in Figure 2. This figure shows the raw data for one participant in 20 seconds in four different types of activities: (1) supine position, (2) Tilting on the right and left (TRL), (3) slow walking, and (4) running.

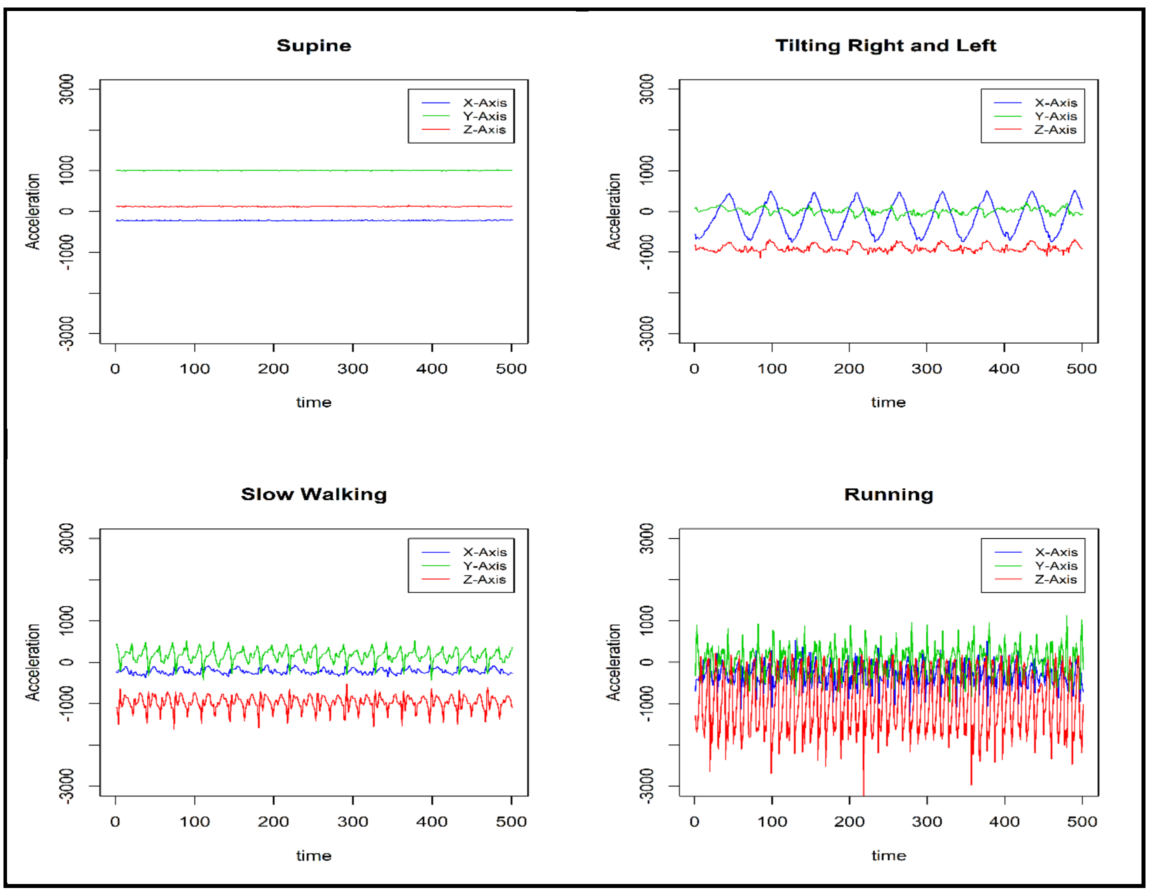


Figure 2. Raw Data Acceleration in Four Different Activities in 20 Seconds

From Figure 2, we can see, in the supine position, we do not have many changes in each axis compared to other three activities. In tilting to the right and left the side of the body we have more changes in X-axis compared to other axes. In slow walking, we have more fluctuation in Y-axis and Z-axis, and in the running, we have more changes in each axis, but the amount of changes in Z-axis is more than other axes.

After data labeling for each participant, we separated the data based on the type of activities. Since the data from the accelerometer is time series data, we can not change the order of data. Hence, we added each activity data from all subjects together without modifying the order of data. Figure 3 shows two sample of new datasets for two different activities from five people in X-axis. Since the accelerometer measures the static and dynamic acceleration, individual characteristics such as height affect the row data.

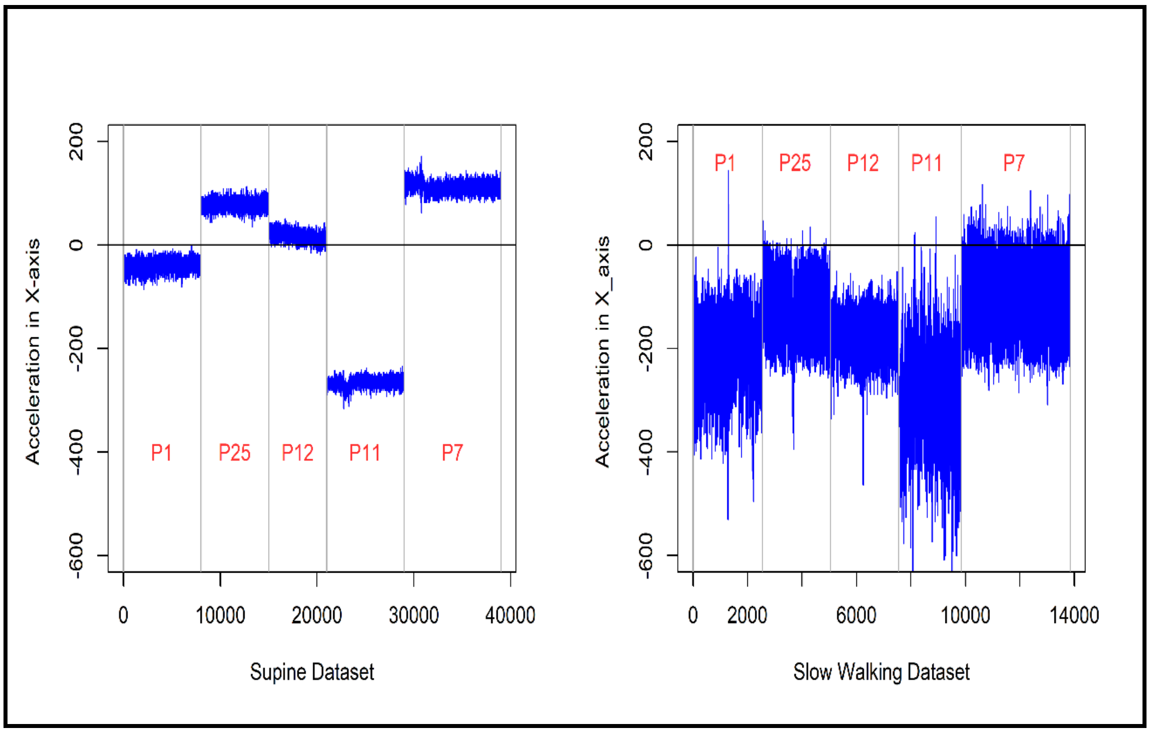


Figure 3. Visualization of Supine and Slow Walking Dataset for 5 People in X-Axis

Figure 3 shows how participants characteristics affect the row data. Although the pattern of data in a certain activity is similar for all participants, we can see the recorded acceleration for each person placed in different positions in the plots. This problem is not related to the one axis, and we have the same issue in other two axes. Since we will use machine learning algorithms, this issue is one of the problems which causes the model to depend on the participant's characteristics. In the next sections of this chapter, we will explain how we can fix this issue.

Table 4 has summarized the number of data points and equivalent time which we have at each dataset of activities for all participants. The length of all datasets are not equal. We collected more data in “dynamic activities” compared to “static activities” because we had more changes in activities.

Table 4. Number of Data Points and Equivalent Time in Each Activity Dataset

No.	Name of Dynamic and Static Activities	Number of Datapoints	Total Time (minute)
1	Lying down in the supine position	153950	102.63
2	Lying down on the left side of the body (LD-L)	30070	20.05
3	Lying down on the right side of the body (LD-R)	34270	22.85
4	Prone position	38120	25.41
5	Tilting on the right side of the body on standing position (T-R)	21020	14.01
6	Tilting on the left side of the body on standing position (T-L)	23070	15.38
7	Bending forward	28770	19.18
8	Standing	42453	28.30
9	Twisting on the right and left	31270	20.85
10	Tilting on the right and left (TRL)	28218	18.81
11	Bending forward and backward (Bending-UD)	33271	22.18
12	Squatting	31020	20.68
13	Slow walking	56370	37.58
14	Fast walking	59870	39.91
15	Running	29471	19.65

III.2. Feature Extraction

The triaxial accelerometer of Equivital generates the time series data in three dimensions, vertical movement axis (Z), lateral movement axis (X), and longitudinal movement axis (Y). The Equivital records the accelerations in each axis, every 40 ms. Building a model based on this data to recognize the activities on 40 ms, is impossible. When we talk about an activity, we consider some repetitive movements during the time. For instance, one full squat consists of sitting and standing up to the previous position. The time to complete a single squat takes more than at least one or two seconds. Hence, we should consider a time window of data and label it as a certain activity. On the other hand, we need to extract some features from raw data in that time window to provide more information for classifiers. We should select some attributes which participant's characteristics do not have a high effect on. In the next two sections, we introduce two features which we used in this study.

III.2.1. Angles of accelerometer vector

Accelerometer generates the acceleration vector in the time domain ($V_t=(x_t, y_t, z_t)$). The angle of acceleration vector with each axis is computable for each sample with equation below:

$$\theta x = Arccos\left(\frac{x}{R}\right), \theta y = Arccos\left(\frac{y}{R}\right), \theta z = Arccos\left(\frac{z}{R}\right)$$

Where R is the length of a vector which is represented by a three-component matrix at time t.

$$R = |(x, y, z)^T| = \sqrt{(x^2 + y^2 + z^2)}$$

The Equivital sensor provides the acceleration vector per 40 ms. Figure 4 illustrates the concept of angles and acceleration vector which are discussed. We need to consider some features of these angles for the certain time window as classifier predictors.

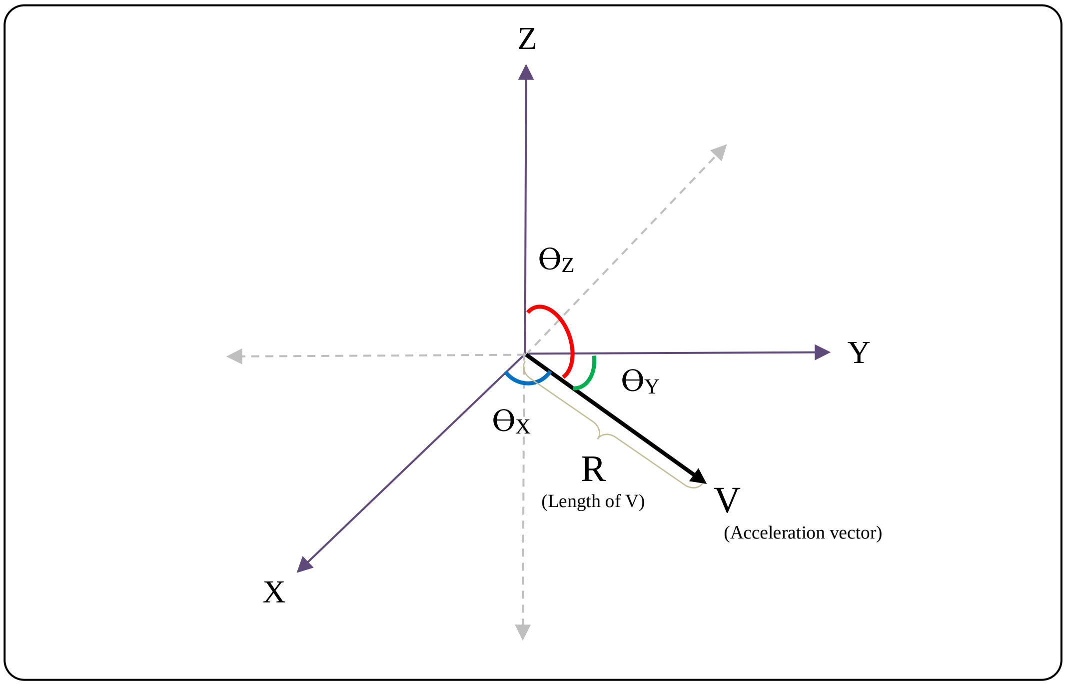


Figure 4. Illustration of Angles and Acceleration Vector

As features, we selected just median of angles in each axis. We selected median instead of mean because the average is sensitive to outlier data. In accelerometer data, very intensive activity in tiny time window may happen which affects the average of acceleration in that time, but the median is not sensitive to an outlier. This feature helps us to remove the effect of some unintentional acceleration in data.

III.2.2. Area under the curve (AUC)

Each time series is comprised of points measured at that point of time. AUC as depicted in Figure 5, is the area under the time series plot. The AUC of accelerometer times series estimates the velocity in each direction. This feature can capture the changes in acceleration; high acceleration can be returned large AUC in each axis. Since the acceleration data from the sensor can take a negative number, the area under the curve for certain periods of time ; (t, t+h) should be calculated by formulation below:

$$AUC_T = AUC_+ + AUC_- \quad (AUC_- \geq 0)$$

Where the AUC_T is the total area under the curve at a given period (h), AUC_+, and AUC_- are the area under the curve in the positive and negative part respectively.

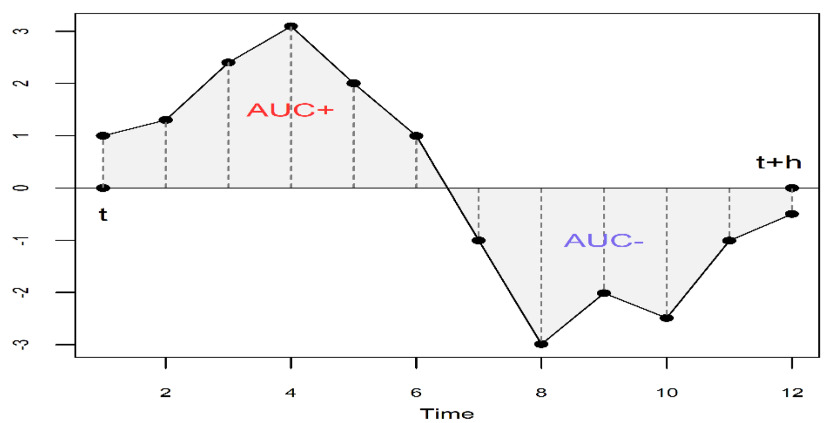


Figure 5. Illustration of Area Under the Curve

The AUC+ and AUC- can be derived from the trapezoid formula considering the zero line as a baseline. The point to use this feature is all data for all participants should place in the same range otherwise the result of this function is not reliable.

III.3. Data Transformation

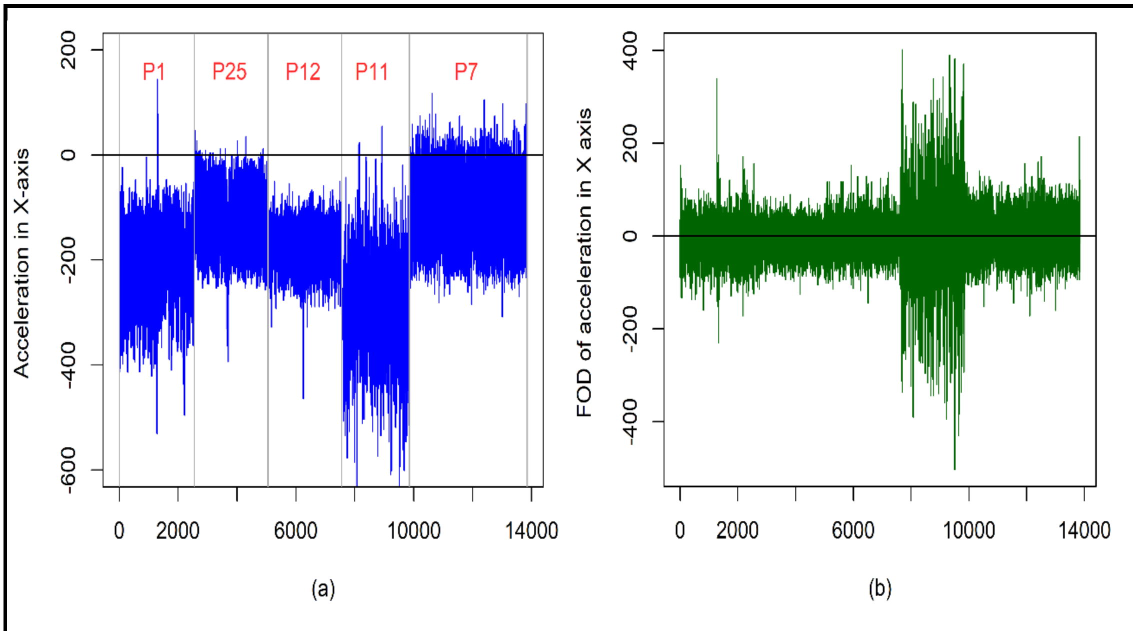
As we mentioned in the previous sections, the participant's characteristics affect the acceleration data. We used a transformation method which can reduce the effect of participant's characteristics and put all acceleration data in the same range for all subjects in a certain activity.

III.3.1. First order differencing (FOD)

Focusing on changes in the acceleration between time t and t-1 can be effective to remove the trace of individual characteristics. Differencing is a classic way to detrendand convert a series to stationary in time series analysis[30]. For each axis, FOD at time t is calculated by formulations below:

$$\nabla \mathrm{x}_t = \mathrm{x}_t - \mathrm{x}_{t-1},\ \nabla \mathrm{y}_t = \mathrm{y}_t - \mathrm{y}_{t-1},\ \nabla \mathrm{z}_t = \mathrm{z}_t - \mathrm{z}_{t-1}$$

Figure 6 illustrates how FOD affects the raw data to eliminate the effect of individual characteristics. FOD causes all acceleration data to oscillate around zero in each axis. This transformation method gives us the opportunity to use the AUC as a predictor in classifiers. Plots in Figure 6 have been drawn based on five different people data in slow walking activity. The plot (a) in this figure shows the raw data of acceleration in X-axis and plot (b) represent the FOD of that data.



(a) Raw data from five different people in slow walking activity in X-axis accelerometer data

(b) FOD transformed data from four different people in slow walking activity in X-axis accelerometer data

Figure 6. Illustration of First Order Differencing on Row Data

From the figure above we can see how the FOD causes all acceleration data from five different subjects for slow walking in the X-axis to fluctuate approximately in the same range. The FOD performs data stationarity regardless of the type of axis. In plot (b) we can see that pattern of some part of data is a little different from other regions and this is related to participant style in walking.

Overall, FOD transformation (1) removes some participant's characteristics effects and (2) causes all data fluctuate around zero which is necessary to use AUC for all activities.

III.4. Model Building and Validation

As we mentioned in previous sections, we need to extract some features of raw accelerometer data in the certain time window. In this section, we explain the procedures we performed to extract those features from each activity dataset.

The selected features for this study were (1) median of angles in each axis, (2) AUC in each axis, in the certain time window.

During the data collection, we measured the average of duration for all activities which performed by all participants. We found that at least 2 seconds is needed to complete one set of some activities such as squat and twisting. In literature review, most studies used the time window between 1.28 to 6.4 seconds. According to all of this information, we selected different time windows to extract all features for this study. We considered the time windows with the length of 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, and 6 seconds.

Since the features for time (t, t+h) depend on starting time (t), we should consider different situations to build a robust model which can recognize the activities without taking into account the starting time. In other words, our model should be able to recognize the activities for all different bunch of data related to the specific activity. The information above has been depicted in Figure 7.

Calculation of all scenarios for each data set is very time-consuming. For instance, the minimum number of data which we have in the activities datasets is 21,020 points, that is related to tilting on the right side of the body on standing position. If we want to consider all possible scenarios just in six seconds time interval (150 data points) and 80% of that as a training dataset, we need to calculate at least 16,666 (21020*.8-150) sample points.

Performing this size sampling and then conducting process related to random-forest takes more than a day in R. Hence, we selected six different sample sizes to find the effect of this feature on the results.We started with 1000 samples and each time added 500 additional samples to that and continued until we reached 5000.

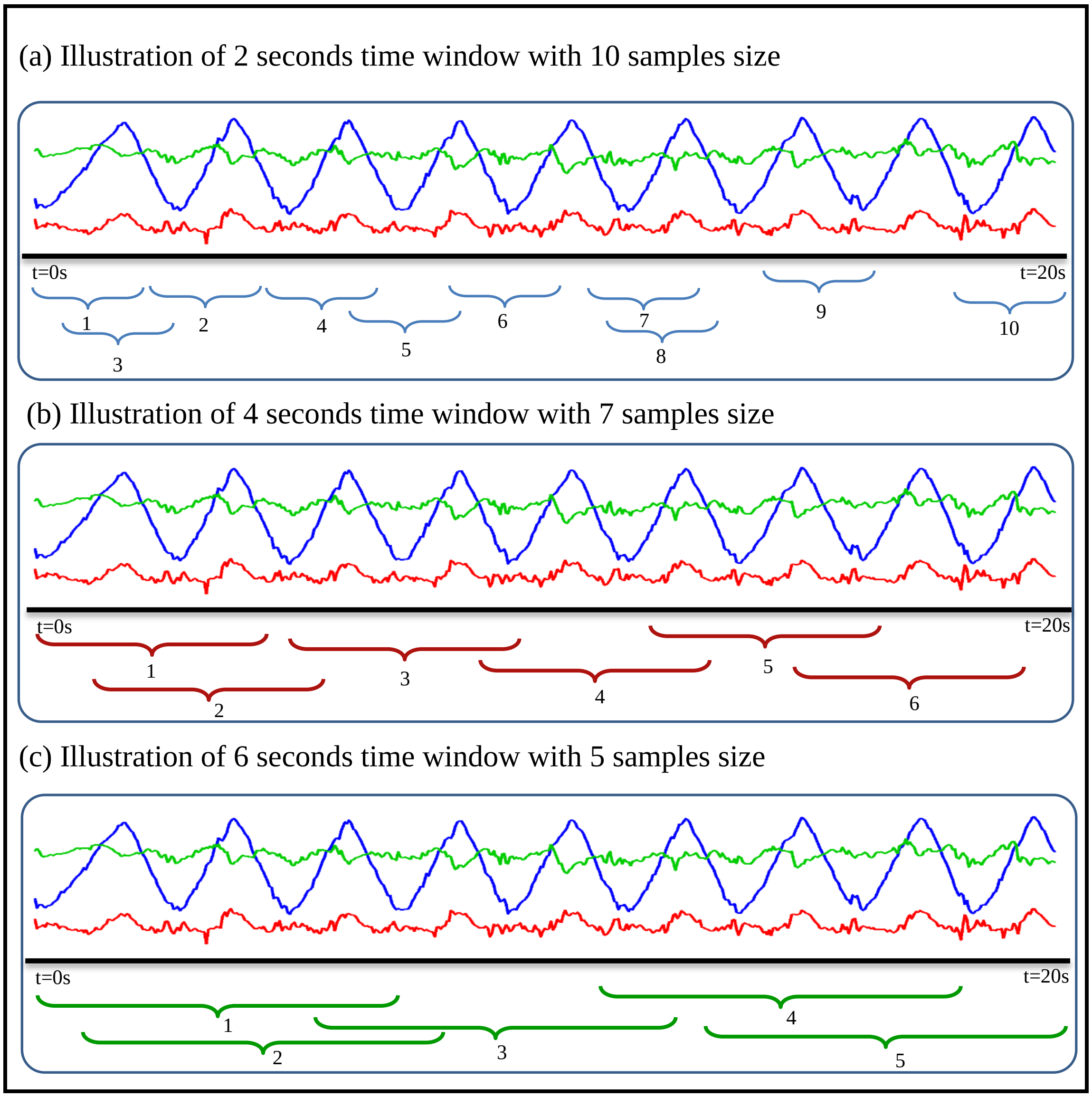


Figure 7. Illustration of Different Time Window and Sample Size

All scenarios based on different time-windows and sample sizes have been summarized in Table 5. In this step, we built different datasets and performed the random-forest algorithm on various time windows.

Table 5. Design Matrix for Random-Forest Models

Sample size		Time window (seconds)								
		2	2.5	3	3.5	4	4.5	5	5.5	6
	1000	Model 1-1	Model 1-2	Model 1-3	Model 1-4	Model 1-5	Model 1-6	Model 1-7	Model 1-8	Model 1-9
	1500	Model 2-1	Model 2-2	Model 2-3	Model 2-4	Model 2-5	Model 2-6	Model 2-7	Model 2-8	Model 2-9
	2000	Model 3-1	Model 3-2	Model 3-3	Model 3-4	Model 3-5	Model 3-6	Model 3-7	Model 3-8	Model 3-9
	2500	Model 4-1	Model 4-2	Model 4-3	Model 4-4	Model 4-5	Model 4-6	Model 4-7	Model 4-8	Model 4-9
	3000	Model 5-1	Model 5-2	Model 5-3	Model 5-4	Model 5-5	Model 5-6	Model 5-7	Model 5-8	Model 5-9
	3500	Model 6-1	Model 6-2	Model 6-3	Model 6-4	Model 6-5	Model 6-6	Model 6-7	Model 6-8	Model 6-9
	4000	Model 7-1	Model 7-2	Model 7-3	Model 7-4	Model 7-5	Model 7-6	Model 7-7	Model 7-8	Model 7-9
	4500	Model 8-1	Model 8-2	Model 8-3	Model 8-4	Model 8-5	Model 8-6	Model 8-7	Model 8-8	Model 8-9
	5000	Model 9-1	Model 9-2	Model 9-3	Model 9-4	Model 9-5	Model 9-6	Model 9-7	Model 9-8	Model 9-9

In total, we performed 81 different models in all datasets. In modeling, we divided the datasets into training and testing parts. We considered 80% of each dataset as training data set and 20% as a testing dataset.

The efficiency of each model was evaluated by confusion matrix which visualizes the performance of algorithms. Specificity, sensitivity, precision (positive predictive value), and the negative predictive value is calculated for each static and dynamic activity class to compare the efficiency of these two algorithms. Table 6 shows an example confusion matrix for two categories.

Table 6. Confusion Matrix for Two Classes

		Actual Classes	
		1	0
Predicted Classes	1	True Positive (TP)	False Positive (FP)
	0	False Negative (FN)	True Negative (TN)

The accuracy, sensitivity and precision (positive predictive value) are calculated by the equations below[31].

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

$$Sensitivity = Se = \frac{TP}{TP + FN} \quad , \quad \text{Precision} = \text{PPV} = \frac{TP}{TP + FP}$$

Type specific-performance recognition will be measured by the F-score which is computed by formulation below[13]:

$$F - score = 2 * (\frac{\text{Se} * \text{PPV}}{Se + PPV})$$

Where the Se and PPV are the sensitivity and positive predictive value (precision) in each activity type, respectively.

After performing all models, we selected the best time window and sample size which gave us the highest accuracy of the model. Also, we performed the support vector machine (SVM) algorithms on the selected time window and sample size. We compared the result of random-forest and SVM, and according to the results, we selected the best model as our final model.

In the next step, we combined some of the static activities based on the definition of sedentary behavior and physical activity; (1) lying down in the supine position, (2) lying down on the left side of the body, (3) lying down on the right side of the body, (4) prone position, and (5) bending were considered as the Lying-Down category. We included the bending in this category because based on sensor position on the chest, bending is very similar to the prone position.

Also, we combined the standing, tilting on the right side of the body on standing position, and tilting on the left side of the body on standing position to the new category which was named standing-star.

As a final step, we combined the Lying-Down category and standing-star category as sedentary behavior category and evaluate the performance of models in new categories.

CHAPTER IV

MODELING RESULTS

This chapter provides the results of models discussed in the previous chapter.

The results for each model are summarized in two tables; (1) the confusion matrix between the predicted value and actual value of the test data and (2) the table of sensitivity, specificity, precision (positive predictive value), negative predictive value, F-score in each category, the overall accuracy of model and confidence interval for that. In section 2 of this chapter, we have compared the results of random-forest models based on sample size and time windows to select the best model. In the last two sections, we compared the results of random-forest and SVM for certain sample size and time windows, and also we compared the results of these two classifiers by combining some classes.

IV.1. Descriptive Analysis of Random-Forest Models

In this section, we selected 9 models out of 81 performed models. The results of all models based on overall accuracy are depicted at next section. We selected the models at which the sample size was 1000, 3000, and 5000 points in each static and dynamic activity. As time window, we selected models with time window 2, 4, and 6 seconds. Note that each confution matrix will have a total of $0.2 \times$ sample size $\times$ Number of activities data points.

IV.1.1. Model -1-1

For this model, The sample size was 1000 points in each static and dynamic activity dataset. Also, 2 seconds was considered as time window to extract the features. Table 7 shows the confusion matrix and statistics of this model on the testing data.

Table 7. Confusion Matrix and Statistics of Model 1-1

Model 1-1	sample size =1000 time window= 2 s	Actual Value														
		Static Activities								Dynamic Activities						
		Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Predicted Value	Supine	159	0	0	0	0	0	0	5	0	0	2	0	0	0	0
	LD-R	2	211	0	0	0	0	0	0	0	0	0	0	0	0	0
	LD-L	2	0	194	0	0	0	0	0	0	0	0	0	0	0	0
	Prone	3	0	0	188	0	0	20	0	0	0	0	0	0	0	0
	T-R	1	0	0	0	202	0	0	0	0	0	0	0	0	0	0
	T-L	1	0	0	0	0	198	0	2	0	0	0	0	0	0	0
	Bending	1	0	0	10	0	0	185	0	0	1	0	0	0	0	0
	Standing	11	0	0	0	0	0	0	185	0	3	2	0	1	0	6
	TRL	1	0	0	0	0	0	0	0	186	2	1	0	0	0	0
	Bending-UD	3	0	0	0	0	0	0	0	1	176	22	0	0	0	0
	Squatting	4	0	0	0	0	0	0	0	1	28	156	3	0	0	0
	Slow walking	1	0	0	0	0	0	0	0	0	0	5	197	17	0	0
	Fast walking	1	0	0	0	0	0	0	0	0	0	2	13	168	2	0
	Running	1	0	0	0	0	0	0	0	0	0	0	0	2	191	0
	Twisting	11	0	0	0	0	0	0	11	0	2	3	1	0	0	193

Model 1-1	Overall Accuracy: 0.929							95% CI : (0.919 , 0.938)							
	Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Sensitivity	0.79	1.00	1.00	0.95	1.00	1.00	0.90	0.91	0.99	0.83	0.81	0.92	0.89	0.99	0.97
Specificity	1.00	1.00	1.00	0.99	1.00	1.00	1.00	0.99	1.00	0.99	0.99	0.99	0.99	1.00	0.99
Percision	0.96	0.99	0.99	0.89	1.00	0.99	0.94	0.89	0.98	0.87	0.81	0.90	0.90	0.98	0.87
NPV	0.98	1.00	1.00	1.00	1.00	1.00	0.99	0.99	1.00	0.99	0.99	0.99	0.99	1.00	1.00
F-score	0.86	1.00	0.99	0.92	1.00	0.99	0.92	0.90	0.98	0.85	0.81	0.91	0.90	0.99	0.92

From Table 7 we can see the overall accuracy of the model is 92.9%. The model has reliable performance in most categories.The best results are related to LD-R and T-R. The lowest performance is related to squatting with F-score equal 0.81.

IV.1.2. Model -1-5

For this model, The sample size was 1000 points in each static and dynamic activity. Also, 4 seconds was considered as time window to extract the features. Table 8 shows the confusion matrix and statistics of this model on the testing data.

Table 8. Confusion Matrix and Statistics of Model 1-5

Model 1-5 *sample size =1000* *time window= 4 s*		Actual Value														
		Static Activities								Dynamic Activities						
		Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Predicted Value	Supine	159	0	0	0	0	0	0	1	0	0	2	0	0	0	0
	LD-R	1	211	0	0	0	0	0	0	0	0	0	0	0	0	0
	LD-L	4	0	194	0	0	0	0	0	0	0	0	0	0	0	0
	Prone	1	0	0	189	0	0	6	0	0	0	0	0	0	0	0
	T-R	1	0	0	0	202	0	0	0	0	0	0	0	0	0	0
	T-L	0	0	0	0	0	198	0	2	0	0	0	0	0	0	0
	Bending	2	0	0	9	0	0	199	0	0	0	0	0	0	0	0
	Standing	9	0	0	0	0	0	0	195	0	2	0	0	1	0	2
	TRL	0	0	0	0	0	0	0	0	188	0	0	0	0	0	0
	Bending-UD	3	0	0	0	0	0	0	0	0	192	26	0	0	0	0
	Squatting	2	0	0	0	0	0	0	0	0	18	163	2	1	0	1
	Slow walking	6	0	0	0	0	0	0	0	0	0	1	208	9	0	0
	Fast walking	4	0	0	0	0	0	0	0	0	0	0	4	177	1	0
	Running	1	0	0	0	0	0	0	0	0	0	0	0	0	192	0
	Twisting	9	0	0	0	0	0	0	5	0	0	1	0	0	0	196

Model 1-5	Overall Accuracy: 0.954							95% CI : (0.946 , 0.961)							
	Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Sensitivity	0.79	1.00	1.00	0.95	1.00	1.00	0.97	0.96	1.00	0.91	0.84	0.97	0.94	0.99	0.98
Specificity	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.99	0.99	0.99	1.00	1.00	0.99
Percision	0.98	1.00	0.98	0.96	1.00	0.99	0.95	0.93	1.00	0.87	0.87	0.93	0.95	0.99	0.93
NPV	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.00
F-score	0.87	1.00	0.99	0.96	1.00	0.99	0.96	0.95	1.00	0.89	0.86	0.95	0.95	0.99	0.96

From Table 8 we can see the overall accuracy of the model is 95.4%. The model has the reliable performance in most categories. The best results are related to LD-R, T-R and Tilting on the right and left (TRL).

Comparing the results of F-scores between model 1-1 and model 1-5 shows that increasing the length of time window has a positive effect on the F-score and accuracy of the model. In other words, we need more data points to extract features. Features which were extracted from longer time window contain more information about static and dynamic activities.

IV.1.3. Model -1-9

For this model, the sample size was 1000 points in each static and dynamic activity. Also, 6 seconds was considered as time window to extract the features. Table 9 shows the confusion matrix and statistics of this model on the testing data.

Table 9. Confusion Matrix and Statistics of Model 1-9

Model 1-9

sample size =1000
time window= 6 s

	Actual Value														
	Static Activities								Dynamic Activities						
Predicted Value	Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Supine	174	0	0	0	0	0	0	0	0	0	0	0	1	0	0
LD-R	2	211	0	0	0	0	0	0	0	0	0	0	0	0	0
LD-L	1	0	194	0	0	0	0	0	0	0	0	0	0	0	0
Prone	1	0	0	193	0	0	9	0	0	0	0	0	0	0	0
T-R	0	0	0	0	202	0	0	0	0	0	0	0	0	0	0
T-L	1	0	0	0	0	197	0	5	0	0	0	0	0	0	0
Bending	0	0	0	5	0	0	196	0	0	0	0	0	0	0	0
Standing	9	0	0	0	0	1	0	197	0	0	0	0	1	0	4
TRL	0	0	0	0	0	0	0	0	188	0	0	0	0	0	0
Bending-UD	4	0	0	0	0	0	0	0	0	200	16	0	0	0	0
Squatting	1	0	0	0	0	0	0	0	0	10	176	0	0	0	0
Slow walking	2	0	0	0	0	0	0	0	0	0	1	212	3	0	0
Fast walking	2	0	0	0	0	0	0	0	0	0	0	2	183	0	0
Running	1	0	0	0	0	0	0	0	0	0	0	0	0	193	0
Twisting	4	0	0	0	0	0	0	1	0	2	0	0	0	0	195

	Overall Accuracy: 0.970							95% CI : (0.963 , 0.976)							
Model 1-9	Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Sensitivity	0.86	1.00	1.00	0.97	1.00	0.99	0.96	0.97	1.00	0.94	0.91	0.99	0.97	1.00	0.98
Specificity	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.99	1.00	1.00	1.00	1.00	1.00
Percision	0.99	0.99	0.99	0.95	1.00	0.97	0.98	0.93	1.00	0.91	0.94	0.97	0.98	0.99	0.97
NPV	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00
F-score	0.92	1.00	1.00	0.96	1.00	0.98	0.97	0.95	1.00	0.93	0.93	0.98	0.98	1.00	0.97

From the table above we can see the overall accuracy of the model is 97.4%. The model has a reliable performance in most categories. The best results are related to LD-R, LD-L, T-R, Running and Tilting on the right and left (TRL). By increasing the time window, all F-scores have been improved.

IV.1.4. Model -5-1

For this model, The sample size equals 3000 points in each static and dynamic activity. Also, 2 seconds was considered as time window to extract the features. Table 10 shows the confusion matrix and statistics of this model on the testing data.

Table 10. Confusion Matrix and Statistics of Model 5-1

Model 5-1 *sample size =3000* *time window= 2 s*	Actual Value														
	Static Activities								Dynamic Activities						
Predicted Value	Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Supine	478	0	0	0	0	0	0	4	0	0	0	0	1	0	0
LD-R	5	576	0	0	0	0	0	0	0	0	0	0	0	0	0
LD-L	6	0	607	0	0	0	0	0	0	0	0	0	0	0	0
Prone	8	0	0	614	0	0	21	0	0	1	0	0	0	0	0
T-R	0	0	0	0	624	0	0	0	0	0	0	0	0	0	0
T-L	2	0	0	0	0	596	0	8	0	0	0	0	0	0	0
Bending	3	0	0	10	0	0	605	0	0	0	0	0	0	0	0
Standing	34	0	0	0	0	1	0	590	0	3	2	0	4	0	24
TRL	7	0	0	0	0	0	0	0	565	0	1	0	0	0	1
Bending-UD	7	0	0	0	0	0	0	0	1	548	57	1	0	0	0
Squatting	7	0	0	0	0	0	0	1	2	50	500	5	2	1	2
Slow walking	17	0	0	0	0	0	0	0	0	0	8	603	29	1	0
Fast walking	7	0	0	0	0	0	0	0	0	0	2	16	525	4	0
Running	4	0	0	0	0	0	0	0	0	0	0	0	3	597	0
Twisting	11	0	0	0	0	0	0	15	3	4	7	0	0	0	559

	Overall Accuracy: 0.954							95% CI : (0.949 , 0.958)							
Model 5-1	Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Sensitivity	0.80	1.00	1.00	0.98	1.00	1.00	0.97	0.95	0.99	0.90	0.87	0.96	0.93	0.99	0.95
Specificity	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.99	0.99	0.99	1.00	1.00	1.00
Percision	0.99	0.99	0.99	0.95	1.00	0.98	0.98	0.90	0.98	0.89	0.88	0.92	0.95	0.99	0.93
NPV	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.00
F-score	0.89	1.00	1.00	0.97	1.00	0.99	0.97	0.92	0.99	0.90	0.87	0.94	0.94	0.99	0.94

From Table 10 we can see the overall accuracy of the model is 95.4%. The model has the reliable performance in most categories. The best results are related to LD-R, LD-L, T-R.

Comparing the results of this model with model 1-1 shows by increasing the sample size without a change in the time window, the performance of model has been improved. Although this result is better than model 1-1, the performance of the model in squatting is deficient. This result shows there is a positive correlation between sample size and model accuracy and F-scores. The results of model 5-5 and 5-9 in the next steps provide more evidence for this hypothesis.

IV.1.5. Model -5-5

For this model, The sample size was 3000 points in each static and dynamic activity. Also, 4 seconds was considered as time window to extract the features. Table 11 shows the confusion matrix and statistics of this model on the testing data.

Table 11. Confusion Matrix and Statistics of Model 5-5

Model 5-5 *sample size =3000* *time window= 4 s*		Actual Value														
		Static Activities								Dynamic Activities						
		Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Predicted Value	Supine	469	0	0	0	0	0	2	3	0	1	0	3	0	0	0
	LD-R	6	576	0	0	0	0	0	0	0	0	0	0	0	0	0
	LD-L	4	0	607	0	0	0	0	0	0	0	0	0	0	0	0
	Prone	3	0	0	621	0	0	10	0	0	0	0	0	0	0	0
	T-R	5	0	0	0	624	0	0	0	0	0	0	0	0	0	0
	T-L	2	0	0	0	0	594	0	10	0	0	0	0	0	0	0
	Bending	6	0	0	3	0	0	614	0	0	0	0	0	0	0	0
	Standing	37	0	0	0	0	3	0	599	0	3	0	0	0	0	6
	TRL	2	0	0	0	0	0	0	0	571	0	0	0	0	0	0
	Bending-UD	3	0	0	0	0	0	0	0	0	564	43	0	0	0	0
	Squatting	14	0	0	0	0	0	0	0	0	37	532	5	2	0	0
	Slow walking	15	0	0	0	0	0	0	0	0	0	1	611	14	0	0
	Fast walking	13	0	0	0	0	0	0	0	0	0	0	6	547	1	0
	Running	4	0	0	0	0	0	0	0	0	0	0	0	0	602	0
	Twisting	13	0	0	0	0	0	0	6	0	1	1	0	1	0	580

	Overall Accuracy: 0.967							95% CI : (0.964 , 0.971)							
Model 5-5	Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Sensitivity	0.79	1.00	1.00	1.00	1.00	0.99	0.98	0.97	1.00	0.93	0.92	0.98	0.97	1.00	0.99
Specificity	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.99	0.99	1.00	1.00	1.00	1.00
Percision	0.98	0.99	0.99	0.98	0.99	0.98	0.99	0.92	1.00	0.92	0.90	0.95	0.96	0.99	0.96
NPV	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.00
F-score	0.87	0.99	1.00	0.99	1.00	0.99	0.98	0.95	1.00	0.93	0.91	0.97	0.97	1.00	0.98

From the table above we can see the overall accuracy of the model is 96.7%. The model has a reliable performance in most categories. The best results are related to LD-L, T-R, TRL, and Running.

IV.1.6. Model -5-9

For this model, The sample size was 3000 points in each static and dynamic activity. Also, 6 seconds was considered as time window to extract the features. Table 12 shows the confusion matrix and statistics of this model on the testing data.

Table 12. Confusion Matrix and Statistics of Model 5-9

Model 5-9 *sample size =3000 time window= 6 s*		Actual Value														
		Static Activities								Dynamic Activities						
		Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Predicted Value	Supine	479	0	0	0	0	0	0	1	0	0	0	0	2	0	0
	LD-R	1	576	0	0	0	0	0	0	0	0	0	0	0	0	0
	LD-L	1	0	607	0	0	0	0	0	0	0	0	0	0	0	0
	Prone	3	0	0	619	0	0	4	0	0	0	0	0	0	0	0
	T-R	8	0	0	0	624	0	0	0	0	0	0	0	0	0	0
	T-L	1	0	0	0	0	595	0	6	0	0	0	0	0	0	0
	Bending	5	0	0	5	0	0	622	0	0	0	0	0	0	0	0
	Standing	30	0	0	0	0	2	0	604	0	0	0	0	0	0	3
	TRL	5	0	0	0	0	0	0	0	570	0	0	0	0	0	1
	Bending-UD	6	0	0	0	0	0	0	0	0	592	25	0	0	0	0
	Squatting	8	0	0	0	0	0	0	0	0	14	548	1	1	0	0
	Slow walking	19	0	0	0	0	0	0	0	0	0	4	618	4	0	0
	Fast walking	8	0	0	0	0	0	0	0	0	0	0	6	556	1	0
	Running	4	0	0	0	0	0	0	0	0	0	0	0	1	602	0
	Twisting	18	0	0	0	0	0	0	7	1	0	0	0	0	0	582

Model 5-9	Overall Accuracy: 0.977							95% CI : (0.973 , 0.980)							
	Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Sensitivity	0.80	1.00	1.00	0.99	1.00	1.00	0.99	0.98	1.00	0.98	0.95	0.99	0.99	1.00	0.99
Specificity	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Percision	0.99	1.00	1.00	0.99	0.99	0.99	0.98	0.95	0.99	0.95	0.96	0.96	0.97	0.99	0.96
NPV	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
F-score	0.89	1.00	1.00	0.99	0.99	0.99	0.99	0.96	0.99	0.96	0.95	0.97	0.98	1.00	0.97

From the table above we can see the overall accuracy of the model is 97.7%. The model has a reliable performance in most categories. The best results are related to LD-L, LD-R, T-R, and Running.

Based on results of model 5-5 and model 5-9, we can see there is more evidence that shows the model accuracy increases by increasing the sample size in the same time window. All models with sample size equal to 3000 have better performance compared to the model with sample size equal to 1000 in the same time window.

In the next section, we bring the results of other three model with sample size equal to 5000 with 2,4, and 6 seconds time window. The 5000 sample size was the highest sample size which we test in this study.

IV.1.7. Model -9-1

For this model, The sample size was 5000 points in each static and dynamic activity. Also, 2 seconds was considered as time window to extract the features. Table 13 shows the confusion matrix and statistics of this model on the testing data.

Table 13. Confusion Matrix and Statistics of Model 9-1

Model 9-1 *sample size =5000* *time window= 2 s*		Actual Value														
		Static Activities								Dynamic Activities						
		Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Predicted Value	Supine	794	0	0	0	0	0	0	3	0	0	0	1	3	0	0
	LD-R	13	996	0	0	0	0	0	0	0	0	0	0	0	0	0
	LD-L	9	0	1016	0	0	0	0	0	0	0	0	0	0	0	0
	Prone	7	0	0	974	0	0	15	0	0	1	0	0	0	0	0
	T-R	3	0	0	0	977	0	0	0	0	0	0	0	0	0	0
	T-L	1	0	0	0	0	979	0	13	0	0	0	0	0	0	0
	Bending	5	0	0	8	0	0	991	0	0	0	0	0	0	0	0
	Standing	56	0	0	0	0	1	0	960	0	5	2	0	4	3	18
	TRL	5	0	0	0	0	0	0	2	1001	1	7	0	0	0	2
	Bending-UD	10	0	0	0	0	0	0	0	1	930	86	0	0	0	3
	Squatting	22	0	0	0	0	0	0	3	1	63	949	7	2	1	3
	Slow walking	29	0	0	0	0	0	0	0	0	0	7	965	44	3	0
	Fast walking	7	0	0	0	0	0	0	0	0	0	2	32	916	12	0
	Running	9	0	0	0	0	0	0	0	0	0	0	0	1	972	0
	Twisting	30	0	0	0	0	0	0	46	3	5	6	2	2	0	950

Model 9-1	Overall Accuracy: 0.958							95% CI : (0.954 , 0.961)							
	Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Sensitivity	0.79	1.00	1.00	0.99	1.00	1.00	0.99	0.93	1.00	0.93	0.90	0.96	0.94	0.98	0.97
Specificity	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.99	0.99	0.99	1.00	1.00	0.99
Percision	0.99	0.99	0.99	0.98	1.00	0.99	0.99	0.92	0.98	0.90	0.90	0.92	0.95	0.99	0.91
NPV	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.00
F-score	0.88	0.99	1.00	0.98	1.00	0.99	0.99	0.92	0.99	0.91	0.90	0.94	0.94	0.99	0.94

From the table above we can see the overall accuracy of the model is 95.8%. The model has a reliable performance in most categories. The best results are related to LD-L, T-R.

IV.1.8. Model -9-5

For this model, The sample size was equal 5000 points in each static and dynamic activity. Also, 4 seconds was considered as time window to extract the features. Table 14 shows the confusion matrix and statistics of this model on the testing data.

Table 14. Confusion Matrix and Statistics of Model 9-5

Model 9-5

sample size =5000
time window= 4 s

Predicted Value	Actual Value: Static Activities: Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	Dynamic Activities: TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Supine	842	0	0	1	0	0	1	3	0	0	0	5	2	1	1
LD-R	2	996	0	0	0	0	0	0	0	0	0	0	0	0	0
LD-L	5	0	1016	0	0	0	0	0	0	0	0	0	0	0	0
Prone	8	0	0	974	0	0	12	0	0	0	0	0	0	0	0
T-R	3	0	0	0	977	0	0	0	0	0	0	0	0	0	0
T-L	2	0	0	0	0	975	0	8	0	0	0	0	0	0	0
Bending	6	0	0	7	0	0	993	0	0	0	0	0	0	0	0
Standing	46	0	0	0	0	5	0	1010	0	1	0	0	1	0	10
TRL	5	0	0	0	0	0	0	0	1006	0	0	0	0	0	0
Bending-UD	10	0	0	0	0	0	0	0	0	960	63	0	0	0	0
Squatting	14	0	0	0	0	0	0	0	0	44	990	2	1	0	0
Slow walking	17	0	0	0	0	0	0	0	0	0	5	990	23	0	0
Fast walking	6	0	0	0	0	0	0	0	0	0	0	7	944	1	0
Running	7	0	0	0	0	0	0	0	0	0	0	0	1	989	0
Twisting	27	0	0	0	0	0	0	6	0	0	1	3	0	0	965

Overall Accuracy: 0.975 95% CI : (0.972 , 0.977)

Model 9-5	Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Sensitivity	0.84	1.00	1.00	0.99	1.00	0.99	0.99	0.98	1.00	0.96	0.93	0.98	0.97	1.00	0.99
Specificity	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00
Percision	0.98	1.00	1.00	0.98	1.00	0.99	0.99	0.94	1.00	0.93	0.94	0.96	0.99	0.99	0.96
NPV	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
F-score	0.91	1.00	1.00	0.99	1.00	0.99	0.99	0.96	1.00	0.94	0.94	0.97	0.98	0.99	0.98

From Table 14 we can see the overall accuracy of the model is 97.5%. The model has a reliable performance in most categories. The best results are related to LD-R, LD-L, T-R and Tilting on the right and left (TRL).

F-scores shows the performance of the model in most categories is improved by increasing the sample size and time window. We have one static activity with F-score under 96%. Performance of model in all activities is more than 94%. By investigating in misclassifications, we can see that prone position has been misclassified as bending position and vice versa. We have the same situation in slow and fast walking categories. The most misclassifications in fast walking have occurred in the slow walking category and vice versa.

IV.1.9. Model -9-9

For this model, The sample size was equal 5000 points in each static and dynamic activity. Also, 6 seconds was considered as time window to extract the features. Table 15 shows the confusion matrix and statistics of this model on the testing data.

Table 15. Confusion Matrix and Statistics of Model 9-9

Model 9-9 *sample size =5000* *time window= 6 s*		Actual Value														
		Static Activities								Dynamic Activities						
		Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Predicted Value	Supine	863	3	0	1	0	0	0	3	1	0	0	0	1	0	0
	LD-R	4	993	0	0	0	0	0	0	0	0	0	0	0	0	0
	LD-L	8	0	1016	0	0	0	0	0	0	0	0	0	0	0	0
	Prone	5	0	0	977	0	0	5	0	0	0	0	0	0	0	0
	T-R	6	0	0	0	977	0	0	0	0	0	0	0	0	0	0
	T-L	1	0	0	0	0	978	0	14	0	0	0	0	0	0	0
	Bending	5	0	0	4	0	0	1001	0	0	0	0	0	0	0	0
	Standing	36	0	0	0	0	2	0	1007	0	1	0	0	2	0	10
	TRL	6	0	0	0	0	0	0	0	1002	0	0	0	0	0	0
	Bending-UD	6	0	0	0	0	0	0	0	0	982	40	0	0	0	0
	Squatting	9	0	0	0	0	0	0	0	2	22	1016	0	1	0	0
	Slow walking	18	0	0	0	0	0	0	0	0	0	2	1001	7	0	0
	Fast walking	10	0	0	0	0	0	0	0	0	0	1	3	961	1	0
	Running	5	0	0	0	0	0	0	0	0	0	0	0	0	990	0
	Twisting	18	0	0	0	0	0	0	3	1	0	0	3	0	0	966

Model 9-9	Overall Accuracy: 0.982							95% CI : (0.979 , 0.984)							
	Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Sensitivity	0.86	1.00	1.00	0.99	1.00	1.00	1.00	0.98	1.00	0.98	0.96	0.99	0.99	1.00	0.99
Specificity	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Percision	0.99	1.00	0.99	0.99	0.99	0.98	0.99	0.95	0.99	0.96	0.97	0.97	0.98	0.99	0.97
NPV	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
F-score	0.92	1.00	1.00	0.99	1.00	0.99	0.99	0.97	1.00	0.97	0.96	0.98	0.99	1.00	0.98

From the table above we can see the overall accuracy of the model is 98.2%. The model has a reliable performance in most categories. The best results are related to LD-R, LD-L, T-R, running and Tilting on the right and left (TRL).

IV.2. Summary of Random-Forest Models

In the previous section, we provided the details of some random-forest models and discussed the performance of models in each category of static and dynamic activities. In this section, we compare the accuracy of all 81 models and assess the effect of the time window and sample size on the accuracy of models. Table 16 shows the overall accuracy which was achieved on random-forest models.

Table 16. Overall accuracy of Random Forest Models

		Time window (seconds)								
		2	**2.5**	**3**	**3.5**	**4**	**4.5**	**5**	**5.5**	**6**
Sample size	**1000**	0.930	0.948	0.954	0.952	0.954	0.961	0.963	0.960	0.970
	1500	0.942	0.949	0.953	0.952	0.962	0.965	0.967	0.966	0.972
	2000	0.944	0.956	0.952	0.961	0.967	0.964	0.976	0.970	0.976
	2500	0.949	0.957	0.962	0.964	0.968	0.966	0.976	0.971	0.977
	3000	0.954	0.963	0.965	0.966	0.968	0.972	0.976	0.977	0.977
	3500	0.955	0.964	0.966	0.969	0.972	0.973	0.977	0.977	0.976
	4000	0.954	0.964	0.969	0.973	0.975	0.973	0.975	0.978	0.980
	4500	0.960	0.968	0.969	0.973	0.973	0.977	0.978	0.979	0.980
	5000	0.958	0.967	0.971	0.972	0.975	0.979	0.980	0.980	0.982

From the table above, we can see the accuracy of random forest models increases when both variables, time window and sample size have some increase in their value. The lowest accuracy is related to the model with 1000 sample size and 2 seconds time window. The best performance is related to model with 5000 sample size and 6 seconds time window. All information in Table 16 is depicted in Figure 7.

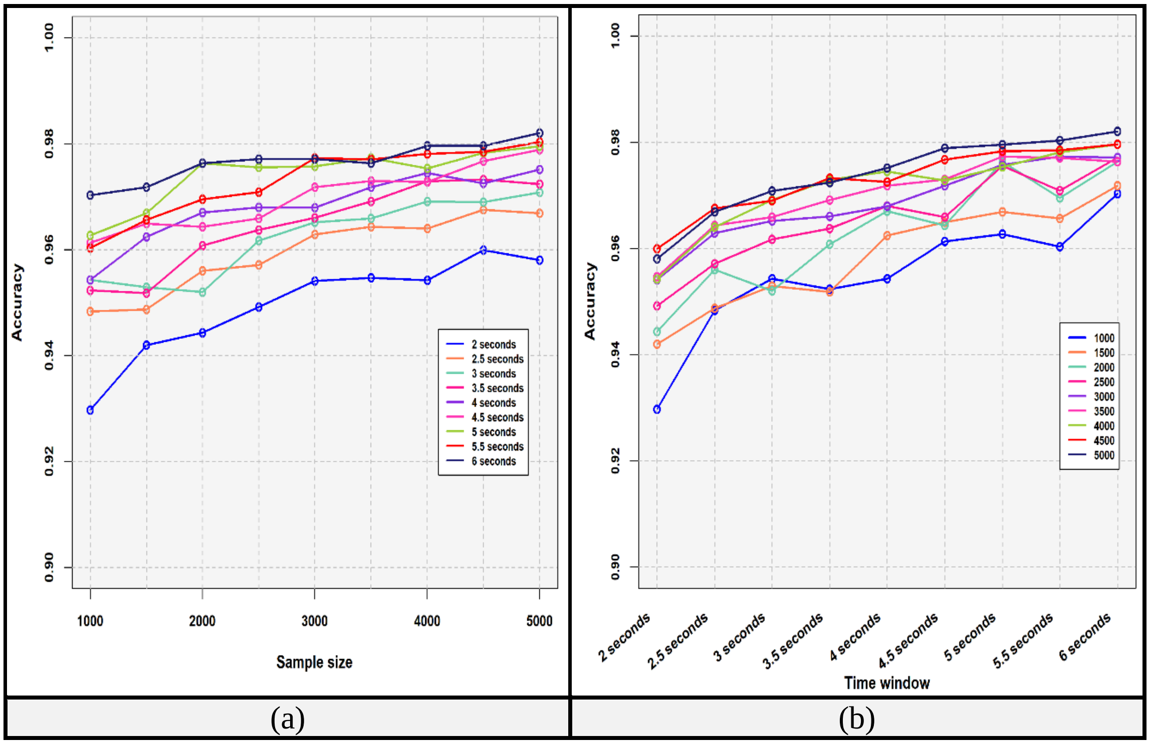


Figure 8. Accuracy of Models Based on Time Window and Sample Size

Figure 7- (a) shows, all models in 2 seconds time window have had the lowest accuracy compared to other time windows. In contrast, the highest results were related to 6 seconds time window in the same sample size.

Figure 7-(b) represents, the models with 1000 sample size have had the lowest accuracy compared to other sample sizes in the same time window. In opposite, the highest accuracy is related to 5000 sample size in the same time window.

The results from Table 7 show, the best random-forest model is related to sample size 5000 and 6 seconds time window with overall accuracy 98.2%. We can see, the accuracy of models in some levels of the time window and sample size are very close to each other which is presented in Figure 9.

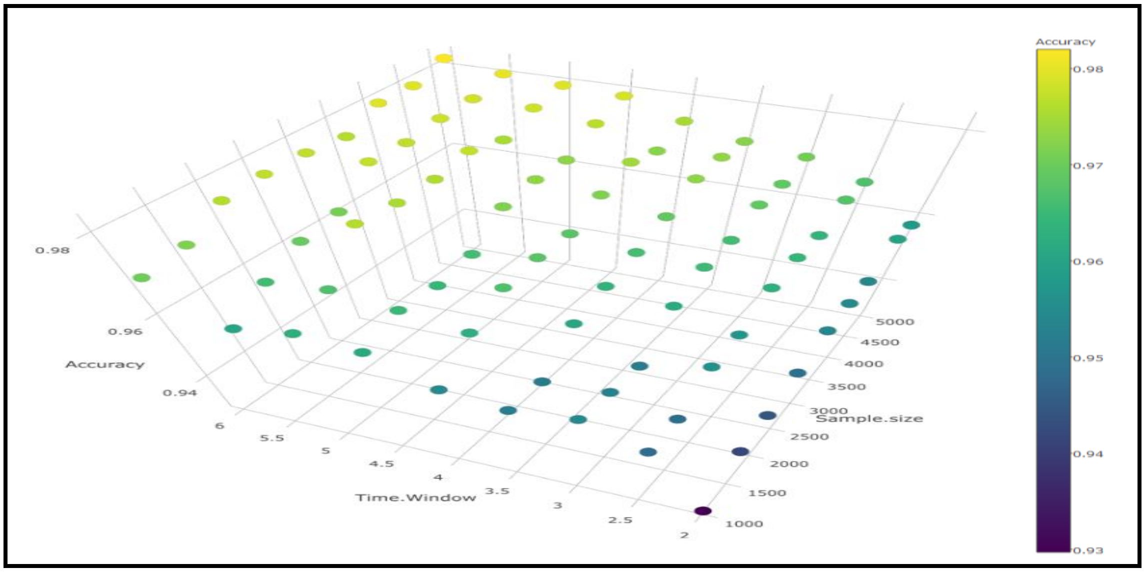


Figure 9. 3D Scatter Plot of Accuracy, Time Window, and Sample Size

These results show we can have several choices to select the model. If the overall accuracy of the model is the most important feature, we can choose the model with the highest accuracy. Otherwise, we can consider a certain level of accuracy and then based on other variables which are time window and sample size, select the model. For instance, if the recognizing the activities in the certain time window is important, we can define a threshold for accuracy and find the best sample size which provides the defined accuracy in that time window. The other scenario can happen when we do not have enough data. In this situation, we can fix the sample size to find the best time window to achieve the desired accuracy. Runnig time to build the model in 2 second time window and 1000 sample size is at list 3 minutes and running time to build the model based on 6 second time window and 5000 sample size is aproximatly 46 minutes. The predict time for both models takes less than 1 second.

IV.3. Results of SVM Model

In this section, we provided the results of SVM model only for 5000 sample size and 6 seconds time window and compared them with the random-forest model in the same sample size and time window. Table 17 shows the results of SVM model.

Table 17. Confusion Matrix and Statistics of SVM Model

SVM Model *sample size =5000* *time window= 6 s*		Actual Value														
		Static Activities								Dynamic Activities						
		Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Predicted Value	Supine	827	0	0	2	0	0	0	3	0	0	3	2	0	0	0
	LD-R	4	996	0	0	0	0	0	0	0	0	0	0	0	0	0
	LD-L	9	0	1016	0	0	0	0	0	0	0	0	0	0	0	0
	Prone	6	0	0	957	0	0	157	0	0	0	0	0	0	0	0
	T-R	6	0	0	0	977	0	0	0	0	0	0	0	0	0	0
	T-L	1	0	0	0	0	980	0	18	0	0	0	0	0	0	0
	Bending	4	0	0	23	0	0	849	0	0	0	0	0	0	0	0
	Standing	58	0	0	0	0	0	0	998	0	7	0	0	2	0	66
	TRL	6	0	0	0	0	0	0	0	1004	0	0	0	0	0	0
	Bending-UD	6	0	0	0	0	0	0	0	0	924	92	0	0	0	0
	Squatting	13	0	0	0	0	0	0	0	2	74	962	1	0	0	0
	Slow walking	15	0	0	0	0	0	0	0	0	0	2	994	20	0	0
	Fast walking	12	0	0	0	0	0	0	0	0	0	0	7	950	0	0
	Running	4	0	0	0	0	0	0	0	0	0	0	0	0	991	0
	Twisting	29	0	0	0	0	0	0	8	0	0	0	3	0	0	910

SVM Model	Overall Accuracy: 0.955							95% CI : (0.952 , 0.958)							
	Supine	LD-R	LD-L	Prone	T-R	T-L	Bending	Standing	TRL	Bending-UD	Squatting	Slow walking	Fast walking	Running	Twisting
Sensitivity	0.83	1.00	1.00	0.97	1.00	1.00	0.84	0.97	1.00	0.92	0.91	0.99	0.98	1.00	0.93
Specificity	1.00	1.00	1.00	0.99	1.00	1.00	1.00	0.99	1.00	0.99	0.99	1.00	1.00	1.00	1.00
Percision	0.99	1.00	0.99	0.85	0.99	0.98	0.97	0.88	0.99	0.90	0.91	0.96	0.98	1.00	0.96
NPV	0.99	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.00
F-score	0.90	1.00	1.00	0.91	1.00	0.99	0.90	0.92	1.00	0.91	0.91	0.98	0.98	1.00	0.94

From Table 17 we can see the overall accuracy of the model is 95.5%. The model has a reliable performance in most categories. The best results are related to LD-R, LD-L, T-R, running and Tilting on the right and left (TRL).The comparison of F-scores between SVM model and the best random-forest model has been depicted in Figure 9.

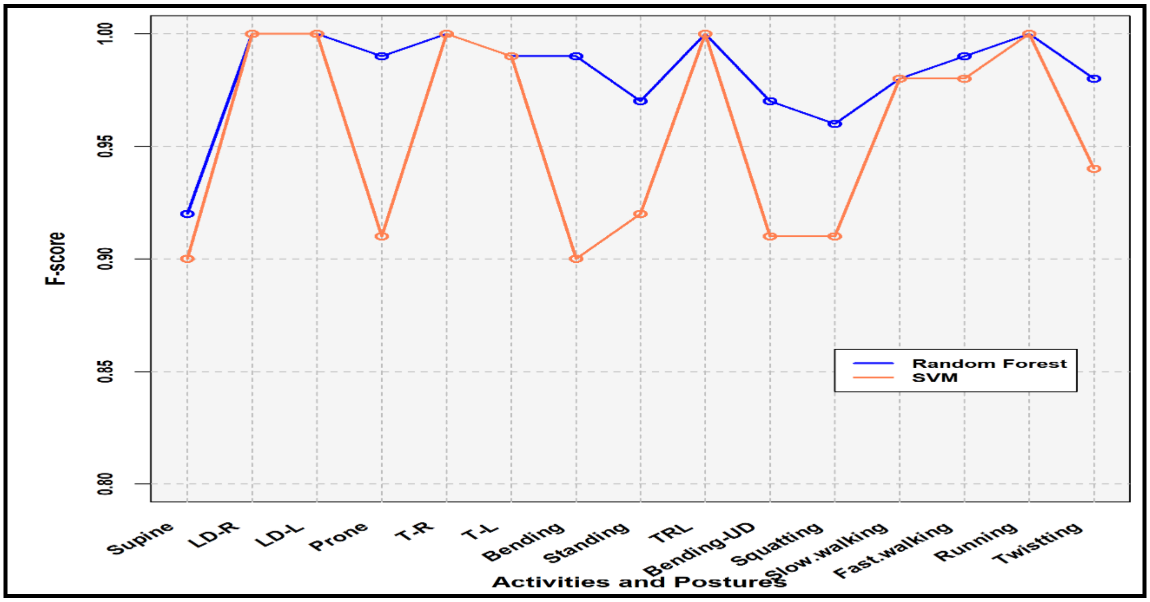


Figure 10. F- Score Results for Random Forest and SVM Models in Each Category of Static and Dynamic Activities

Figure above shows both models have the same performance in some categories such as running, slow walking, TRL, T-R, T-L, LD-R, and LD-L. The performance of the random-forest model in other classes is better than SVM model. Both models have good performance in overall. This information shows, the selected features can capture the most information about static and dynamic activities.

IV.4. Modeling Results for Combined Categories

In this section, we provided the results of the models on new categories which defined in chapter III. The time window and sample size for both models were 6 seconds and 5000, respectively.

After first combination we had eight categories; Lying-down, Standing –star, TRL, Bending-UD, Squatting, Walking, and Twisting. Table 18 shows the performance of each model in all categories.

Table 18. Statistics of Random-Forest and SVM Model in Eight Categories

Model	Random-Forest				SVM			
Overall Accuracy	0.986				0.972			
	Lying-Down	Standing-star	TRL	Bending-UD	Squatting	walking	Running	Twisting
Random-forest F-scores	0.99	0.99	1.00	0.97	0.97	0.99	1.00	0.98
SVM F-scores	0.98	0.98	1.00	0.91	0.91	0.99	1.00	0.94

From the table above, we can see that by incorporating of some categories, the overall accuracy of both models increased. The overall accuracy for random-forest and SVM model was 98.6% and 97.2%, respectively. The performance of both models in most categories are very similar but in squatting, Bending-UD, and twisting, the random forest model performed better than SVM model. This information provides more evidence which implies that our features contain the sufficient information to categories the activities.

In the next step, we combined the Lying-Down and Standing-star categories as a sedentary category. Table 19 demonstrates the performance of both models in the seven categories.

Table 19. Statistics of Random-Forest and SVM Model in Seven Categories

Model	Random-Forest			SVM			
Overall Accuracy	0.989			0.976			
	Sedentary	TRL	Bending-UD	Squatting	walking	Running	Twisting
Random-forest F-scores	0.99	1.00	0.97	0.96	0.99	1.00	0.98
SVM F-scores	0.99	1.00	0.91	0.91	0.99	1.00	0.94

From the table above, we can see that by incorporating of some categories, the overall accuracy of both models increased. The overall accuracy for random-forest and SVM model was 98.9% and 97.6%, respectively. The performance of both models in most categories are very similar but in squatting, Bending-UD, and twisting, the random forest model performed better than SVM model.

CHAPTER V

DISCUSSION, CONCLUSION, AND CONTRIBUTIONS

This chapter provides a summary of our procedures to extract some features from raw data, building the classification models, and the results obtained from the models. Lastly, potential areas of future work are presented.

V.1. Discussion

Nowadays, the positive correlation between sedentary behaviors and some chronic disease is known as a fact. During the past decades, using accelerometers to recognize the human activity has increased significantly [15,16,17,18,19,20]. Scientists used the accelerometer in different positions on the body. Some studies chose sitting, standing, walking, and running as daily activities. There are several studies which evaluate the performance of different classifiers in accelerometer data[19]. Most of them considered several features of raw data as explanatory variables to build the model at which participants' characteristics affect the accelerometer data, but those studies did not consider it as an important issue. In this study, we used the first order differencing (FOD) to remove the participant's characteristics and considered just two features as explanatory variables. We used the median of angles and area under the curve (AUC) in the certain time window. In our study, we considered fifteen different static and dynamic activities which some of them have not been performed in other studies. We evaluate the effect of the time window and sample size in the accuracy of random-forest models to recognize the activities.

Results for our models showed the selected features could capture more information about activities. Increasing the sample size can improve the accuracy of models. The performance of both classifiers was reassuring. We achieved 98.2% and 95.5% overall accuracy in fifteen different categories with random-forest and SVM models, respectively. The results of models after combining of some activities showed we have improvement in overall accuracy.

V.2. Summary of Thesis Research

In conclusion, this thesis utilized a transformation method which can remove the subject's characteristics from raw data. Also, the concept of area under the curve helped us to extract some features which contain more information of activities. Our models in both classifiers were able to recognize the activities with high accuracy. The results of random-forest models in the certain time window and the sample size was better than SVM models. The results showed our selected features provide sufficient information about activities in laboratory data.

V.3. Future Work

All work which we have done in this study was based on the laboratory data. However, in the real world, we need to evaluate the results of our model on the real-life data. The results of performing these models on real-life data are very substantial. The predictions for some situations which have not been considered in the laboratory data can be helpful to find a better way to improve our models.

V.4. Contributions

The thesis research made three major contributions to the knowledge of activity recognition and feature selection.

Previous studies have had more focus on five or six different static and dynamic activities such as sitting, standing, walking, and running. In this study, we selected the fifteen different static and dynamic activities which were designed based on accelerometer position on the body. We considered activities which have more acceleration on the upside of the body because our accelerometer was attached to the participant's chest. To the best of our knowledge, some of these activities, like squatting, tilting left and right side of the body, twisting and bending up and down have not been performed by other studies.

Our thesis research also made contributions through the number of features which were used in classifiers. Previous studies have used at least seven different features as predictors in classifiers. The accuracy for study with seven feature was less than 90%. In this study, we just used two features. The area under the curve was the feature which has not been used in other studies. We achieved to at least 95.5% and 98.2% accuracy in SVM, and random-forest models in fifteen categories which implies that the selected features for this study can capture the most important information of activities for classifiers.

The previous studies have extracted all features from raw accelerometer data without considering this fact that participant’s characteristics affect the data. In this study (1) we considered the features which are not related to the subject characteristics and (2) we used the transformation method which reduced the effect of participant’s characteristics from raw data.

REFERENCES

[1] C. J. Caspersen, K.E. Powell, G.M. Christenson, “Physical Activity, Exercise, and Physical Fitness: Definitions and Distinctions for Health-Related Research,” *Public Health Reports*, vol. 100, no. 2, pp. 126–130, 1985.

[2] S. B. R. Networ, “Letter to the Editor: Standardized Use of the Terms ‘Sedentary’ and ‘Sedentary Behaviours’,” *Applied Physiology, Nutrition, and Metabolism*, vol. 37, no. 3, pp. 540–542, 2012.

[3] B. E. Ainsworth, W. L. Haskell, A. S. Leon, D. R. Jacobs Jr, H. J. Montoye, J. F. Sallis, R. S. Paffenbarger Jr, “Compendium of Physical Activities: Classification of Energy Costs of Human Physical Activities.” *Medicine and Science in Sports and Exercise*, vol. 25, no. 1, pp. 71-80, 1993.

[4] Alberta Centre for Active Living. [Online]. Available: https://www.centre4activeliving.ca/news/2015/11/standing-sedentary-behaviour/. [Accessed: 02-Oct-2017].

[5] “Measuring Physical Activity,” The Nutrition Source, 31-Jan-2017. [Online]. Available: https://www.hsph.harvard.edu/nutritionsource/mets-activity-table/. [Accessed: 17-Sep-2017].

[6] P. F. Saint-Maurice, Y. Kim, G. J. Welk, G. A. Geasser, “Kids are Not Little Adults: What MET Threshold Captures Sedentary Behavior in Children?” *European Journal of Applied Physiology*, vol. 116, no. 1, pp. 29-38, 2016.

[7] “SparkFun Triple Axis Accelerometer Breakout - ADXL335,” *Learn at SparkFun Electronics*. [Online]. Available: https://learn.sparkfun.com/tutorials/accelerometer-basics.[Accessed: 19-Sep-2017].

[8] “A Beginner's Guide to Accelerometers,” *Dimension Engineering*. [Online]. Available: https://www.dimensionengineering.com/info/accelerometers. [Accessed: 19-Sep-2017].

[9] “Physical Activity,” *World Health Organization*. [Online]. Available: http://www.who.int/topics/physical_activity/en. [Accessed: 13-Sep-2017].

[10] W. L. Haskell, S. N. Blair, and J. O. Hill, “Physical activity: Health Outcomes and Importance for Public Health Policy,” *Preventive Medicine*, vol. 49, no. 4, pp. 280–282, 2009.

[11] A. A. Thorp, N. Owen, M. Neuhaus, and D. W. Dunstan, “Sedentary Behaviors and Subsequent Health Outcomes in Adults,” *American Journal of Preventive Medicine*, vol. 41, no. 2, pp. 207–215, 2011.

[12] “Physical Activity Guidelines,” *Physical Activity Guidelines*. [Online]. Available: https://health.gov/paguidelines/. [Accessed: 20-Aug-2017].

[13] R. R. Pate, M. Pratt, S. N. Blair, W. L. Haskell, C. A. Macera, C. Bouchard, D. Buchner, W. Ettinger, G. W. Heath, A. C. King, A. Kriska, A. S. Leon, B. H. Marcus, J. Morris, R. S. Paffenbarger, K. Patrick, M. L. Pollock, J. M. Rippe, J. Sallis, J. H. Wilmore, “Physical Activity and Public Health. A Recommendation from the Centers for Disease Control and Prevention and the American College of

Sports Medicine,” *JAMA: The Journal of the American Medical Association*, vol. 273, no. 5, pp. 402–407, Jan. 1995.

[14] F. B. Hu, R. J. Sigal, J. W. Rich-Edwards, G. A. Colditz, C. G. Solomon, W. C. Willett, F. E. Speizer, and J. E. Manson, “Walking Compared With Vigorous Physical Activity and Risk of Type 2 Diabetes in Women,” *Jama*, vol. 282, no. 15, pp.1433–1439, 1999.

[15] T. Bastian, A. Maire, J. Dugas, A. Ataya, C. Villars, F. Gris, E. Perrin, Y. Caritu, M. Doron, S. Blanc, P. Jallon, and C. Simon, “Automatic Identification of Physical Activity Types and Sedentary Behaviors from Triaxial Accelerometer: Laboratory-Based Calibrations Are Not Enough,” *Journal of Applied Physiology*, vol. 118, no. 6, pp. 716–722, 2015.

[16] T. G. Pavey, N. D. Gilson, S. R. Gomersall, B. Clark, and S. G. Trost, “Field Evaluation of a Random Forest Activity Classifier For Wrist-Worn Accelerometer Data,” *Journal of Science and Medicine in Sport*, vol. 20, no. 1, pp. 75–80, 2017.

[17] M. E. Rosenberger, W. L. Haskell, F. Albinali, S. Mota, J. Nawyn, and S. Intille, “Estimating Activity and Sedentary Behavior from an Accelerometer on the Hip or Wrist,” *Medicine & Science in Sports & Exercise*, vol. 45, no. 5, pp. 964–975, 2013.

[18] A. G. Bonomi, A. H. Goris, B. Yin, and K. R. Westerterp, “Detection of Type, Duration, and Intensity of Physical Activity Using an Accelerometer,” *Medicine & Science in Sports & Exercise*, vol. 41, no. 9, pp. 1770–1777, 2009.

[19] A. Bayat, M. Pomplun, and D. A. Tran, “A Study on Human Activity Recognition Using Accelerometer Data from Smartphones,” *Procedia Computer Science,* vol. 34, pp. 450–457, 2014.

[20] S. I. D. Vries, F. G. Garre, L. H. Engbers, V. H. Hildebrandt, and S. V. Buuren, “Evaluation of Neural Networks to Identify Types of Activity Using Accelerometers,” *Medicine & Science in Sports & Exercise*, vol. 43, no. 1, pp. 101–107, 2011.

[21] B. C. Galland, G. J. Kennedy, E. A. Mitchell, and B. J. Taylor, “Algorithms for Using an Activity-Based Accelerometer for Identification of Infant Sleep-Wake States During Nap Studies,” *Sleep Medicine*, vol. 13, no. 6, pp. 743–751, 2012.

[22] I. C. Gyllensten and A. G. Bonomi, “Identifying Types of Physical Activity With a Single Accelerometer: Evaluating Laboratory-Trained Algorithms in Daily Life,” *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 9, pp. 2656–2663, 2011.

[23] C. B. Anderson, M. Hagstromer, and A. Yngve, “Validation of the PDPAR as an Adolescent Diary: Effect of Accelerometer Cut Points,” *Medicine & Science in Sports & Exercise*, vol. 37, no. 7, pp. 1224–1230, 2005.

[24] K. Zhang, P. Werner, M. Sun, F. X. Pi-Sunyer, and C. N. Boozer, “Measurement of Human Daily Physical Activity,” *Obesity Research*, vol. 11, no. 1, pp. 33–40, 2003.

[25] A. Mackey, N. Stott, and S. Walt, “Reliability and Validity of an Activity Monitor (IDEEA) in the Determination of Temporal–Spatial Gait Parameters in Individuals with Cerebral Palsy,” *Gait & Posture*, vol. 28, no. 4, pp. 634–639, 2008.

[26] L. Bao and S. S. Intille, “Activity Recognition from User-Annotated Acceleration Data,” *Lecture Notes in Computer Science Pervasive Computing*, pp. 1–17, 2004.

[27] A. Mannini, S. S. Intille, M. Rosenberger, A. M. Sabatini, and W. Haskell, “Activity Recognition Using a Single Accelerometer Placed at the Wrist or Ankle,” *Medicine & Science in Sports & Exercise*, vol. 45, no. 11, pp. 2193–2203, 2013.

[28] Human Activity Recognition from Accelerometer Data Using Convolutional Neural Network - IEEE Conference Publication. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7881728. [Accessed: 25-Oct-2017].

[29] 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, Jan. 2010.

[30] R. H. Shumway and D. S. Stoffer, Time Series Analysis and Its Applications: with R Examples. Cham, Switzerland: Springer, 2017.

[31] D.M.W. Powers, “Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation,” *Machine Learning Technologies*, vol. 1, no. 1, pp. 37–63, 2011.