

Extraction of Information from Inertial Sensors to Aid Health Assessment

A Thesis

Presented to
the faculty of the School of Engineering and Applied Science
University of Virginia

in partial fulfillment
of the requirements for the degree

Master of Science

by

Sriram Raju Dandu

August

2016

APPROVAL SHEET

The thesis
is submitted in partial fulfillment of the requirements
for the degree of
Master of Science



AUTHOR

The thesis has been read and approved by the examining committee:

Dr. John Lach

Advisor

Dr. James Aylor

Dr. Laura Barnes

Accepted for the School of Engineering and Applied Science:



Craig H. Benson, Dean, School of Engineering and Applied Science

August

2016

Extraction of Information from Inertial Sensors to Aid Health Assessment

Sriram Raju Dandu

August, 2016

Abstract

Tracking human activity and fitness have been the latest trends in the world of wearable technology. Inertial sensors offer a practical and low cost method for objectively monitoring human movements, and particularly have the applicability to monitor free-living subjects. Shoes, smart phones, watches and other wearables equipped with inertial sensors are able to capture human motion in a non-obtrusive way. Similar sensors are placed on several positions on the human body to monitor a range of different movements, including gait, sit-to-stand transfers, postural sway and limb movements. Inertial sensors have provided a simple and reliable approach to collect human kinematics data. The inertial data collected by these sensors contains valuable information related to human movements, which can be used to monitor and assess the health of a person. But the field of wearable technology still struggles to develop a methodology which can derive relevant information, from the raw inertial data, to aid in health assessment.

This work presents a framework to extract information from raw inertial data corresponding to human motion. The framework is a sequential learning process coupled with application-specific knowledge. This framework has three stages: data preparation, feature engineering and information extraction. The first stage converts the raw data into usable data, the second stage extracts relevant attributes from the data and the third stage derives information from these attributes to aid health assessment. A number of techniques can be used at each of the stages, and the appropriate ones are chosen by employing application-specific knowledge as the cornerstone.

I will be implementing the framework on two real world case studies; one studies the relationship between inertial features and gait pathology in Multiple Sclerosis and the other one deals with classification of head impacts in contact sports using inertial sensors. The information derived from both the case studies has the potential to serve as a tool to monitor the health of a person.

Acknowledgement

First and foremost I offer my sincerest gratitude to my advisor, John Lach, who has supported me throughout my Masters' program with his patience and knowledge whilst allowing me the room to work in my own way. This thesis would not have been completed without his encouragement and effort. I simply could not wish for a better or friendlier advisor, and am grateful to him for taking me on as his student. I would also like to thank James Aylor and Laura Barnes for being a part of Masters' Thesis committee.

I have been blessed with a friendly and cheerful group of teammates at INERTIA. My special thanks to Jiaqi Gong and Matthew Engelhard, who guided me through the case studies covered in this thesis. I must acknowledge Luis Lopez and Victor Sobral for helping me get accustomed to the American culture in my early days. I would also like to thank Ben Ghaemmaghami for his help in editing this thesis.

I would like to sincerely thank all the medical collaborators I worked with: Myla Goldman, Jason Druzgal, Tom Bowman and Bryson Reynolds.

My deepest gratitude goes to my parents and family, without whose support I would not have come halfway across the world. I also would like to acknowledge University of Virginia and my friends for the everlasting and fun memories.

Finally, this work was supported in part by the National Science Foundation under grants IIS-1065262, IIS-1231712, and IIS-1418622.

Table of Contents

ABSTRACT.....	II
ACKNOWLEDGEMENT.....	III
1 INTRODUCTION.....	1
1.1 CASE STUDY 1	3
Preface	3
Data.....	4
Contribution.....	5
1.2 CASE STUDY 2	6
Preface	6
Data.....	7
Contribution.....	8
1.3 THESIS OUTLINE.....	8
2 DATA PREPARATION.....	9
2.1 BACKGROUND	9
2.2 CASE STUDY 1	10
2.3 CASE STUDY 2	14
3 FEATURE ENGINEERING.....	16
3.1 BACKGROUND	16
3.2 CASE STUDY 1	17
3.3 CASE STUDY 2	20
4 INFORMATION EXTRACTION.....	25
4.1 BACKGROUND	25
4.2 CASE STUDY 1	26
4.3 CASE STUDY 2	34
5 CONCLUSION	52

5.1 CASE STUDY 1	52
5.2 CASE STUDY 2	53
5.3 FUTURE WORK	54
APPENDIX.....	56
A. PERFORMANCE METRICS	56
R^2	56
Adjusted R^2	56
Akaike information criterion (AIC)	57
Bayesian information criterion (BIC)	57
p value.....	57
B. CLASSIFICATION METRICS.....	58
Sensitivity	58
Specificity	58
ROC & AUROC	58
C. DIAGNOSTIC CURVES	59
Residuals vs Fitted	59
Normal Q-Q	59
Scale-Location	59
Residuals vs Leverage.....	60
D. STATISTICAL MEASURES OF THE CLINICAL SCORES	60
REFERENCES.....	62
LIST OF PUBLICATIONS	70

Chapter 1

Introduction

Over the past decade, the use of mobile sensors has become an integral part of the modern lifestyle. With the advent of micro-electro-mechanical systems (MEMS), the physical size of sensors was reduced to the order of millimeters, thereby making them easier to integrate with systems. Systems are able to analyze situations and make decisions with help of sensors, thereby transforming them into smart systems. Sensors are now employed in a wide range of applications from detecting DNA sequences to monitoring the environmental conditions on Mars.

One of the applications, where sensors have been immensely useful is fitness tracking of human subjects. Over the past few years, kinematic measurements have been widely used by clinicians for medical diagnosis[1]–[5]. Such measures are used to quantify both normal and pathological movements, quantify the degree of impairment, plan rehabilitation strategies and assess the effect of various interventions. Small, low-powered inertial sensors have helped with the collection of kinematic measurements from human motions and provided the potential for dynamic three-dimensional motion analysis. These sensors measure rigid kinematic motion, such as acceleration by accelerometers and angular velocity by gyroscope. They are currently used in real world scenarios to gather accurate and high resolution data corresponding to human motion.

Inertial sensors are used to study human kinematics for a wide variety of applications [1]. Lieber et al [4] described the role played by accelerometers in recording and predicting falls, recording and predicting freezing of gait, and evaluating postoperative recovery of subjects. Goodwin et al [6] developed a dynamical system using inertial sensors to study the spinal motion. Jamsa et al [7] showed that there exists a significant relationship between accelerometer-based data and proximal femur bone mineral

density. Wong et al [8] monitored the human trunk movements using inertial sensors. Daukantas et al [9] utilized an accelerometer based sensor for extracting information related to the performance of a swimmer. Gong and Lach [10] employed inertial sensors to enhance the robotic surgical training process. The location of the sensors and the information extracted from the inertial data are different across applications, but they share elements such as noise and human motion data. This noise may be due to improper placement of the sensor, infidelity of sensor recordings, unpredictable real-world scenarios or the large degree of freedom human action possesses. Thus, identifying subtle human movements and deriving corresponding information to address a problem is an arduous task. In order to tackle this issue a framework which incorporates domain-specific knowledge should be used.

A general framework for extraction of information from raw inertial data is already established for human activity recognition. There has been a lot of research done in this area and some of the highly cited papers in this domain are [11]–[15]. Randell et al [11] computed 4 features and employed neural networks to cluster the activities. Bao et al [13] used decision tree classifier with temporal features like mean, energy, frequency-domain entropy, and correlation of acceleration data as features to classify activities. Laerhoven Van et al [12] employed an algorithm which takes the sensor data as input and generates a probability/confidence value for a particular activity. Similarly, Kern et al [14] exploited bayes classifier with the raw sensor data as input to recognize activities. Ward et al [15] used the number of peaks within a 100 ms sliding window, the mean amplitude of these peaks, and the raw x-axis data as features for a hidden Markov model to classify activities. All these papers tend to follow a common procedure for activity recognition, i.e. feature engineering followed by implementation of a classifier. The authors tend to do an exhaustive search within the available features and the classifiers to choose the right ones. But they do not discuss the reasons behind choosing the specific features or the specific classifier.

In this work, I will integrate application-specific knowledge into the framework to derive features and utilize machine learning techniques which are coherent with the problems statement. I will demonstrate the working of the framework and simultaneously implement it on two different case studies. The

framework I utilize is a sequential learning process, where investigation follows a logical stepwise path to find solutions. The sequential learning framework (SLF) in my work consists of three steps; data preparation, feature engineering and information extraction. The SLF coupled with application-specific knowledge will address the two case studies mentioned below.

1.1 Case Study 1

Preface

Multiple sclerosis (MS) is a chronic progressive neurological disorder affecting about 2.5 million people globally [16]. Impaired mobility is a common symptom of MS even at lower levels of the disease and has significant negative effects on quality of life [17]. As a result, walking assessment is critical in tracking the progression of the MS disease.

There are various techniques to study human gait that are often quick and simple to use; however, such systems often lack valuable kinematic data. Tape measures and goniometers provide information in single planes and only for static positions. Electro-goniometers and inclinometers offer solutions for more than one plane, as well as provide dynamic data; however, the physical design of such sensors can restrict motion. Therefore, because of measurement drawbacks, it remains difficult for the clinician to gain information about dynamic three-dimensional movements. In contrast, laboratory systems are complex and expensive, but are capable of capturing high quality three-dimensional movements. Two laboratory systems commonly found within the literature to study gait are force plate systems[18]–[20] and video-based optoelectronic systems[2], [5], [21], but both the systems are time-consuming and complex [22]. Thus, employing inertial sensors is a better alternative to study gait with higher precision, compared to the aforementioned systems. Inertial sensors are portable, inexpensive and can record high resolution gait data. Figure 1 shows some of the inertial sensors that are used to study human gait.



Figure 1: Examples of commonly used inertial sensors for human motion tracking

Recently, the gait impairment in MS subjects has been measured by performing the Six Minute Walk (6MW) Test. Under the 6MW, the subject is asked to walk for six minutes as swiftly as possible to identify motor fatigue. Chetta et al [23] focused on the cardiorespiratory response during the 6MW in 11 MS subjects with mild disability. The authors showed that 6MW distance correlated to disability score, but not subjective fatigue. Goldman et al [24] evaluated a modified 6MW in MS subjects with varying disability and controls to assess test reliability. The authors also assessed 6MW distance correlation to subjective measures of fatigue, physical function and ambulation. However, the correlation of 6MW distance was limited to certain physiological parameters, and there remains a need to identify other gait features from 6MW that may provide additional insight into gait impairment in MS. In order to address this issue, the motion of the subjects during the 6MW must be monitored with a higher precision.

The next chapters will explain the techniques that were incorporated into the sequential learning framework (SLF) to extract gait features from 6MW and determine their physiological significance.

Data

The gait and mobility of 115 participants were assessed at the University Of Virginia Department Of Neurology. There were 86 MS subjects and 29 controls. The subjects were asked to perform the 6MW in

a hallway with an inertial sensor attached to their dominant hip. For this study, an off the shelf inertial sensor was used to collect the gait data, namely, Actigraph [25]. This inertial sensor has a triaxial accelerometer with a sampling frequency of 30 Hz. The clinicians recorded the distance covered by the subjects for each minute manually, which is the average walking speed for the corresponding minute. The clinicians also conducted three clinical tests, for every 6MW, which include patient-completed questionnaires and functional tests. The three clinical assessments were:

1. The Expanded Disability Status Scale (EDSS) [26] is a ten-point scale, where zero indicates a normal neurological exam and a ten indicates death from MS. This physician-assessed composite score is based on walking ability and neurological examination of seven functional systems: pyramidal, cerebellar, brainstem, sensory, vision, bowel and bladder, and cognitive.
2. The Multiple Sclerosis Walking Scale (MSWS) [27] is a self-reported measure of the impact of MS on walking ability consisting of 12 physical items. The total score ranges from 0 to 100, with higher values indicating greater impairment.
3. The Modified Fatigue Impact Scale (MFIS) [28] is another self-reported measure concerning the impact of fatigue on daily life. The total score is the sum of the scores for the 21 items, each ranging from 0 to 4. Individual subscale scores for physical, cognitive, and psychosocial functioning can also be generated by calculating the sum of specific sets of items.

Contribution

I derived metrics based on the subtle variations among the gait cycles across the six minutes and later established a relationship between the metrics and the gait pathology of the subject.

1.2 Case Study 2

Preface

Concussions* are common in contact sports such as football, soccer and lacrosse, with approximately 1.6 – 3.8 million sport-related concussions reported in the USA annually [29]. In many cases, concussions are caused by a blow to the head or the head and upper body being violently shaken. Concussion symptoms include headache, lack of concentration, and loss of memory, judgment, balance, and/or coordination. Every concussion incurs injures the brain to some extent [30]. Previously, research found statistically significant correlation between the number of head impacts and the resulting neurophysiological deficit [31]. Furthermore in [32], authors showed that repeated sub-concussive impacts have also been associated with blood-brain barrier disruption, indicating elevated risks for cognitive deficits. However, the exact etiology of concussion is unclear.

The identification of risk factors that predispose an athlete to concussion will help medical experts understand the underlying mechanisms of concussion and aid in the improvement of prevention strategies. One way to identify the root cause of concussions is by identifying the relation between number of direct hits to the player's head and brain injury. In order to study this, there is a need to find the number of direct hits to the head (true hits) a player sustains during a game/season. A simple and straightforward way to check the number of true hits is by watching the game and manually counting the hits. But it requires a significant amount of human effort to monitor every player on field and keep track of the hits. So, an automated way to identify the true hits will be effective.

An inertial sensor can be used to collect motion data of players' heads and an automated algorithm can be implemented to differentiate the true hits from clacks†. There is no available work in the literature, to the best of the author's knowledge, which tried to identify the true hits from real world accelerometer data.

* Concussion is a traumatic brain injury characterized by immediate and transient alteration in brain function.

† Shaking of head due to impacts, which are not related to play, such as kicking and jumping are called clacks.

So, establishing a technique for detecting the true head impacts will serve as research tool to help investigate relationship between direct head impacts and concussions experienced by a player.

Upon investigating several temporal features, the estimation of stimuli* was regarded as the best feature to identify true hits. This is because the manner in which players react after sustaining a similar impact to the head can be quite varied. This reaction of the players depends on a several human factors like their age, physical fitness, impact preparedness, etc. Thus, the inertial data may not be the same if two players sustain the same hit. In this work, I will present an algorithm using SLF for estimating the stimuli of the impacts.

Data

An off the shelf inertial sensor, X2 Biopatch [33], was used to track and record the acceleration of the head of the players. The sensor was placed on the mastoid process part of the skull, rather than on a helmet [34]. A helmet does not provide a clear picture of the motion of the skull because of the various protective layers embedded in them. Placing a sensor on the mastoid process gives us more precise motion data of the skull and will not obstruct game play. Figure 2 shows the sensor placed on a person. The sampling frequency of the sensor was 1000 Hz. The data collection study took place at Lynchburg College, where 16 players from Women Lacrosse team participated in the study.



Figure 2: X2 Biopatch placed on the players' head

* A stimulus is the event that causes the head motion. Later sections will explain its significance and derivation.

The sensor was used for 51 sessions, including games and practice sessions. The video footage was provided for some games and practice sessions, which was used to label the impacts as true hits and clacks. The X2 Biosystems software computed the following the metrics, which were available for analysis:

1. Peak linear acceleration (PLA)
2. Peak rotation acceleration (PRA)
3. Peak linear velocity (PLV)

Contribution

I have estimated the stimuli (that causes the head movement) of the recorded head impacts. Further, I built a classifier using this estimated signal to differentiate the true hits from clacks and compared the performance with other features.

1.3 Thesis Outline

Given the outlook and impact of this research, a summary of the thesis can be stated as follows:

This work utilizes a framework incorporated with application-specific knowledge to extract the meaningful information from raw inertial data corresponding to human motion. This framework is implemented on two medical case studies to derive information which can aid in health assessment.

The rest of this thesis is organized as follows: After the introduction of the framework and the case studies in Chapter 1, the steps taken to convert the raw data into usable data are presented in Chapter 2. Then, Chapter 3 contains the techniques employed to extract the features in both the case studies. Chapter 4 studies the relationship between the derived features and the available ground truth using machine learning techniques. Chapter 5 will conclude the thesis and layout the future work.

Chapter 2

Data Preparation

2.1 Background

In order to analyze data for making predictions/classification, one of most important and strenuous step is data preparation. Data Preparation is the process of collecting, cleaning, and consolidating data into usable files or data tables for use in analysis. The process of preparing data generally entails correcting any errors (typically from human and/or machine input), filling in nulls and incomplete data, and transforming data into the desired format for analysis. The eventual output of the step is to have corrected data along with the corresponding ground truth.

Data preparation is a common step in big data analysis [35], web usage mining [36] and text mining [37]. Cooley et al [36] proposed data preparation techniques and algorithms to convert raw web server logs into user session files in order to perform web usage mining . The authors presented models to encode both the web site developers' and users' view of how a web site should be used to identify web site users, user sessions, and page accesses that are missing from a web server log. Mclellan et al [37] employed data preparation techniques, developed for several Centers for Disease Control and Prevention, to organize textual data. The authors laid out a set of rules and considerations for managing text data.

There is far lesser focus on this step when dealing with human motion data. While handling inertial data, the biggest task lies in segmenting the required data and labelling them accordingly. This step plays an important role when the data collection happens in the real world as most often the ground truth is not available in the desired format.

In this chapter, for case study 1, the 6MW data was derived from the raw accelerometer data and the clinical scores were assigned to the respective 6MW test. For case study 2, the impacts recorded by the sensor were labelled as a true hit/clack for analysis.

2.2 Case Study 1

The accelerometer data, walking speed and clinical scores were manually surveyed to find missing data and other errors. Subjects with missing clinical scores and walking speed data were not included in the analysis. So that the clinical scores and walking speed associated with the accelerometer data, is available for analysis.

The first task was to extract the 6MW walking data from the accelerometer data. Since the sensors start recording the data even before the subjects start walking, it tends to collect data which is not related to the 6MW test. In order to extract the 6MW data from the accelerometer data, a GUI was developed in MATLAB® to manually separate the region of interest. The user had to manually mark the start point and the end point of the 6MW using the cursor. Figure 3 displays the screenshot of the GUI. 6MW inertial data was extracted for all the subjects using this GUI.

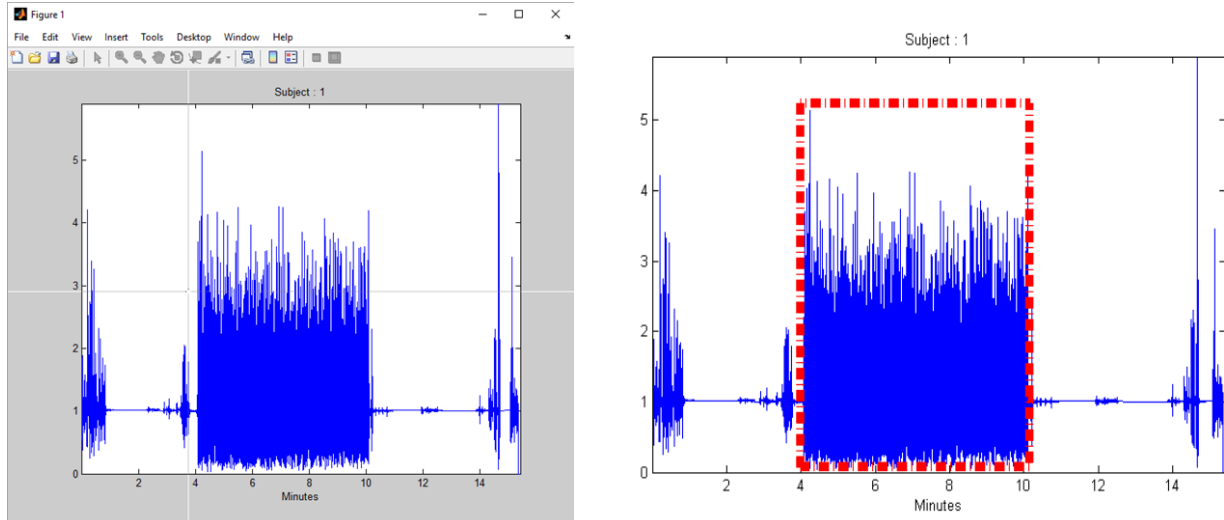


Figure 3: The left image shows the GUI interface where the perpendicular lines are used select the start and the end points of the 6MW and the red dotted lines in the right image encompass the region of interest.

Once the 6MW accelerometer data was obtained, all the gait cycles had to be extracted to identify the subtle variations among gait cycles. Gait is comprised of sequential gait cycles and each such gait cycle is composed of a sequence of events that mark the transition from one gait phase to another. In terms of temporal domain parameters, the two most relevant events in a normal gait cycle are the initial heel contact or heel strike (HS) and terminal contact or toe off (TO). So detecting HS and TO accurately is of vital importance in clinical gait analysis to study the variation of gait patterns in high resolution. The gait patterns of the subjects in our study are quite variable, as shown in Figure 4, due to the heterogeneity of gait pathology in MS.

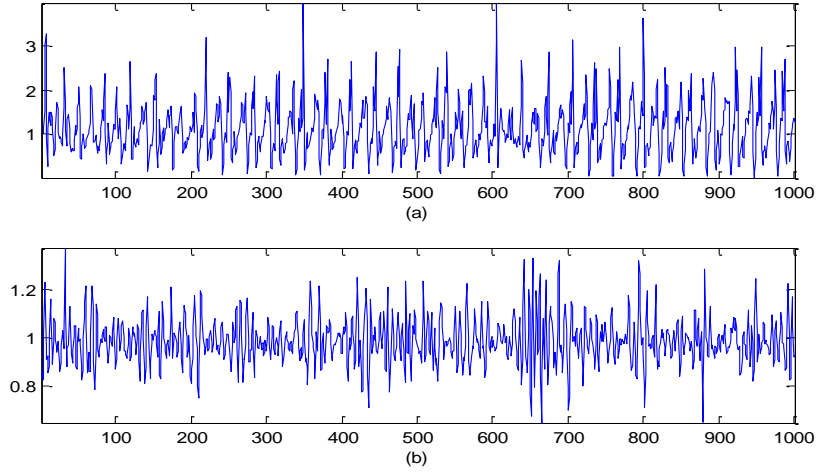


Figure 4: Magnitude* of six minute walk accelerometer data. (a) and (b) show 1000 data points from two different subjects.

In order to handle the wide variety of the gait patterns, our gait cycle extraction algorithm employs Fourier analysis to derive the gait cycles from the 6MW inertial data. Walking is predominantly a low frequency activity and, as such, Fourier coefficients representing high frequency signal content have very low amplitude. Most of the useful information closely related to the impact acceleration is contained in the band below 17 Hz [38]. Moreover, frequencies greater than 0.25 Hz enable us to effectively separate body acceleration (BA) and gravity acceleration (GA) components [39]. For our purpose, a bandpass filter with the frequency range 1 – 3 Hz is applied to the 6MW accelerometer data. Figure 5 shows the frequency spectrum of the 6MW data after bandpass filtering.

* Magnitude = $\sqrt{x^2 + y^2 + z^2}$

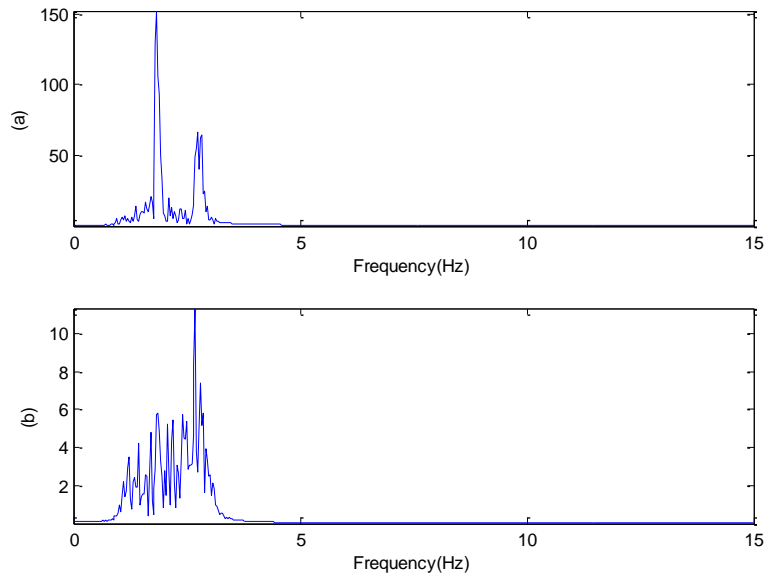


Figure 5: Frequency of 6MW after applying the bandpass filter for the data sample shown in Figure 4.

The maximum peak of the frequency spectrum was considered as the fundamental cycle frequency. This fundamental frequency is directly related to the average step rate of the six-minute walk. The gait length was calculated by dividing the sampling frequency of the accelerometer by the fundamental frequency of the 6MW. To reduce computational complexity, data segments of twice the corresponding gait length were extracted, with no overlap, from 6MW inertial data. Later, these segments of data were considered as the gait cycles. The gait cycles for each of minute from 6MW were obtained and stored. Figure 6 shows the gait cycles extracted using this technique.

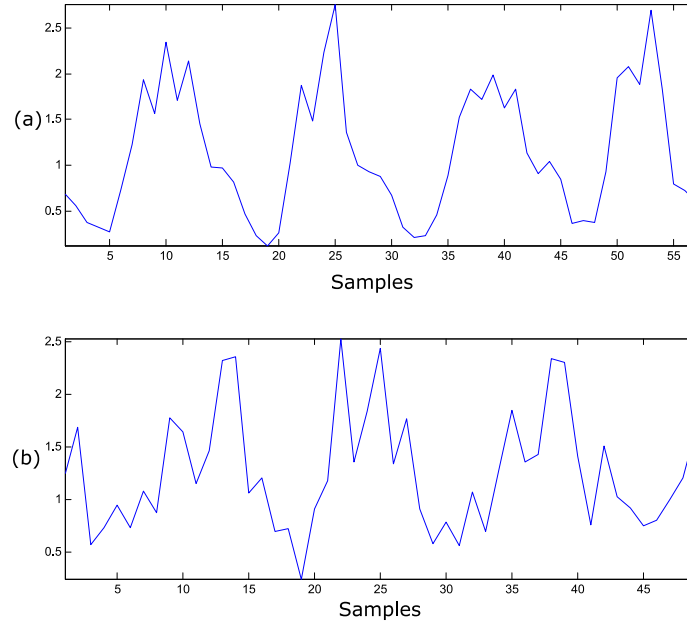


Figure 6: Gait cycle chunks from data samples shown in Figure 4.

For measuring the variability between gait cycles, the cycles should be in-phase. In other words, the gait cycles should start from the same instance such as heel strike. But the drawback of this gait cycle extraction technique is that all the extracted gait cycles might not be in-phase. Some gait cycles may start with heel strike and others with toe off. This inconsistency across gait cycles is solved in chapter 3 by aligning the gait cycles using a phase invariant DTW method.

2.3 Case Study 2

For recording the head impacts in lacrosse, players were equipped with a X2 Biopatch sensor throughout the game. The sensor continuously reads the values from the accelerometer, and saves the data if the acceleration is above a threshold (i.e. 10g). The sensor saves the ten samples prior and eighty-nine samples post to the threshold point. The recorded 100 samples of data are called impacts. For each of the impacts, the X2 Biosystems software computed the PLA, PRA and PLV values. There were 1862 impacts recorded in total and the video footage was used to label each impact. But out of the 51 sessions, video

footage for just 40 sessions was available. Labelling each of the impacts required manual effort due to inconsistent timing, poor video quality, absence of players in the video frame, etc. These issues were addressed through manual supervision, and finally achieved 252 labelled impacts which could be used for analysis. Out of the 252 impacts, 214 were true hits and 38 were clacks. Figure 7 shows an example of a true hit and clack^{*}.

The stimulus for each of the labelled impacts is estimated in the next chapter using a linear dynamical model (LDM).

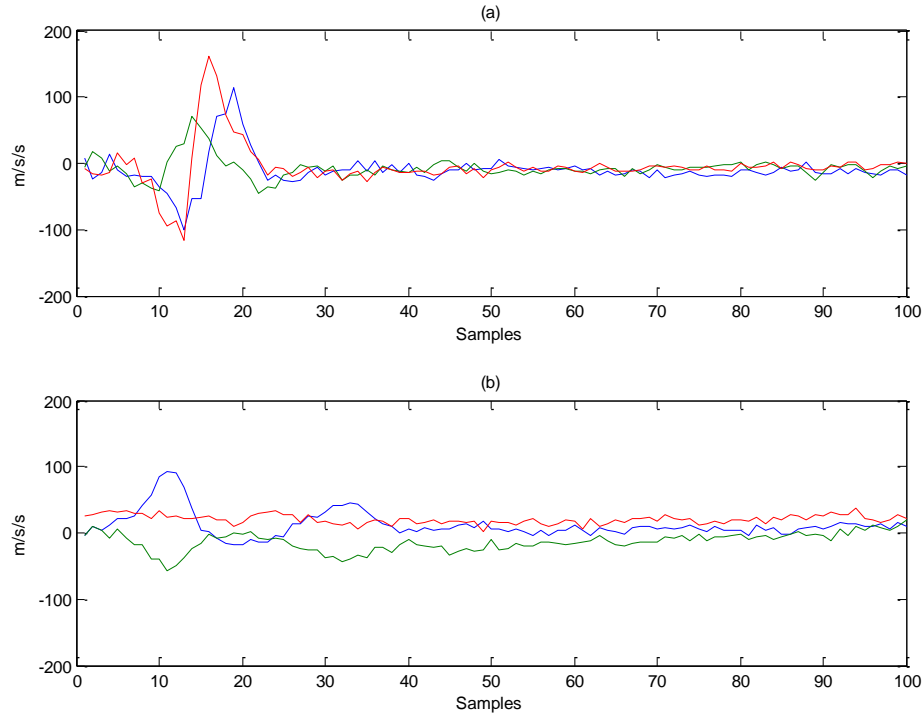


Figure 7: Sensor output of (a) true hit and (b) clack.

^{*} Shaking of head due to impacts, which are not related to play, such as kicking and jumping are called clacks.

Chapter 3

Feature Engineering

3.1 Background

A feature is a piece of information that can be used for prediction/classification. The features derived from the data will influence the results that are achieved from the predictive models. The process of using domain knowledge of the data to create attributes is called feature engineering. There are a number of feature extraction tools available in the literature which can be applied to inertial data.

In [40]–[42], authors derived time-domain features from inertial data which include the mean, median, variance, skewness, kurtosis and inter- quartile range. Aminian et al [43] computed the cross-correlation coefficients between accelerometer axis. Foerster et al [44] employed fast Fourier transform (FFT) to calculate the median of the spectral distribution, whereas Preece et al [45] used a subset of the FFT coefficients as features. Huynh et al [46] derived the spectral energy, which is the sum of the squared FFT coefficients, whereas Bao et al [13] computed the frequency-domain entropy, which is the normalized information entropy of the FFT components. In [38], [47], [48] authors used wavelet analysis, on inertial data, to calculate wavelet parameters based on the sum of the squares or RMS* of specific detail coefficients. Similarly, Wang et al [49] derived wavelet parameters using simple statistical measures, such as standard deviation and RMS, of specific approximation and detail coefficients.

In this chapter, for case study 1, features which measured the variability of the gait cycles across 6 minutes were computed. For case study 2, upon examining the temporal features derived from the inertial

* RMS stands for Root Mean Square

data and understanding the complexity of the problem, the stimuli of the impacts were estimated using a linear dynamical model.

3.2 Case Study 1

To study the subtle changes in gait cycles across the six minutes, the gait cycles from minutes 2 – 6 were compared with the gait cycles of minute 1. The changes of gait patterns were studied for every subject, rather than comparing gait cycles across subjects as walking patterns vary amongst subjects. The most common way to compare two sequences is by computing the Euclidean distance. But Euclidean distance is ineffective when the sequences are temporally scaled. Thus to avoid this issue, we employed an alternative algorithm, dynamic time warping (DTW), to measure the similarity between two temporal sequences which may vary in speed. An example of showing the difference between the working of DTW and Euclidean distance is shown in Figure 8. Assume that the green and blue curves are two different sequences. The black lines show how the two distances compare the signals. Euclidean distance follows a linear temporal mapping between the two sequences, whereas DTW distance adopts a non-linear temporal mapping.

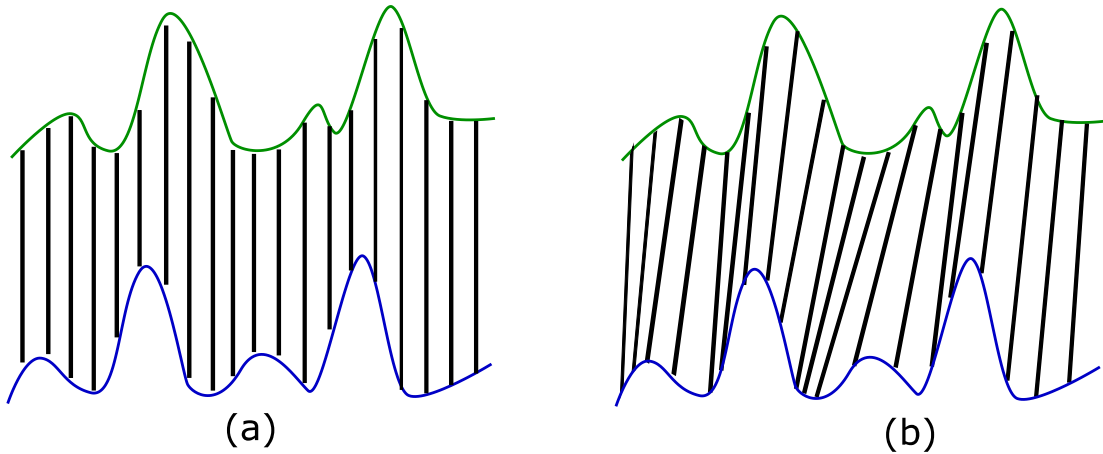


Figure 8: (a) represents Euclidean distance and (b) represents DTW

DTW is commonly used to analyze temporal sequences of video, audio, and graphics data. A well-known application is in automatic speech recognition [50], to cope with different speaking speeds. The algorithm determines an optimal match between two given sequences and returns a distance measure known as the DTW distance. The sequences are warped nonlinearly in the time dimension to match each other as closely as possible. We have defined the number of warps to align both the sequences as the warping length.

Since the gait cycles from a particular 6MW were of the same length, let's consider P and Q to be two gait cycles of length n each.

$$P = [p_1, p_2, p_3, \dots, p_n]$$

$$Q = [q_1, q_2, q_3, \dots, q_n]$$

where p_i and q_i are 3-dimensional vectors with x, y and z acceleration values respectively.

These two sequences P and Q were provided as inputs to the DTW algorithm and the outputs were the DTW distance between P and Q and the warped sequences P^w and Q^w , given by the following:

$$P^w = [(p_1)^{a_1}, (p_2)^{a_2}, (p_3)^{a_3} \dots (p_n)^{a_n}]$$

$$Q^w = [(q_1)^{b_1}, (q_2)^{b_2}, (q_3)^{b_3} \dots (q_n)^{b_n}]$$

where $(s)^k$ means k repetitions of the sample s.

The DTW distance is the Euclidean distance between P^w and Q^w , shown in (1). The warping length between the sequences was computed using the equation (2),

$$\text{DTW Distance} = |P^w - Q^w|_2 \quad (1)$$

$$\text{Warping Length} = \frac{1}{2} \left(\sum_i a_i + \sum_j b_j \right) \quad (2)$$

The hurdle of misalignment of gait cycles, as mentioned in chapter 2, was tackled using a phase-invariant implementation of dynamic time warping (DTW). To make sure the gait cycles are in-phase, one of the gait cycles (either P or Q) was shifted in a cyclic manner till the best match was found, before running the DTW algorithm. Keogh lower bounding and early DTW termination were used to find a computational optimal solution [51].

For each subject, all the gait cycles from the first minute were matched with every other gait cycle from rest of the minutes using the DTW algorithm explained above. For instance, to compute the DTW score and Warp score for second minute, all the gait cycles from minute 2 were matched with all the cycles from minute 1. Now each gait cycle from minute 2 has a DTW distance and a warping length with respect to each gait cycle from minute 1. The least DTW distance and warping length were chosen for each gait cycle from minute 2. Then, the corresponding medians across all the chosen DTW distances and warping lengths are computed respectively and considered as “DTW score” and “Warp Score” related to minute 2. The same approach was repeated for rest of the minutes. Following this procedure, the DTW score and Warp score for minutes 2 to 6 were computed with respect to first minute for every 6MW.

The gait features (DTW score and Warp score), the walking speed and the clinical scores of every 6MW were used for analysis in the next chapter to determine the physiological significance of the gait features.

3.3 Case Study 2

A number of linear and non-linear features* like mean, standard deviation, skew, energy, etc. were extracted from the impacts' data. Figure 9 shows the boxplots for a few of the features and there is high overlap between both classes (true hits/clacks).

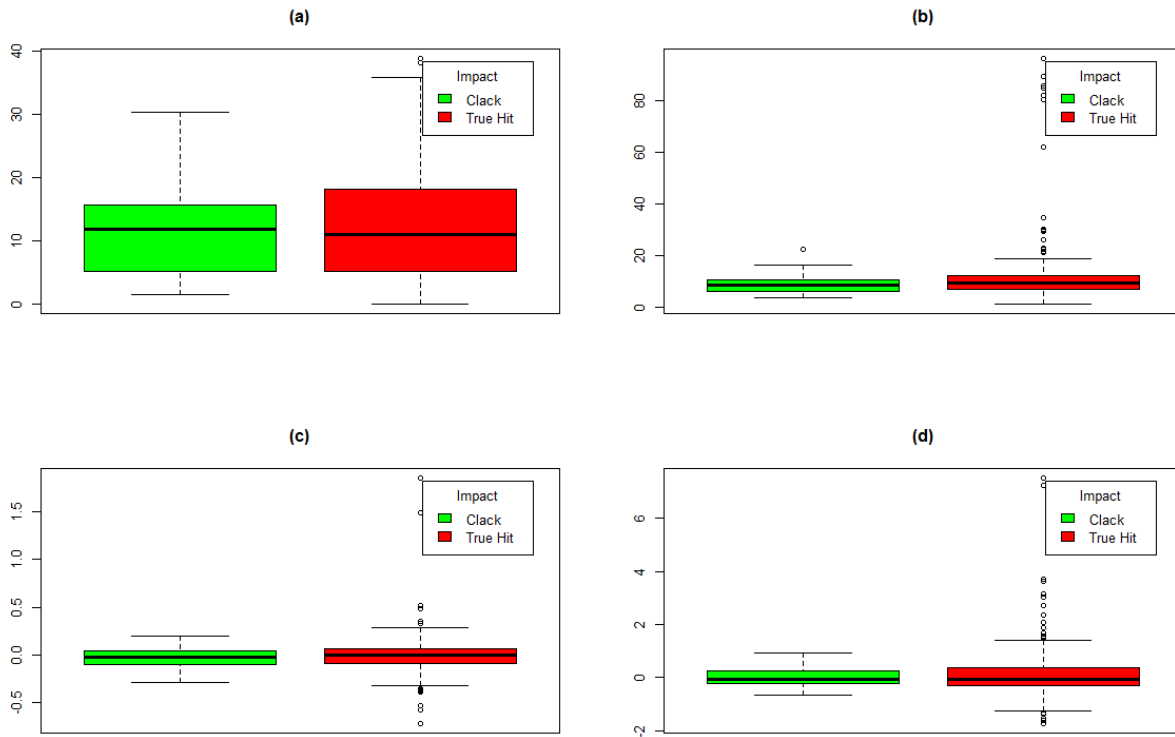


Figure 9: Boxplots of temporal features. (a) and (b) correspond to minimum and kurtosis of the acceleration, whereas (c) and (d) correspond to mean and skew of the gradient of acceleration.

The boxplots of the software computed PLA, PRA and PLV can be viewed in Figure 10. There is very little separation between the classes.

* Refer to Table 6

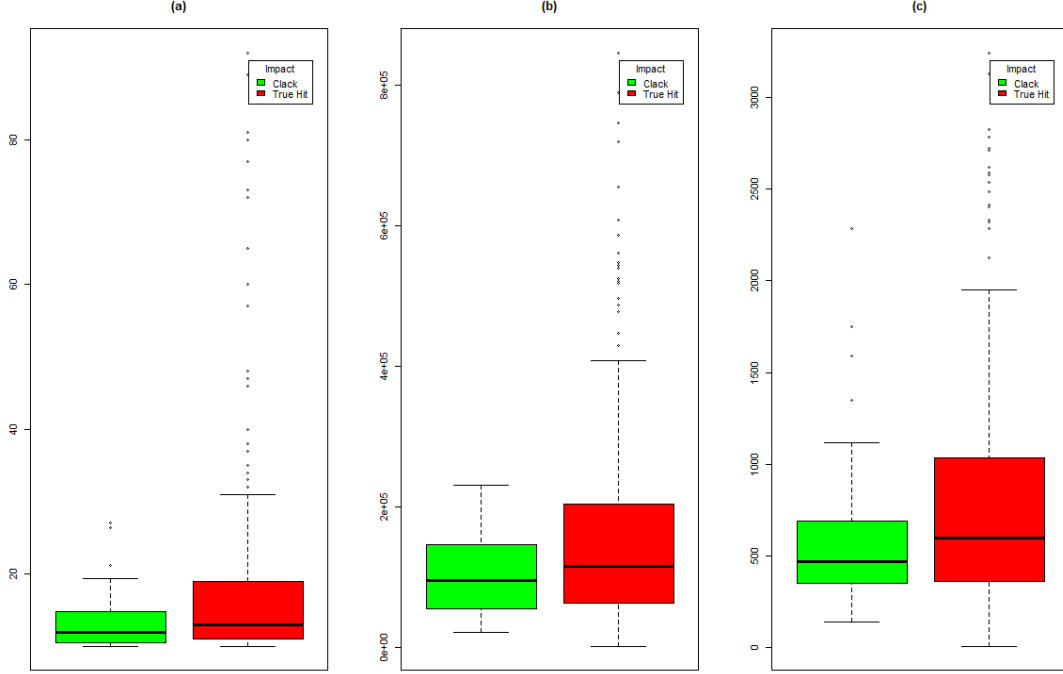


Figure 10: Boxplots of (a) PLA, (b) PRA and (c) PLV of the impacts

There was no distinction between the true hits/clacks using temporal features such as PLA, PRA, PLV, minimum and kurtosis of the acceleration, and mean and skew of the gradient of acceleration. One of the reasons for the lack of separation between the classes might be due to the dynamic environment on the field - the players' heads move rigorously while running, jumping, attacking, throwing and defending. Though the true hits and clacks occur due to two different motives, the inertial data corresponding to the both of them possess similar characteristics. So it can be assumed that the human factors play a major role, in converting the driving forces into sensor recordings. Thus, the driving forces/stimuli* would be the key to differentiate the true hits from clacks.

Models like Hidden Markov Models (HMMs) [52] can be used to estimate the stimuli of the impacts but these models experience an exponential increase in parameters as more signal history is encoded. Thus dynamical models were used to estimate the stimuli. In this work, human motion analysis was viewed as a

* A stimulus is the event that causes the head motion.

blind system identification where head motion is an unknown dynamical model driven by an unknown input. Intuitively, the dynamical model represents physical characteristics of an actor, such as mass and inertia, whereas the input represents the driving signal, a signature of the action.

The dynamical model used for estimation of stimuli from the inertial data has been inspired from Raptis' work [53]. The model, proposed in [53], falls into the class of linear dynamical models, where the task of motion modeling has been posed as a system identification problem under bounded energy and sparsity constraints [54]. The motivations to employ this model are two-fold. Firstly, the inertial data observed from human actions is non-stationary and is multivariate time series. This model accommodates such non-stationary and multivariate time series data. Secondly, the core hypothesis of this model is that multivariate time series $z(t)$ are outputs of a linear time invariant dynamical system driven by a one dimensional sparse and bounded input $u(t)$. But the human actions are non-linear and cannot be modelled by linear equations. Since the impacts recorded in our data were of just 0.1 seconds duration, the human actions were assumed to be linear in such small durations and abide the hypothesis of the model [55]. Thus, $u(t)$ is considered as the stimulus for the impact $z(t)$.

Generally, when the sensor records an impact with high acceleration values for a very few samples it corresponds to a direct impact to the head. Figure 7, shows an example of a true hit and clack. A good way to take advantage of these high acceleration values over short intervals is by computing the gradient of the signal.

So deriving the jerk signal from the raw inertial signal will amplify the sudden movements of the skull across all the directions. Thus, the jerk signal $y(t)$ is considered as the output of the linear dynamical model, which is driven by $u(t)$. The jerk signal is computed by taking the gradient of the raw accelerometer signal $z(t)$, which is shown below:

$$y(t) = \nabla_{x,y,z} z(t)$$

where $\nabla_{x,y,z}$ stands for first order derivative of each of the axes x , y and z . Figure 11 displays the jerk signal of the accelerometer data shown in Figure 7.

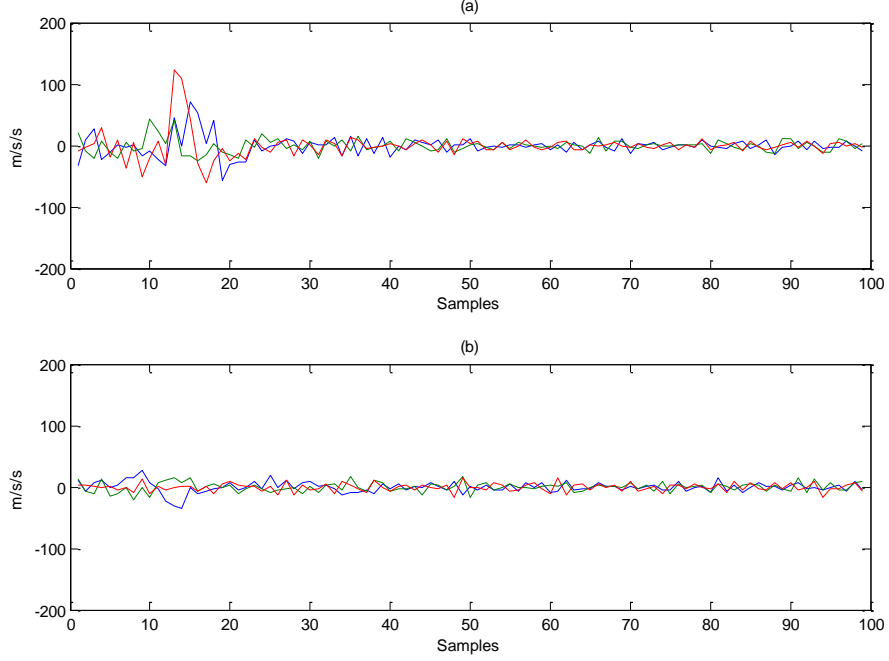


Figure 11: Gradient of the accelerometer signal shown Figure 7

The linear dynamical model was defined by the system matrices A, B, C and a state vector $x(t)$. The model can be expressed as follows:

$$\left. \begin{aligned} \|u(\cdot)\|_{l_0} &\leq k \\ |u(t)| &< 1 \\ \Sigma_t \|CA^t B\|_1 &\leq \mu \\ x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \right\} \quad (3)$$

where $\|u\|_{l_0}$ is the number of nonzero elements in the sequence u .

One aspect of the model that plays an important role is bounding the input signal. This forces variables such as the amplitude of an action into the system matrices, thus resulting in inputs u that are more comparable across individuals. An expectation – maximization (EM) algorithm [53] was applied to estimate the A, B, C , and x matrices under the assumption of system linearity and time invariance. Initially, u was randomly chosen abiding to the conditions in (3) and then A, B, C , and x were estimated. Then the sequence u was updated using the estimated A, B, C , and x matrices. This process iteratively continued until error between the estimated y and the actual y was reduced. Once the error converged, the input sequence u was obtained, which is the stimulus for the reaction z . Figure 12 displays an example of the estimated stimulus for an impact. The stimuli for all the verified impacts were computed using this linear dynamical model* approach.

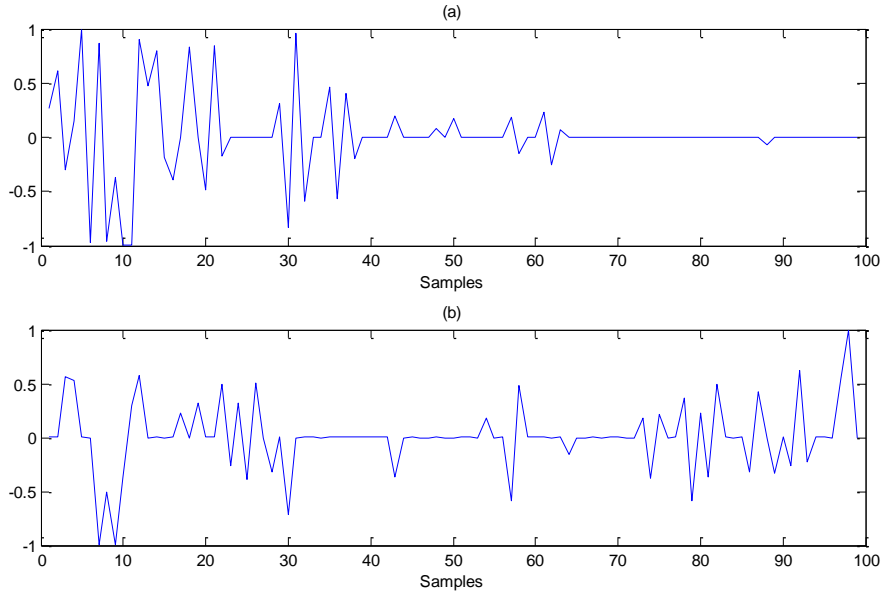


Figure 12: The $u(t)$ of the linear dynamical model where (a) and (b) correspond to a true hit and clack respectively.

In the next chapter, a classifier was trained to differentiate the true hits from clacks using the estimated stimuli. The performance of the estimated stimuli was compared with several other features.

* The linear dynamical model was built in MATLAB® R2013b (Intel i7-4770 CPU, 8GB RAM). The average computation time for each stimulus was 6.91 seconds. The algorithm is computationally expensive, so it will require additional constraints to implement on a smaller processor for real-time applications.

Chapter 4

Information Extraction

4.1 Background

Information, in simple words, is the answer to a question. Information extraction is the process of finding factual evidence, by identifying instances and relationships between entities, to solve a problem. This step in SLF focuses on modeling and knowledge discovery for predictive/classification purposes using the features derived from the earlier step. Information extraction exploits machine learning techniques and statistical analysis to build models that learn from and make predictions on the data. Machine learning techniques exploits algorithms that iteratively learn from data and allows computers to find hidden insights without being explicitly programmed where to look. These techniques are used for classification, regression, clustering, dimensionality reduction and density estimation.

Previously, several machine learning and statistical measures were exploited to understand relationships between human motion and inertial features. In [11], [41] and [56] authors employed neural networks to identify human actions like walking, standing, running, sitting, walking upstairs and downstairs. Jamsa et al [7] used pearson's correlation coefficients to study the association between acceleration peaks and bone mineral density changes. Wang et al [49] exploited multi-layer perceptron neural networks to differentiate walking patterns. Maurer et al [40] adopted online nearest neighborhood and linear discriminant analysis to differentiate human activities like walking, standing, running, sitting, ascending and descending. Kern et al [14] employed Bayes classifier to recognize activities like walking, sitting, standing, typing and writing. Herren et al [42] exploited multiple regression analysis to predict the incline and the speed of the runner. Parkka et al [57] utilized decision trees to identify rowing, cycling and more.

The type of information to be derived depends on the application itself. For my work, the extracted information will have to be actionable and useful for the medical experts. This way, the information acts as a tool for the clinicians for assessing and tracking the health of subjects.

In this chapter, for case study 1, the correlation between DTW based scores and gait pathology was determined. For case study 2, a classifier was modelled to differentiate the true hits from the clacks using the estimated stimuli and was compared with other features.

4.2 Case Study 1

In this section, a model was developed to find the relation between the gait features, walking speed and the clinical scores. In order to show that the gait features are closely related with the clinical scores (MFIS, EDSS, MSWS) and significant than walking speed, three sets of six separate models were built. These models were implemented using a least squares regression model. A linear regression model is an approach for modeling the relationship between a scalar dependent variable (Y) and one or more explanatory variables (or independent variables) denoted by X . In my case, the dependent variable (or response variable) represented one of the clinical scores and the independent variables corresponded to the derived gait features or walking speed. Each set of models represented one of the clinical scores with six different combinations of DTW score, Warp score and walking speed. Table 1 shows all the different models.

A linear regression model can be represented as follows:

$$Y = w_1X_1 + w_2X_2 + \dots + w_pX_p$$

In matrix form,

$$Y = wX$$

where w is the weights vector which has to be estimated using the Y and X matrices. Y is the column

vector with the clinical scores and X is the matrix with the features. This method was followed to build the models (F_i^n) in Table 1, where i belongs to each of the clinical scores, i.e. EDSS, MSWS and MFIS and n is the model number.

Table 1: Linear Regression Models

Model No.	Features
F_i^1	DTW Score (2,3,4,5,6)
F_i^2	Warp Score (2,3,4,5,6)
F_i^3	Walking Speed (1,2,3,4,5,6)
F_i^4	Walking Speed (1,2,3,4,5,6) + DTW Score (2,3,4,5,6)
F_i^5	Walking Speed (1,2,3,4,5,6) + Warp Score (2,3,4,5,6)
F_i^6	Walking Speed (1,2,3,4,5,6) + DTW Score (2,3,4,5,6) + Warp Score (2,3,4,5,6)

The set of weights w are chosen to minimize the sum of squared errors between the estimate of clinical score, \hat{y} , and the true clinical score, y . The ordinary least squares solution is given in (4),

$$w = (X^T X)^{-1} X^T Y \quad (4)$$

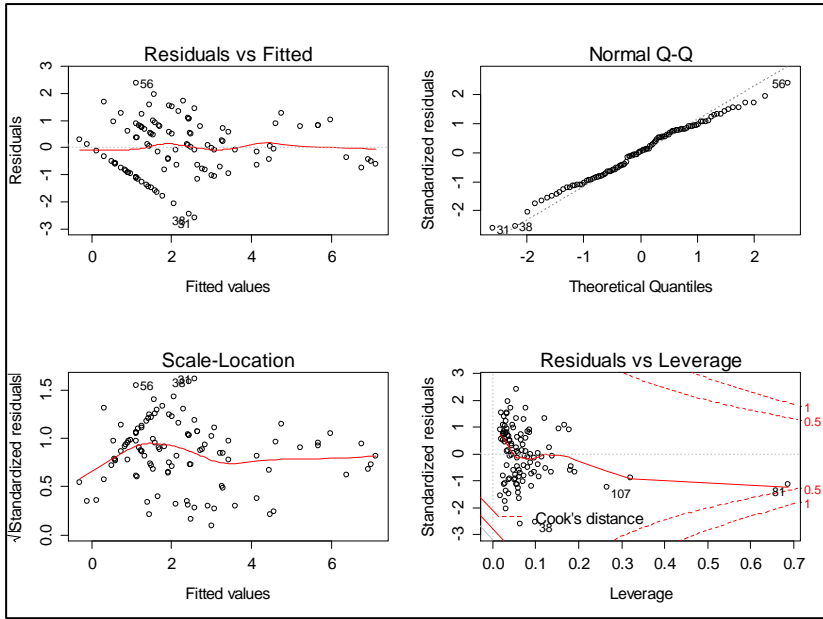
Once the weights were estimated for the all models in all the different sets, model performance was evaluated to determine the best model which explains the MS clinical scores. The following statistical performance metrics* were computed: Adjusted R squared, Akaike information criterion (AIC), Bayesian information criterion (BIC) and p significance value.

Table 2 shows these performance metrics for all the models. From Table 2, the linear regression model F^6 best fits the EDSS clinical score. Similarly, F^5 and F^4 best fit the clinical scores MSWS and MFIS respectively. The diagnostic curves† of the models F^6 , F^5 and F^4 corresponding to EDSS, MSWS and MFIS respectively are shown in Figure 13. The residuals vs fitted and the scale-location plots for all the

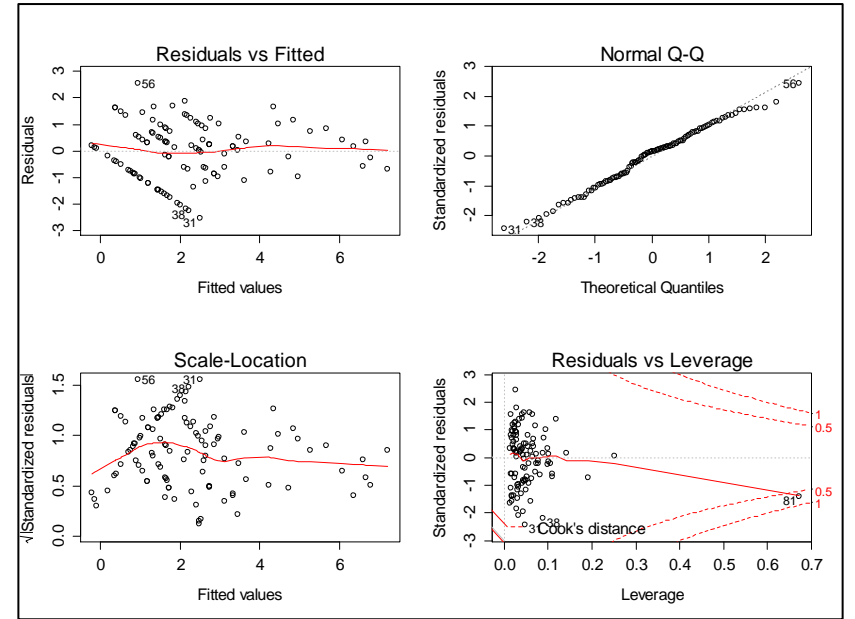
* See Appendix A for definitions

† See Appendix C for definitions

(a)



(b)



(c)

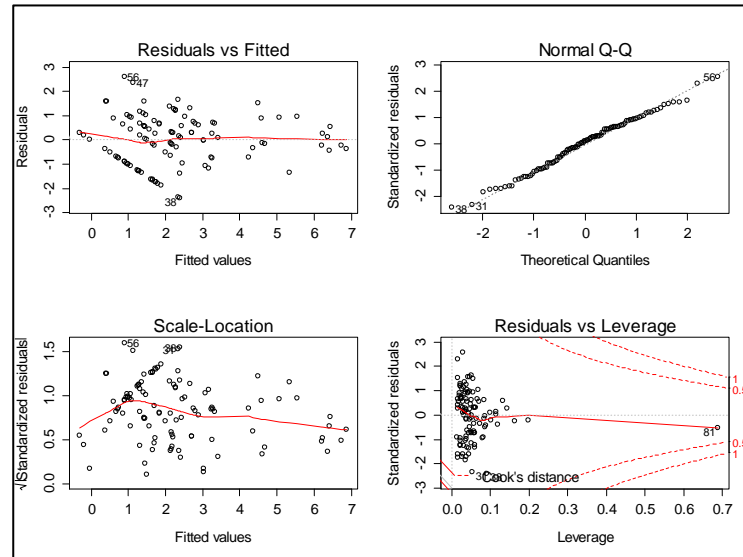


Figure 13: Diagnostic curves of the linear models (a) F_{EDSS}^6 , (b) F_{MSWS}^5 and (c) F_{MFIS}^4

three models do not possess any patterns, which indicates that models abide the linearity and homoscedasticity assumptions. In the normal Q-Q plot for all the three models, the residuals are lined well with the dotted line, which is evidence that the distribution of observations is Gaussian. The residuals vs leverage plot shows that the cook's distance scores of all the observations is less than 1, suggesting that there are no influential points which can manipulate the regression model. All the diagnostic curves provide evidence that the models fit the data well.

The predictors of the best models, that fit the clinical scores, are a combination of walking speed, DTW score and Warp score. This suggests that the walking speed along with DTW score and Warp score perform better in all the three cases, when compared with walking speed alone. Thus, DTW Score and Warp score encompass certain domains of the MS clinical scores that walking speed does not.

Now, to understand what DTW scores and Warp scores explain which walking speed does not, the direction of inference was reversed by evaluating the contribution of various clinical measures to the DTW Score, Warp Score, and walking speed. In other words, DTW score, Warp score and walking speed will act as the response variable and the clinical scores as the independent variables.

Three models were employed to analyze the physiological significance of our features. All models utilized stepwise regression, in which significant features were selected to best explain the dependent variable. Stepwise regression is an approach to select a subset of features for a regression model. It is a semi-automated process of building a model by successively adding or removing features based on the AIC performance metric of the model and the t-statistic of the feature. The focus of stepwise regression is to find the best combination of independent features to fit the dependent variable. Each of the MS clinical scores (EDSS, MFIS and MSWS) are composite scores of individual sub scores [26]–[28]. All these sub-scores of the MS clinical scores (EDSS, MFIS and MSWS) were the features to all the three models, whereas, the two gait features and the walking speed were the dependent variable for each of the models respectively. The model variables are mentioned in Table 3 where t is 1, 2, 3, 4, 5, 6 minutes.

Table 2: Performance Metrics of the Linear Regression Models

	EDSS				MSWS				MFIS			
	Adj. R^2	AIC	BIC	p	Adj. R^2	AIC	BIC	p	Adj. R^2	AIC	BIC	p
F^1	0.23	432.5	451.5	<0.001	0.24	885.78	904.75	<0.001	0.11	965.94	984.78	<0.01
F^2	0.42	401.4	420.4	<0.001	0.51	837.88	856.85	<0.001	0.20	954.87	973.71	<0.001
F^3	0.66	343.1	364.9	<0.001	0.71	783.11	804.86	<0.001	0.38	932.77	953.77	<0.001
F^4	0.68	326.8	345.6	<0.001	0.73	747.92	761.33	<0.001	0.43	890.05	906.03	<0.001
F^5	0.70	322.4	341.2	<0.001	0.77	731.76	745.17	<0.001	0.41	891.70	905.02	<0.001
F^6	0.71	320.1	344.1	<0.001	0.77	731.73	753.28	<0.001	0.43	890.07	912.01	<0.001

Table 3: Stepwise Regression Models

Response Variable	Features
<i>DTW Score_t</i>	EDSS + MFIS + MSWS
<i>Warp Score_t</i>	EDSS + MFIS + MSWS
<i>Walking Speed_t</i>	EDSS + MFIS + MSWS

The technique shown in the flowchart in Figure 14 was exploited to find the clinical scores that are closely related to the gait features and the walking speed. The DTW score, Warp Score and walking speed for every minute was considered as a dependent variable for the stepwise regression model. The reason for considering each minute separately is because previously in [24], it was shown that the pattern of distance covered through each minute in 6MW was related to the severity in MS. So, the clinical scores that are considered significant for a minute might not be the same for a different minute. For this reason, the composite value of the clinical scores was not considered. The “Z” in the flow chart corresponds to one of the following features, DTW Score, Warp Score and Walking speed. Since the DTW algorithm of the 6 minutes was computed with respect to the first minute, the Warp Score and the DTW Score for minute one do not exist. Once the significant clinical scores for each minute were obtained from the model, the most frequently occurring clinical scores* over all the six minutes were selected.

* The statistical measures of the clinical scores are shown in Appendix D.

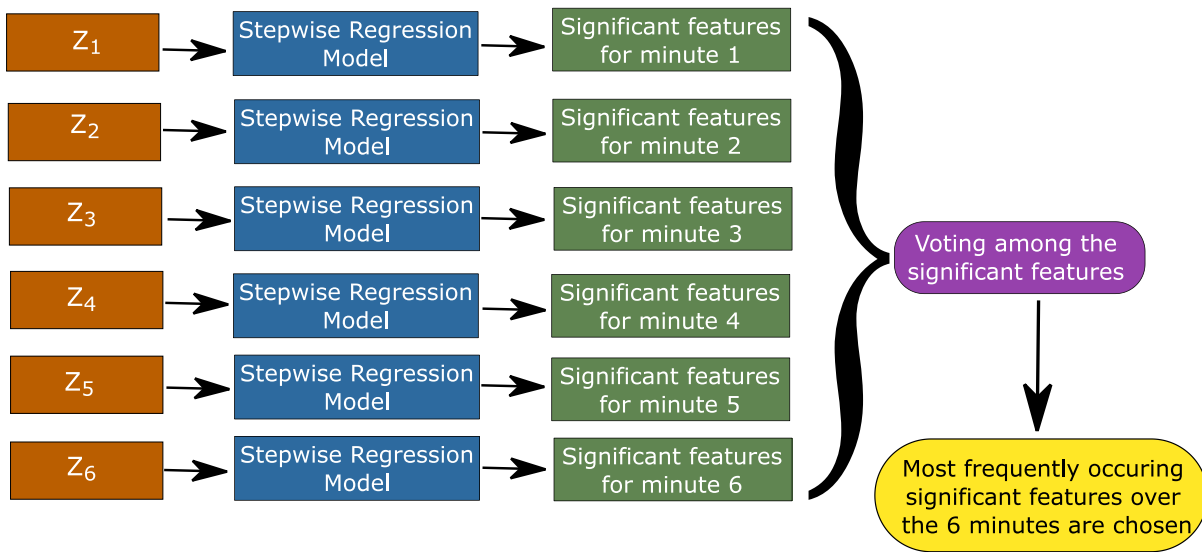


Figure 14: Stepwise Regression Flowchart

Table 4 shows the significant clinical scores that best fit each response variable. The questionnaires and functional tests, that were conducted, are related to cognitive thinking, physical fatigue, balance, coordination, and other symptoms. Gait pathology can be related to muscle weakness, stiffness, and spasticity. For simplicity, all questions and functional tests have been segregated into 3 categories, namely, “Balance”, “Physical Fatigue” and “Others”. The category “Balance” relates to stability while walking, whereas “Physical Fatigue” relates to declining walking performance brought on by exertion. Questions relating to cognitive fatigue were included in the “Others” category. The last column in Table 4 specifies the category corresponding to each clinical score.

Table 4: Clinical Scores Correlated with Gait Features

Gait Features	Clinical Scores		Category
Warp Score	MSWS	Limited your balance when standing or walking?	Balance
	MSWS	Made it necessary for you to use support when walking outdoors?	Balance
	EDSS	Gait ataxia, limb ataxia, tandem walking, etc.	Balance
	EDSS	Urinary incontinence, bowel dysfunction, etc.	Others
	MFIS	I have had difficulty paying attention for long periods of time	Others
	MFIS	I have been less motivated to do anything that requires physical effort	Physical Fatigue
	MFIS	I have had difficulty making decisions	Others
DTW Score	MSWS	Limited your ability to run?	Physical Fatigue
	MFIS	I have been less motivated to do anything that requires physical effort	Physical Fatigue
	MFIS	I have had trouble maintaining physical effort for long periods.	Physical Fatigue
	MFIS	I have had difficulty making decisions.	Others
	MFIS	My muscles have felt weak.	Physical Fatigue
	MFIS	I have needed to rest more often or for longer periods.	Physical Fatigue
Walking Speed	MSWS	Made it necessary for you to use support when walking outdoors	Balance
	MSWS	Affected how smoothly you walk?	Balance
	EDSS	Urinary incontinence, bowel dysfunction, etc.	Others
	EDSS	Hearing loss, dysphagia, etc.	Others
	MFIS	I have been unable to think clearly.	Others
	MFIS	I have been forgetful.	Others
	MFIS	I have been limited in my ability to do things away from home	Others
	MFIS	My muscles have felt weak.	Physical Fatigue
	MFIS	My thinking has been slowed down.	Others
	MFIS	I have needed to rest more often or for longer periods.	Physical Fatigue

Figure 15 shows the contributions of the three categories. This shows that walking speed has the highest contribution from the category “Others”, including several symptoms of cognitive fatigue, whereas the DTW Score and Warp Score have the highest contribution from “Physical Fatigue” and “Balance” respectively. This shows that the walking speed relates to many diverse clinical variables, whereas the DTW Score and Warp Score are primarily a function of gait impairment.

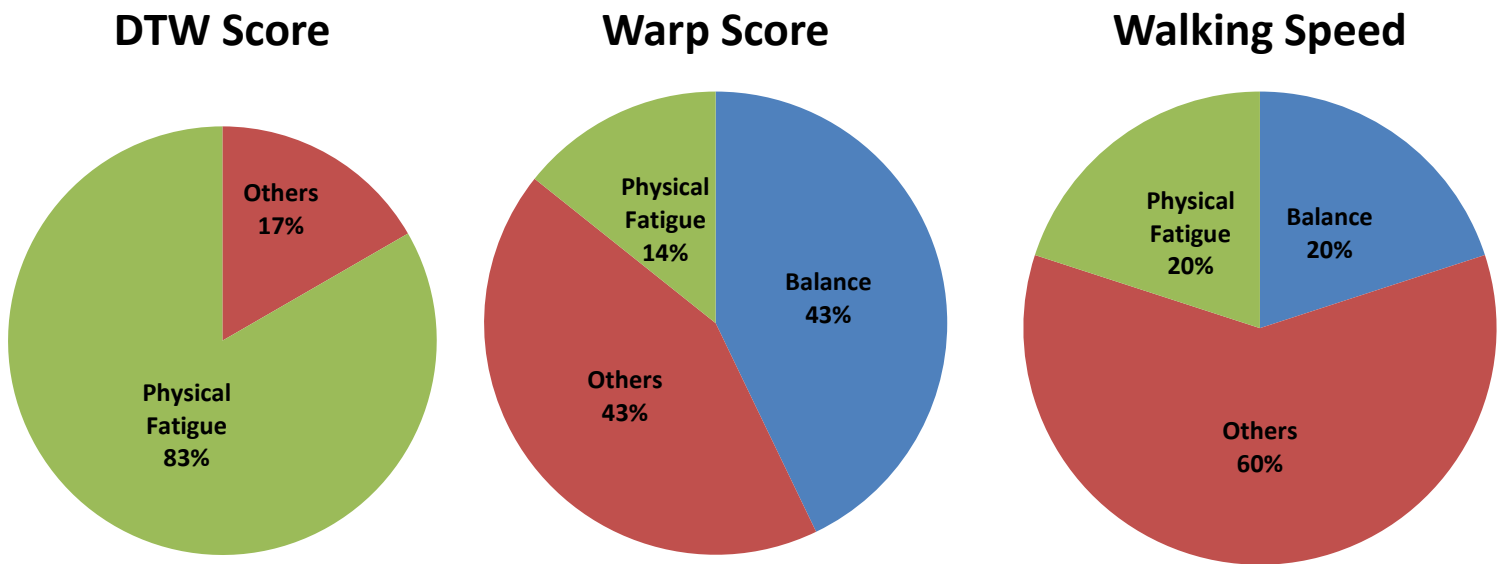


Figure 15: Contributions of the clinical categories to the DTW Score, Warp Score and Walking Speed

4.3 Case Study 2

A classifier was trained to identify the true hits from the clacks. In machine learning, classification is the problem of identifying to which category a new observation belongs, on the basis of a training set of data containing observations (or instances) whose categories are known. Classification is considered an instance of supervised learning, i.e. learning from a set of correctly identified observations and labelling the unidentified observations. Two types of classifiers were implemented:

➤ Probabilistic Classifier

This type of classifiers learns the probability distribution over the set of classes and uses that to make predictions.

➤ Non-Probabilistic Classifier

This type of classifiers does not attempt to model the underlying probability distributions, but tries to build the decision curve itself to make predictions.

One probabilistic classifier and two non-probabilistic classifiers were implemented, which can be found in Table 5.

Table 5: Classifiers*

Probabilistic Classifier	Quadratic Discriminant Analysis
Non- Probabilistic Classifier	Support Vector Machine
	Logistic Regression

The performance of the classifiers was validated using a variant of exhaustive cross validation called leave-one-out cross-validation (LOOCV). This cross validation technique uses one observation as the testing set and the remaining observations as the training set. This is repeated until all the observations have been used as the testing set. This validation technique was chosen due to the small dataset available. It uses the available data more efficiently when compared to other cross validation methods, as only one observation is omitted at each step. All the above classifiers were trained using LOOCV with four different sets of features which are mentioned below:

❖ Set A: Magnitude

- The magnitude of the raw inertial data obtained from the sensor

$$Magnitude = \sqrt{x^2 + y^2 + z^2}$$

* All the classifiers were built in RStudio. Packages ‘MASS’, ‘e1071’ and ‘nnet’ were used to implement QDA, SVM and logistic regression respectively.

- 100 available features for each observation
- ❖ Set B: Jerk Signal
 - The magnitude of the gradient of the raw inertial data from the sensor
 - 99 features available for each observation
- ❖ Set C: Temporal Features
 - 18 different temporal features are computed from the magnitude of the jerk signal and original accelerometer data:

Table 6: Temporal features

Mean {.}	Variance {.}	Max {.}
Min {.}	Skew {.}	Kurtosis {.}
Energy {.}	Variance [.]	Max [.]
Min [.]	Kurtosis [.]	Skew [.]
Mean [.]	Energy [.]	Mahalanobis Distance* [.]
PLA [†]	PRA [†]	PLV [†]

$\alpha\{.\}$ refers to computing the feature “ α ” from the magnitude of the original accelerometer data

$\alpha[.]$ refers to computing the feature “ α ” from the magnitude of the jerk signal

- ❖ Set D: Stimuli
 - The u which is the estimated stimuli using LDM.
 - 99 features available for each observation

Initially, a simple linear regression model was employed to study the significance of the four feature sets. The labels of the impacts were used as the response variable of the model. Table 7 shows the performance metrics for all the feature sets. The original feature set D has the highest R squared and adjusted R squared values, whereas the feature set C has the least AIC and BIC values. This infers that the feature set D encompasses the variance of the dataset but due to the large number of features it tends to have higher

* Mahalanobis distance is the ratio of the difference between the max and mean values to the standard deviation of the signal.

[†] Refer to section 1.2

AIC and BIC value. Since the feature set C has fewer features compared to set D, it has a lesser AIC and BIC metric. This suggests that there might be a number of redundant features or features that are correlated with each other in the feature sets A, B and D. Ideally, one feature is required to differentiate two classes. But feature set A, B and D have approximately 100 features each for classification.

To eliminate the insignificant features, stepwise regression model was chosen. A stepwise regression model chooses the significant features by computing the AIC of the model and the t statistic of the features. After using the stepwise regression model, the feature set A, B and D contain 35, 31 and 28 number of features respectively. Table 7 shows the performance metrics of the linear regression model with the original set and selected set of features.

Table 7: Performance metrics of the original set and selected set of features from sets A, B and D

	R^2		Adj. R^2		AIC		BIC	
	Original features	Selected features	Original features	Selected features	Original features	Selected features	Original features	Selected features
Set A	0.37	0.30	~ 0	0.19	281	180	641	310
Set B	0.37	0.28	~ 0	0.18	279	177	635	294
Set C	0.04	-	~ 0	-	226	-	297	-
Set D	0.45	0.35	0.11	0.27	244	145	601	251

The R^2 metric is higher for the original set of features because of the higher number of features when compared to the selected features. But the adjusted R^2 metric is higher for the selected set of features because this metric penalizes the model for having redundant features. The AIC and BIC values are lower for the selected features as well. This suggests that the reduced number of features contain less noise. Thus, these reduced set of features were used to train the classifiers.

Moreover, the feature set D has the best performance metric compared to the other feature sets, which suggests that it may be a good predictor of true hits.

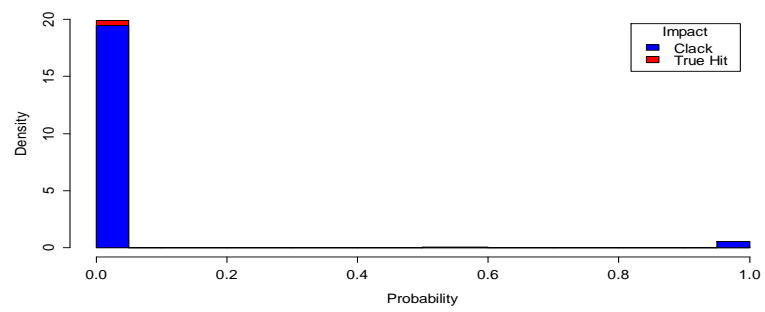
Quadratic Discriminant Analysis

Quadratic Discriminant Analysis (QDA) is a general discriminant function with a quadratic decision boundary which can be used to classify datasets with two or more classes. QDA assumes that the distribution of classes to be Gaussian. A Discriminant analysis technique models the underlying distributions of the classes and assigns the new observation to the class with least the Mahalanobis distance. QDA exploits the variance amongst the classes along with the Mahalanobis distance to classify the test observations.

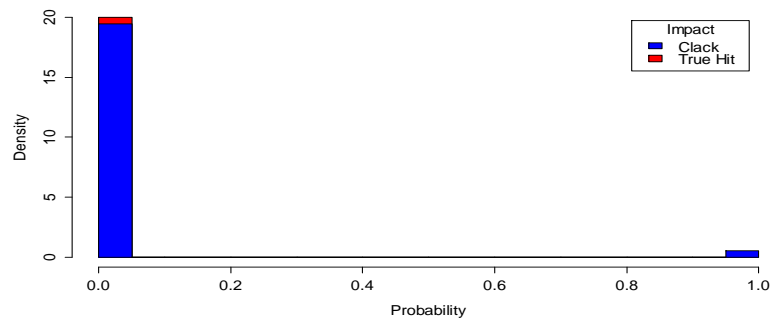
Previously, Liu et al [58] used QDA in their algorithm to detect backward falls, prior to the impact, in elderly population. Palmerini et al [59] tested QDA to analyze postural instabilities in subjects suffering from Parkinson's disease using inertial data. Rebersek et al [60] adopted QDA to detect the onset of gait initiation, the first heel-off and the first toe-off. Ganea et al [61] utilized QDA to investigate the alteration of the gait pattern in children with Duchenne muscular dystrophy, using body-worn inertial sensors.

QDA, with the four feature sets, was implemented on the labelled impacts and the posterior probability for each of the test instances was calculated. Figure 16 shows the distribution of the posterior probabilities all the impacts.

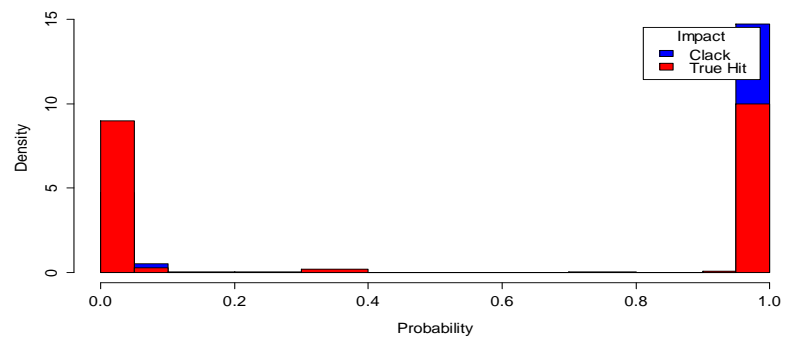
Set A



Set B



Set C



Set D

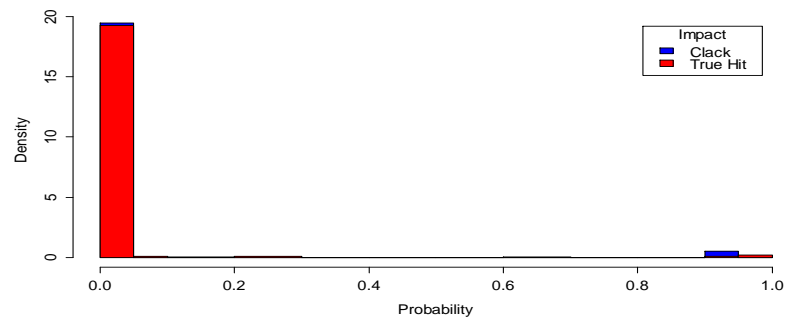


Figure 16: Density plot of posterior probabilities computed using QDA

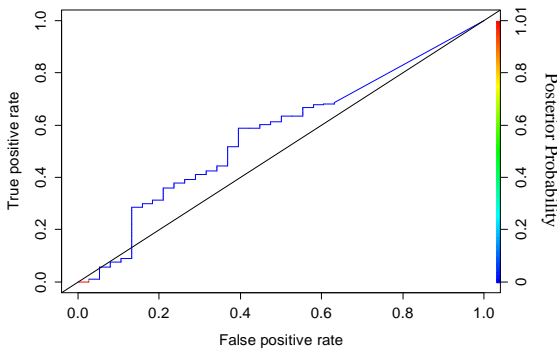
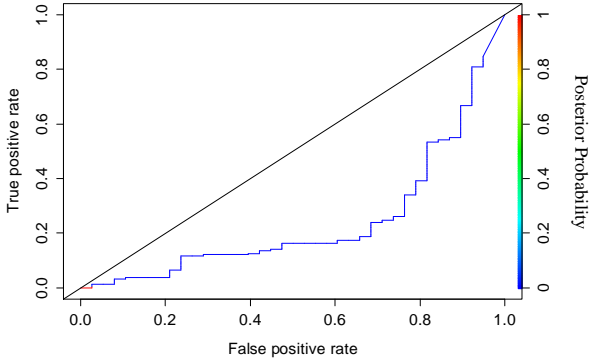
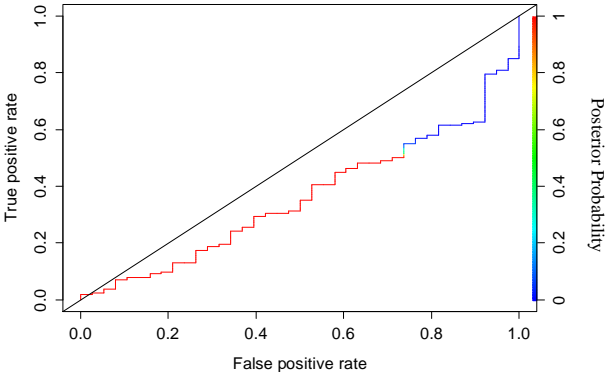
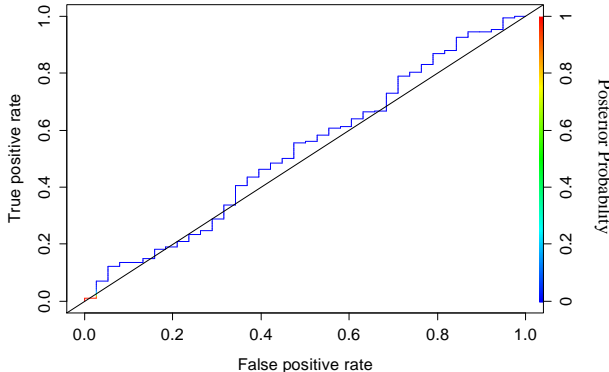
The posterior probabilities computed using QDA for set A, B and D lie near 0. For feature set C, the probabilities look spread out but a distinct separation between true hits and clacks cannot be identified with the naked eye. In order to compare the performance of the feature sets with QDA, ROC* curves were examined along with the area under the curves (AUROC*), which are shown in Table 8.

The AUROC values for all the feature sets are lesser or nearly 0.5, which shows that these features with QDA are no better than random guessing. Also, the ROC curve is completely blue in color for feature sets A, B and D which shows that all the posterior probabilities for true hits and clacks lie near 0.

Overall the QDA performs poorly; it might be due to the overlap between the distributions of the two classes. For this reason, the non-probabilistic classifiers were employed to avoid the modelling of distributions of the classes.

* See Appendix B for definitions

Table 8: ROC and AUROC v alues of QDA for feature set A, B, C and D

Feature Set	Set A	Set B
ROC		
AUROC	0.56	0.24
Feature Set	Set C	Set D
ROC		
AUROC	0.36	0.53

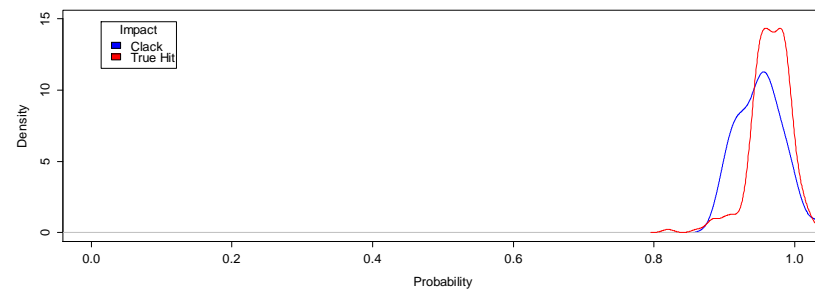
SVM

Support Vector Machine (SVM) performs classification by finding the hyperplane that maximizes the margin between the two classes. The vectors (cases) that define the hyperplane are the support vectors. An SVM model maps the points in space so that the points of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on. To assign probabilities for the test instances Platt Scaling [62] was integrated with SVM. Platt's probabilistic outputs for Support Vector Machines have been popular for applications that require posterior class probabilities. The algorithm approximates the posterior probability using a sigmoid function.

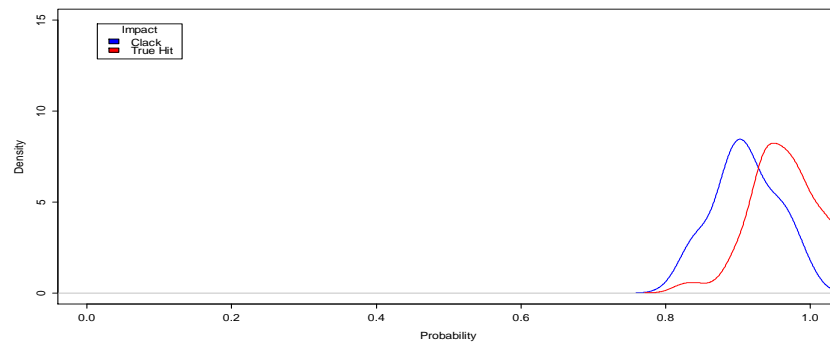
Previously, [63]–[65] adopted SVM to identify human activities using inertial data. Patel et al [66] implemented SVM to estimate the severity of tremor, bradykinesia and dyskinesia using accelerometer data. Kumar et al [67] incorporated SVM into their technique to predict the National Institute of Health Stroke Scale (NIHSS) stroke index of post-acute stroke patients using inertial data. Shibuya et al [68] adopted SVM to build a real time fall detection system.

The SVM algorithm was trained and tested on the labelled dataset along with Platt scaling to obtain the posterior probabilities for all the impacts. Figure 17 shows the distribution of the posterior probabilities of the true hits and the clacks.

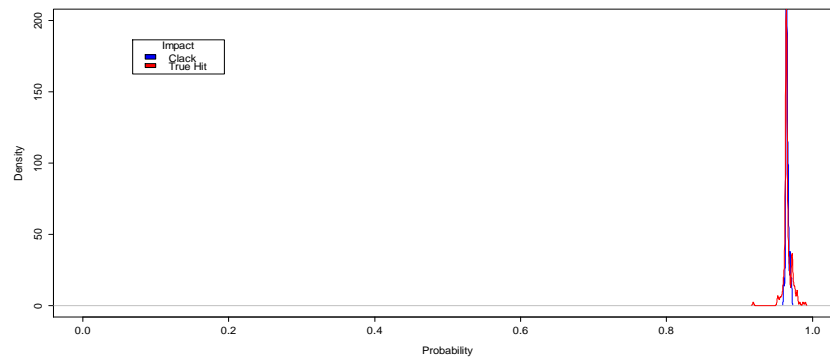
Set A



Set B



Set C



Set D

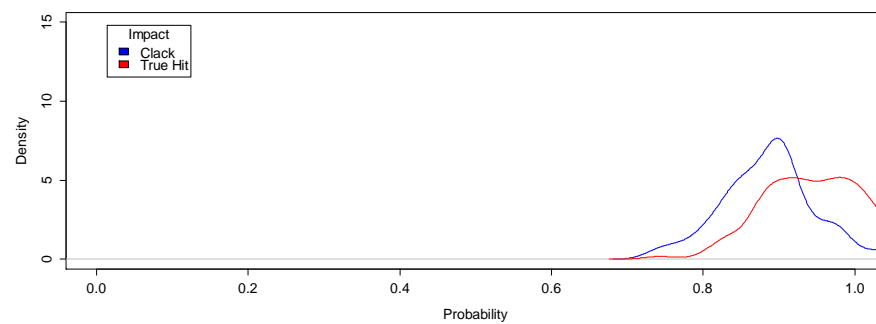


Figure 17: Density plots of the posterior probabilities computed using SVM

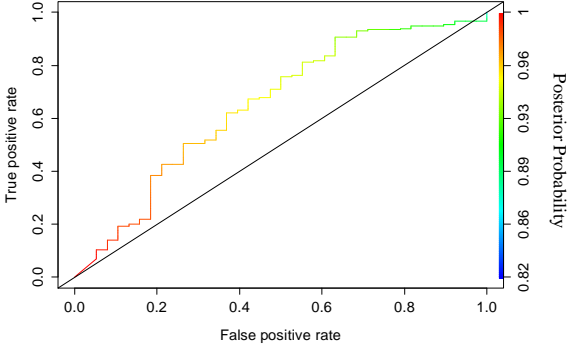
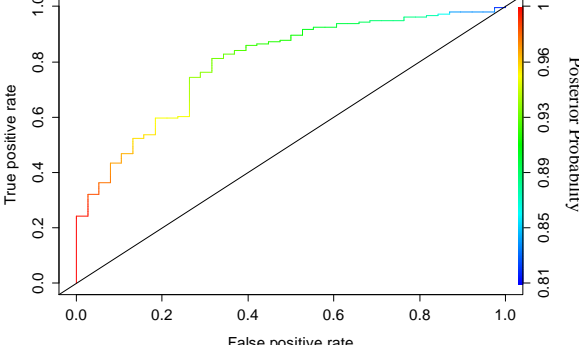
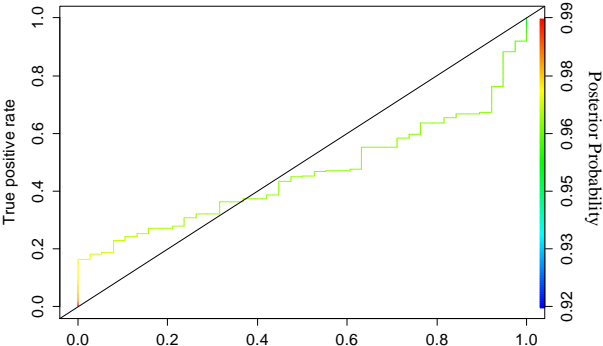
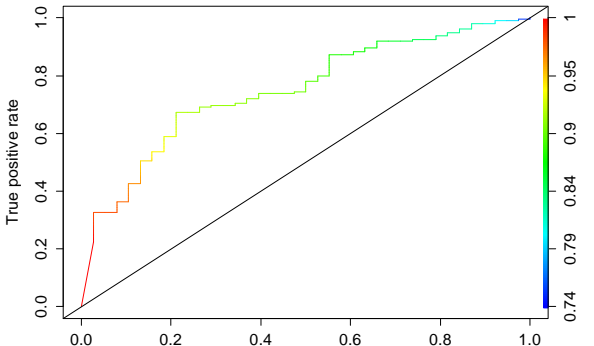
For set C, the density plot is a sharp spike because the posterior probability values for all the test impacts are in the range of 0.96 – 0.97. There is little or no separation between the true hits and clacks using the four feature sets, as shown in Figure 17. For this reason, the best feature set and the optimal threshold cannot be chosen by looking at the plots with the naked eye. In order to handle this issue, ROC* curves were examined along with the area under the curves (AUROC*).

The ROC curves for all the feature sets and the AUROC metrics are shown in the Table 9. The AUROC value for feature set C is less than 0.5, which means that SVM classifier with this feature set was not good than random guessing. This was no surprise because the probability distributions of the true hits and the clacks were overlapping, as shown in Figure 17. The SVM classifier with set B has the highest AUROC value when compared with other feature sets. Thus, SVM classifier performed best with the feature set B.

The optimal threshold to classify the impacts is decided after examining the metrics from the logistic regression classifier.

* See Appendix B for definitions

Table 9: ROC and AUROC values of SVM for feature set A, B, C and D

Feature Set	Set A	Set B
ROC		
AUROC	0.65	0.79
Feature Set	Set C	Set D
ROC		
AUROC	0.45	0.75

Logistic Regression

Logistic regression predicts the probability of an outcome that can only have two values (i.e. a dichotomy). The difference between a linear regression and logistic regression is that the former assumes a linear relationship between the response variable and features, whereas the latter assumes a linear relationship between natural logarithm of odds of the response variable and the features. Figure 18 gives a pictorial view of the difference between linear and logistic regression decision curves.

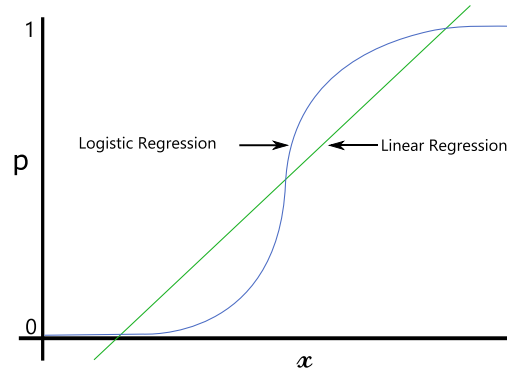


Figure 18: The linear regression and logistic regression decision curves

Since the number of classes is two, the logistic regression assumes that the instances have a Bernoulli distribution and produces a logistic curve, which is limited to values between 0 and 1. The logistic regression model can be represented as shown in (5),

$$\log\left(\frac{p}{1-p}\right) = \sum w_i x_i \quad (5)$$

where p , w and x_i refers to probability, i^{th} weights and i^{th} predictor variable.

Previously, [69], [70] implemented logistic regression to recognize human activity from inertial data. Knighten et al [71] used logistic regression to identify social gestures of humans using wrist-worn smart band. Masse et al [72] adopted logistic regression and inertial sensors to improve postural transition recognition mobility-impaired stroke patients.

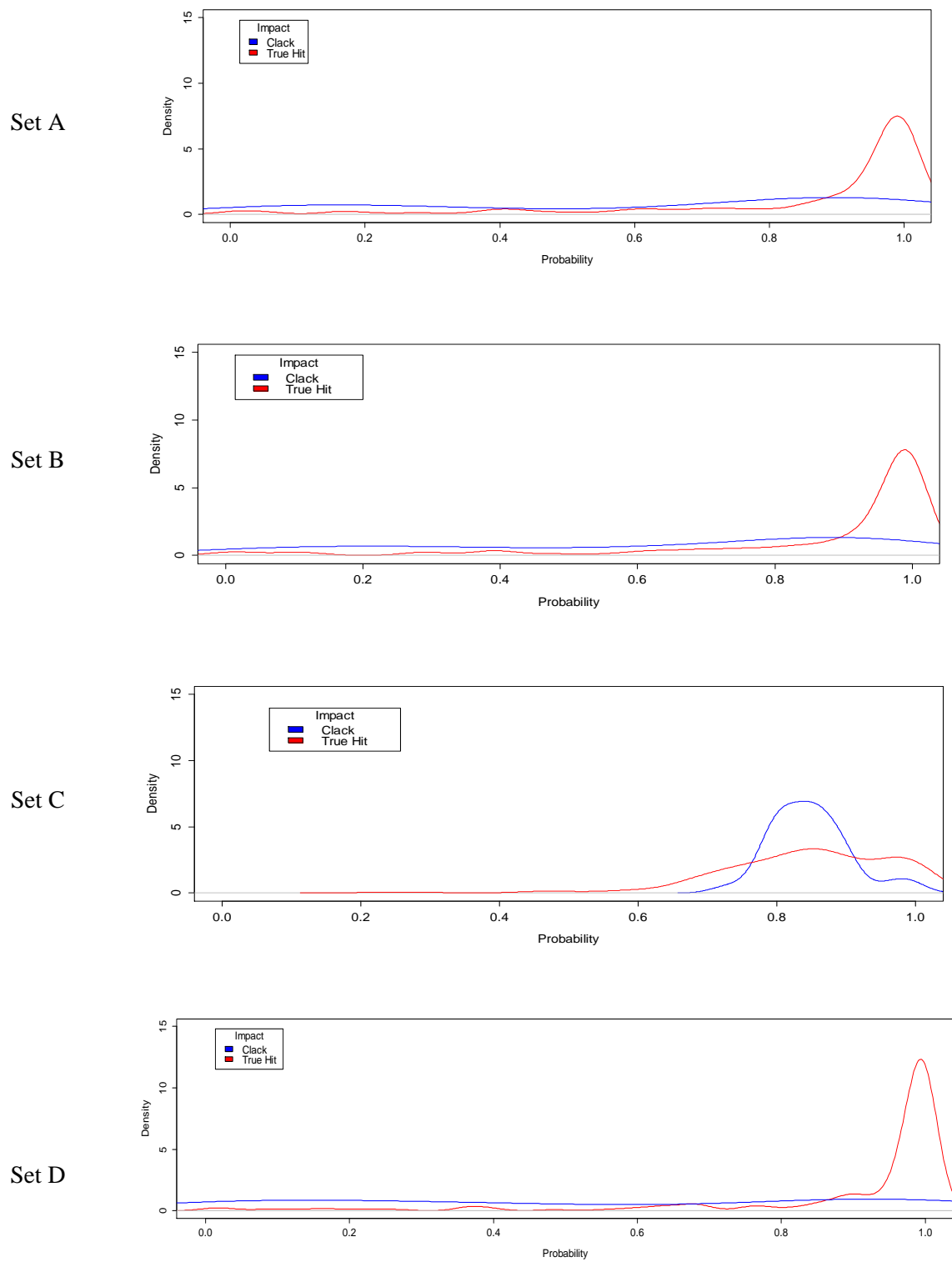


Figure 19: Density plots of the posterior probabilities computed using logistic regression

Logistic regression was tested on all the feature sets and the posterior probabilities were computed for all the impacts. Figure 19 shows the probability distribution of true hits and clacks. The overlap between the distribution of true hits and clacks was highest for feature set C, when compared to feature sets A, B and D. The feature sets A, B and D have a similar distribution, where the true hits were accumulated near 1 and the clacks were evenly spread-out. But the feature set D had more true hits near 1 compared to A and B. The posterior probability distribution of the true hits, for the feature set D, is high near 1, whereas for the clacks it is spread out evenly from 0 through 1. For deciding the best feature set and the optimal threshold, the ROC curves were examined along with the AUROC values.

The ROC curves and the AUROC metrics generated for the available feature sets using a logistic regression are shown in Table 10.

Though the AUROC of set C with logistic regression was higher compared with SVM, it was still low when matched with AUROC values from other feature sets. So the feature set C with all classifiers (i.e. QDA, SVM and Logistic regression) performs poorly. The logistic regression along with set D has the highest AUROC.

Table 10: ROC and AUROC values of logistic regression for feature set A, B, C and D

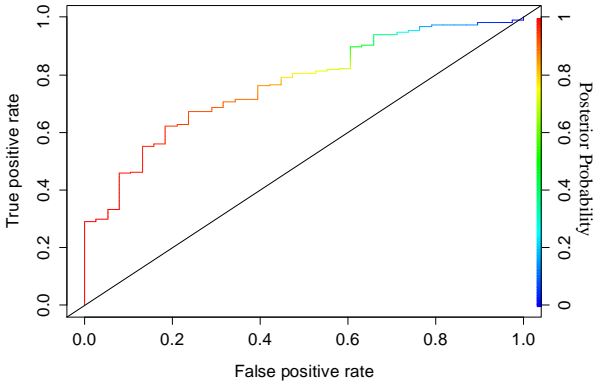
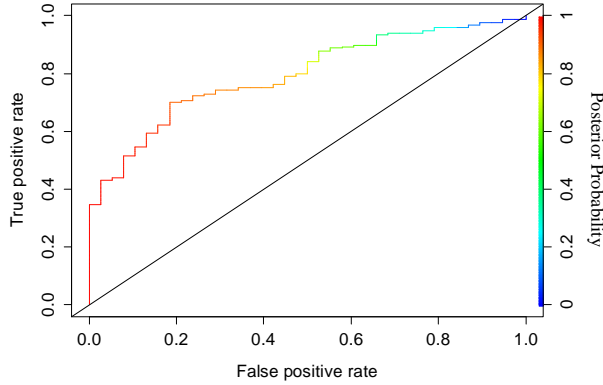
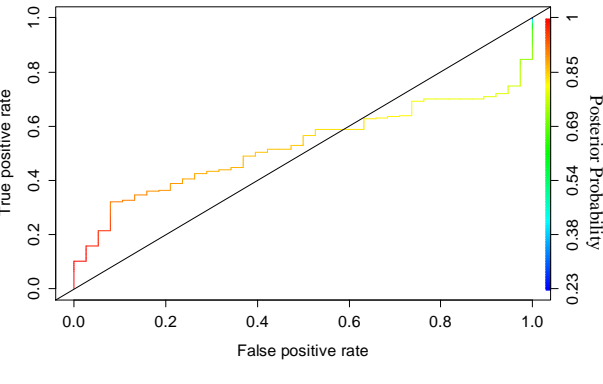
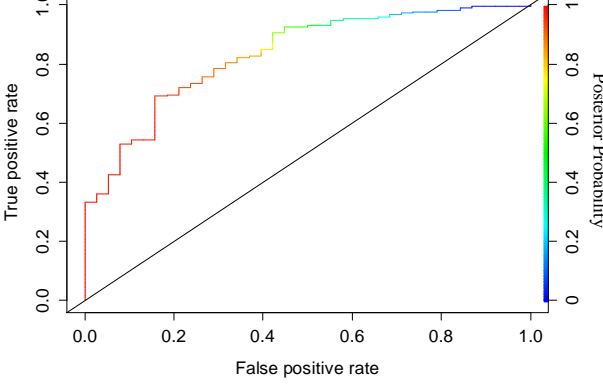
Feature Set	Set A	Set B
ROC		
AUROC	0.76	0.79
Feature Set	Set C	Set D
ROC		
AUROC	0.52	0.83

Table 11 shows the AUROC values for all the combinations between the feature sets and the two non-probabilistic classifiers. The feature set D with logistic regression is the best classifier, among all the possible combinations, with highest AUROC value.

Table 11: Comparison of the AUROC metrics

Feature Set	Classifier	AUROC
Set A	SVM	0.65
Set B	SVM	0.79
Set C	SVM	0.45
Set D	SVM	0.75
Set A	Logistic Regression	0.76
Set B	Logistic Regression	0.79
Set C	Logistic Regression	0.52
Set D	Logistic Regression	0.83

Now the optimal threshold for the logistic regression classifier with set D has to be computed to classify the true hits from clacks. There are several ways to decide the optimal threshold for a classifier; a few ways are when the threshold:

- Maximizes* $(\text{Sensitivity}^* + \text{Specificity}^*)/2$
- Has* Sensitivity = Specificity
- Maximizes* Accuracy
- Has* Sensitivity = 90 %
- Has* Specificity = 80 %

Figure 20 shows the classification metrics like sensitivity, specificity, accuracy and the average of sensitivity and specificity for varying thresholds. The thresholds were chosen between 0 and 1, and the corresponding classification metric can be viewed in Figure 20.

* See Appendix B for definitions

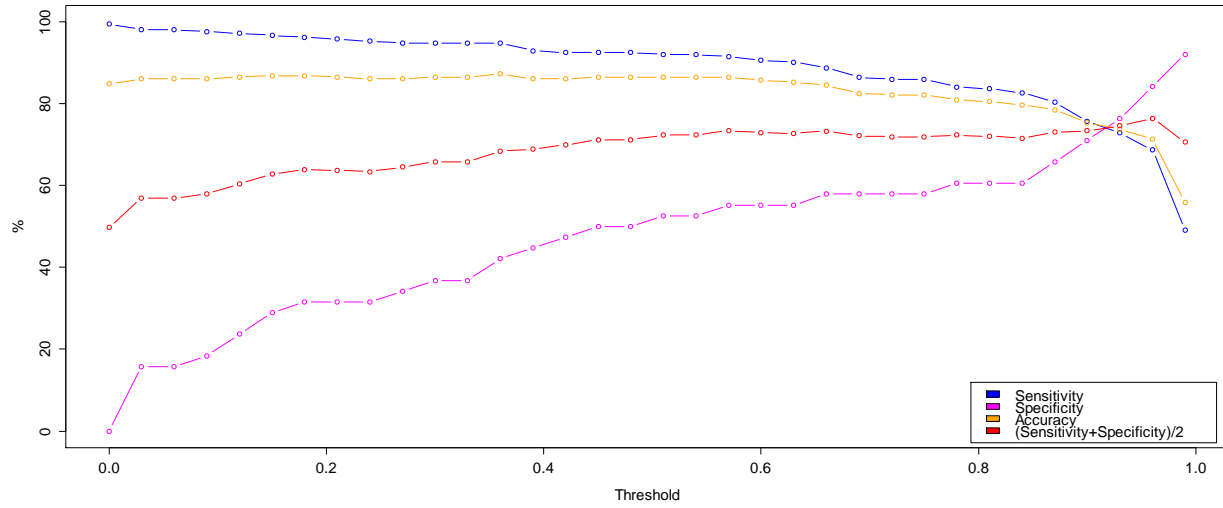


Figure 20: Classification measures with varying thresholds

The thresholds and the corresponding classification metrics for the 5 cases are shown in Table 12. Now, to choose the right threshold, the expertise of a clinician was required. As it is necessary to know what of the measures in Table 12 weigh higher in the clinical domain. After consulting a medical expert, case d was chosen as it is important not to have a large number of false negatives. In other words, it is important to identify most of the true hits at the cost of sometimes labelling clacks as hits as well. So, 0.63 was chosen as the optimal threshold to classify the true hits from clacks for the women lacrosse data; it achieves 90%, 55.26% and 85.31% as sensitivity, specificity and accuracy respectively.

Table 12: Threshold and Classification measures for the different cases

Case	Threshold	Sensitivity	Specificity	Accuracy	(Sensitivity + Specificity)/2
a.	0.96	68.69	84.21	71.42	76.45
b.	0.91	73.68	73.68	74.63	73.68
c.	0.36	94.85	42.11	87.31	68.48
d.	0.63	90	55.26	85.31	72.72
e.	0.45	70.79	80	72.61	75.39

Chapter 5

Conclusion

A sequential learning framework (SLF) integrated with application specific knowledge works towards knowledge discovery at each step, which is an essential process to derive relevant information from raw inertial data. The information generated at each stage is passed through the framework and eventually builds a concrete solution to address the problem. All the three stages defined in the SLF have their own significance, as studied in both case studies. In case study 1, firstly, the gait cycles corresponding to each minute of the 6MW were extracted and the clinical scores were tagged to the respective 6MW. Secondly, DTW based metrics were computed based on the changes in gait cycles across six minutes. Finally, using linear regression and stepwise regression models it was shown that DTW-based scores are a primary representation of gait pathology. In case study 2, firstly, the impacts were labelled as true hit or clack with the help of the video footage. Secondly, the stimuli of the impacts were estimated using a linear dynamical model. Finally, a classifier was developed using the stimuli along with logistic regression to differentiate the true hits from clacks. All the steps, employed in the framework, refined and transformed the data in their respective way to provide meaningful information, which eventually addressed the problems. The information derived from both case studies has the potential to be used as a tool to monitor and assess the health of a person. The subsections below will summarize the information derived from the case studies.

5.1 Case Study 1

Walking speed, DTW Score, and Warp Score are closely related to several patient-reported and exam-defined outcome measurements. The traditional objective measure, walking speed, has the best

correlation with overall disability as it correlates with cognitive fatigue, bladder function, and other more general symptoms of disease progression. The DTW Score and Warp Score are solely representative of the gait pathology as DTW Score is primarily a measure of physical fatigue and, the Warp Score measures balance and physical fatigue. As a result, DTW-based features are a promising tool for gait assessment in other pathologies causing physical fatigue and/or poor balance. Further, predictive models can be built to estimate the gait pathology of a subject by analyzing their respective DTW Score and Warp Score. These new features initially intended for persons with MS can now be applied to a wide variety of gait-related disorders, giving care providers a precise new tool to quantify specific aspects walking impairment.

5.2 Case Study 2

The estimated stimulus is the best performing feature for differentiating the true hits from the clacks. This feature had better classification metrics than features such as original accelerometer signal, jerk signal and the temporal features. This shows that the stimulus, estimated using a linear dynamical model, provides us with a good discrimination between impacts. The rigorous motion of the players' heads causes the true hits and the clacks to have the same temporal characteristics even though the driving motives for true hits and clacks are different. For this reason, the stimuli perform better than the temporal features. Thus, the linear dynamical model has been able to successfully estimate the driving force using the inertial data. Moreover, a logistic regression model with the estimated stimuli, as input, is the best classifier for differentiating the true hits from the clacks. This classifier achieves an overall accuracy of 85.31%, sensitivity of 90%, and specificity of 55.26% with 0.63 as a threshold. This classifier also provides a flexibility to manipulate the performance measures by changing the threshold value. If a higher specificity is required, a higher threshold can be chosen. The true hits identified by the classifier can be used to study the relation of direct hits to the head with concussions and brain injury. Furthermore, this classifier can be

extended to identify true hits in other contact sports like soccer, football, etc. It also has the potential to be incorporated into on-field screening devices for real-time identification of true hits.

5.3 Future Work

The framework presented in this work can be incorporated into a wide range of sensor-based applications for information extraction. However, there are a few aspects of this work, which can be revisited, such as building data preparation standards and evaluating the framework itself. Firstly, standards for data preparation will reduce time and effort for cleaning and labelling the raw data. This will also make sharing data easier and algorithms can be implemented across datasets seamlessly. A few precautionary steps that should be taken during data collection procedure, which will help reduce effort during data preparation, are time synchronization amongst modules, reliable calibration of the sensors, and consistency while recording values manually. Secondly, standards for framework evaluation are required because currently, the correlation between the extracted information and the clinical scores is used as a metric to evaluate the framework. However, this new information may provide a new dimensional solution to the problem which the clinical scores may not consider. Also, “gold standard” clinical scores are not available for most of the health applications. So having a standard to evaluate frameworks would be a good way to judge the credibility of the information.

The methodology built for case study 1 can be extended to other medical domains where gait pathology is a symptom such as rehabilitation, Parkinson’s disease, cerebellar ataxia and many others. The DTW based metrics can be tested in these domains to understand gait disorders. Currently MS subjects visit the hospital every six months for disability assessment. However, this methodology will enable care providers to remotely monitor gait pathology on an ongoing basis.

The approach of case study 2 can be implemented across other sports such as soccer and football, so that every true hit a player sustains during any activity is accounted. This way the relation between true hits and brain trauma injury can be analyzed. Upon understanding the relationship, predictive models can also

be built to foresee the brain health of players before games. This will also make sure that players' are able to recover from brain injuries between games.

Appendix

A. Performance Metrics

R^2

Coefficient of determination (R^2) is a number that indicates the proportion of the variance in the dependent variable that is predictable from the independent variable. Values of R^2 are generally in the range from 0 to 1, where a value closer to 1 indicates that the unexplained variance is less.

Let y be the dependent variable and \hat{y} be the estimated y by the model.

$$R^2 = 1 - \frac{\text{Residual sum of squares}}{\text{Total sum of squares}} = 1 - \frac{\sum_i (y_i - \hat{y})^2}{\sum_i (y_i - \bar{y})^2}$$

where $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$

In simple linear regression, an R^2 of 1 indicates that the regression line perfectly fits the data.

Adjusted R^2

The adjusted R^2 takes into account the phenomenon of the R^2 automatically and spuriously increasing when extra independent variables are added to the model. The adjusted R^2 penalizes additional independent variables by considering the number of independent variables in the model and the number of data points. Its value will always be less than or equal to that of R^2 .

$$\text{Adjusted } R^2 = 1 - (1 - R^2) \frac{n - 1}{n - p}$$

where n is the number of instances and p is the number of independent variables.

Akaike information criterion (AIC)

AIC is a measure of the relative quality of statistical models for a given set of data. Given a collection of models for the data, AIC estimates the performance of each model, relative to each of the other models. Hence, AIC provides a means for model selection.

$$AIC = -2 \cdot \ln(L) + 2k$$

where L is the maximum of the likelihood function for the model and k is number of estimated parameters in the model.

The model with least AIC is preferred. AIC rewards the goodness of fit (as assessed by the likelihood function), but it also includes a penalty that increases with the number of estimated parameters.

Bayesian information criterion (BIC)

BIC is a criterion for model selection among a finite set of models; the model with the lowest BIC is preferred. It is closely related to AIC, but BIC penalizes the complexity of the model more.

$$BIC = -2 \cdot \ln(L) + k \cdot \ln(n)$$

where L is the maximum of the likelihood function for the model, k is number of estimated parameters in the model and n is the number of observations.

The BIC generally penalizes free parameters more strongly than the AIC. Lower BIC implies either fewer explanatory variables, better fit, or both.

p value

The p -value is used for null hypothesis testing. The p value is the probability of finding the observed results when the null hypothesis is true. A null hypothesis proposes that there is no relation between the response variable and features. So, if the p value is less than a significant value it means that the null

hypothesis can be rejected. In other words, it says that there exists a relationship between the features and the response variable.

B. Classification Metrics

Sensitivity

Sensitivity is a statistical measure of the performance for a binary classifier test. It refers to the test's ability to correctly detect the positive cases. In the example of a medical test used to identify a disease, the sensitivity of the test is the proportion of people who test positive for the disease among those who have the disease.

$$\text{Sensitivity} = \frac{\text{Number of true positives}}{\text{Total number of true cases}}$$

Specificity

Specificity is a statistical measure of the performance for a binary classifier test. It relates to the test's ability to correctly detect the negative cases. Consider the example of a medical test for diagnosing a disease. Specificity of a test is the proportion of identified true negatives to the total number of healthy patients known not to have the disease.

$$\text{Specificity} = \frac{\text{Number of true negatives}}{\text{Total number of negative cases}}$$

ROC & AUROC

Receiver Operating Characteristic (ROC) curve, is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied. The curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. The true-positive rate is also known as sensitivity; and the false-positive rate is the fall-out (i.e. $1 - \text{specificity}$). The ROC curve is thus the sensitivity as a function of fall-out.

Area under the Receiver Operative Characteristic (AUROC) curve is a measure that is related to the accuracy of a classifier. AUROC is the area under the curve in a ROC plot. An area of 1 represents a perfect classifier; an area of 0.5 represents random guessing.

C. Diagnostic Curves

Diagnostic curves are used to evaluate the model assumptions and investigate whether or not there are observations with a large, undue influence on the analysis.

Residuals vs Fitted

This plot shows if residuals have non-linear patterns. There could be a non-linear relationship between predictor variables and an outcome variable and the pattern could be visible in this plot. If there are equally spread residuals around a horizontal line without distinct patterns, it indicates that they do not have non-linear relationships.

Normal Q-Q

The normal Q-Q plot, or quantile-quantile plot, is a graphical tool used to assess if a set of data plausibly came from a normal distribution. A Q-Q plot is a scatterplot created by plotting two sets of quantiles against one another. If both sets of quantiles came from the same distribution, the points form a line that is marginally straight.

Scale-Location

This plot shows if the residuals are spread equally along the ranges of predictors. This is used to check the assumption of equal variance (homoscedasticity). The model is homoscedastic if the points are equally (randomly) spread out in the plot.

Residuals vs Leverage

This plot is used to identify the influential observations. These observations could be very influential in manipulating the model even when they contain reasonable range of the values. These observations can alter the results if excluded from analysis.

D. Statistical measures of the clinical scores

Initially, the t and p statistics of the clinical scores from the stepwise regression model were calculated. Then the absolute average of t values and the average of p values, corresponding to the frequently occurring clinical scores across the stepwise regression models, were computed.

Table 13: Statistical measures of the clinical scores

Gait Features	Clinical Scores		Absolute average t value	Average p value
Warp Score	MSWS	Limited your balance when standing or walking?	3.27	< 0.05
	MSWS	Made it necessary for you to use support when walking outdoors?	4.45	< 0.05
	EDSS	Gait ataxia, limb ataxia, tandem walking, etc.	2.35	< 0.05
	EDSS	Urinary incontinence, bowel dysfunction, etc.	1.99	0.05
	MFIS	I have had difficulty paying attention for long periods of time	2.7	< 0.05
	MFIS	I have been less motivated to do anything that requires physical effort	1.82	0.07
	MFIS	I have had difficulty making decisions	2.18	< 0.05
DTW Score	MSWS	Limited your ability to run?	3.04	< 0.05
	MFIS	I have been less motivated to do anything that requires physical effort	1.98	0.05
	MFIS	I have had trouble maintaining physical effort for long periods.	2.03	< 0.05
	MFIS	I have had difficulty making decisions.	2.13	< 0.05
	MFIS	My muscles have felt weak.	2.32	< 0.05
	MFIS	I have needed to rest more often or for longer periods.	2.74	< 0.05

Walking Speed	MSWS	Made it necessary for you to use support when walking outdoors	2.42	< 0.05
	MSWS	Affected how smoothly you walk?	2.24	< 0.05
	EDSS	Urinary incontinence, bowel dysfunction, etc.	1.81	0.07
	EDSS	Hearing loss, dysphagia, etc.	4.23	< 0.05
	MFIS	I have been unable to think clearly.	1.92	0.05
	MFIS	I have been forgetful.	2.03	< 0.05
	MFIS	I have been limited in my ability to do things away from home	1.88	0.06
	MFIS	My muscles have felt weak.	2.33	< 0.05
	MFIS	My thinking has been slowed down.	3.37	< 0.05
	MFIS	I have needed to rest more often or for longer periods.	2.99	< 0.05

A clinical score with a higher t statistic has a stronger relationship with the gait feature. The absolute average of the t statistic for all the clinical scores, mentioned in Table 13, are approximately equal to or higher than 2. The corresponding p values were derived to identify the “strength” of the relationship between the clinical scores and the respective gait features. If the p value is less than 0.05, there is more than 95 % chance that a statistically significant relationship exists between the clinical score and the gait feature. The average p values of the clinical scores shown in Table 13 are nearly equal to or less than 0.05, which provides evidence that the clinical scores are related to the respective gait features.

References

- [1] A. I. Cuesta-Vargas, A. Galán-Mercant, and J. M. Williams, “The use of inertial sensors system for human motion analysis,” *Phys. Ther. Rev.*, vol. 15, no. 6, pp. 462–473, Dec. 2010.
- [2] “Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson’s disease.” [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0966636214000241>. [Accessed: 09-Jun-2016].
- [3] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, “A review of wearable sensors and systems with application in rehabilitation,” *J. NeuroEngineering Rehabil.*, vol. 9, p. 21, 2012.
- [4] B. Lieber, B. E. S. Taylor, G. Appelboom, G. McKhann, and E. S. Connolly, “Motion Sensors to Assess and Monitor Medical and Surgical Management of Parkinson Disease,” *World Neurosurg.*, vol. 84, no. 2, pp. 561–566, Aug. 2015.
- [5] A. K. Roy, Y. Soni, and S. Dubey, “Enhancing effectiveness of motor rehabilitation using kinect motion sensing technology,” in *2013 IEEE Global Humanitarian Technology Conference: South Asia Satellite (GHTC-SAS)*, 2013, pp. 298–304.
- [6] A. F. Burnett, C. J. Barrett, R. N. Marshall, B. C. Elliott, and R. E. Day, “Three-dimensional measurement of lumbar spine kinematics for fast bowlers in cricket,” *Clin. Biomech. Bristol Avon*, vol. 13, no. 8, pp. 574–583, Dec. 1998.
- [7] T. Jämsä, A. Vainionpää, R. Korpelainen, E. Vihriälä, and J. Leppäluoto, “Effect of daily physical activity on proximal femur,” *Clin. Biomech. Bristol Avon*, vol. 21, no. 1, pp. 1–7, Jan. 2006.
- [8] W. Y. Wong and M. S. Wong, “Trunk posture monitoring with inertial sensors,” *Eur. Spine J. Off. Publ. Eur. Spine Soc. Eur. Spinal Deform. Soc. Eur. Sect. Cerv. Spine Res. Soc.*, vol. 17, no. 5, pp. 743–753, May 2008.
- [9] S. Daukantas, V. Marozas, and A. Lukosevicius, “Inertial sensor for objective evaluation of swimmer performance,” 2008, pp. 321–324.

- [10]J. Gong and J. Lach, “Motion marker discovery from inertial body sensors for enhancing objective assessment of robotic surgical skills,” in *2015 International Symposium on Bioelectronics and Bioinformatics (ISBB)*, 2015, pp. 215–218.
- [11]C. Randell and H. Muller, “Context awareness by analysing accelerometer data,” in *The Fourth International Symposium on Wearable Computers*, 2000, pp. 175–176.
- [12]K. V. Laerhoven, A. Schmidt, and H. W. Gellersen, “Multi-sensor context aware clothing,” in *Sixth International Symposium on Wearable Computers, 2002. (ISWC 2002). Proceedings*, 2002, pp. 49–56.
- [13]L. Bao and S. S. Intille, “Activity Recognition from User-Annotated Acceleration Data,” in *Pervasive Computing*, A. Ferscha and F. Mattern, Eds. Springer Berlin Heidelberg, 2004, pp. 1–17.
- [14]N. Kern, B. Schiele, and A. Schmidt, “Multi-sensor Activity Context Detection for Wearable Computing,” in *Ambient Intelligence*, E. Aarts, R. W. Collier, E. van Loenen, and B. de Ruyter, Eds. Springer Berlin Heidelberg, 2003, pp. 220–232.
- [15]J. A. Ward, P. Lukowicz, G. Troster, and T. E. Starner, “Activity Recognition of Assembly Tasks Using Body-Worn Microphones and Accelerometers,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1553–1567, Oct. 2006.
- [16]“WHO | Atlas: Multiple Sclerosis Resources in the World 2008,” *WHO*. [Online]. Available: http://www.who.int/mental_health/neurology/atlas_multiple_sclerosis_resources_2008/en/. [Accessed: 06-Jun-2016].
- [17]“Contribution of impaired mobility and general symptoms to the burden of multiple sclerosis. - PubMed - NCBI.” [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20082242>. [Accessed: 06-Jun-2016].
- [18]C. M. O’Connor, S. K. Thorpe, M. J. O’Malley, and C. L. Vaughan, “Automatic detection of gait events using kinematic data,” *Gait Posture*, vol. 25, no. 3, pp. 469–474, Mar. 2007.
- [19]K. R. Kaufman, C. Hughes, B. F. Morrey, M. Morrey, and K.-N. An, “Gait characteristics of patients with knee osteoarthritis,” *J. Biomech.*, vol. 34, no. 7, pp. 907–915, Jul. 2001.

- [20]J. A. Zeni Jr., J. G. Richards, and J. S. Higginson, “Two simple methods for determining gait events during treadmill and overground walking using kinematic data,” *Gait Posture*, vol. 27, no. 4, pp. 710–714, May 2008.
- [21]H. Stolze, J. P. Kuhtz-Buschbeck, C. Mondwurf, A. Boczek-Funcke, K. Jöhnk, G. Deuschl, and M. Illert, “Gait analysis during treadmill and overground locomotion in children and adults,” *Electroencephalogr. Clin. Neurophysiol. Mot. Control*, vol. 105, no. 6, pp. 490–497, Dec. 1997.
- [22]A. Muro-de-la-Herran, B. García-Zapirain, and A. Méndez-Zorrilla, “Gait Analysis Methods: An Overview of Wearable and Non-Wearable Systems, Highlighting Clinical Applications,” *Sensors*, vol. 14, no. 2, pp. 3362–3394, Feb. 2014.
- [23]A. Chetta, A. Rampello, E. Marangio, S. Merlini, F. Dazzi, M. Aiello, F. Ferraro, A. Foresi, M. Franceschini, and D. Olivieri, “Cardiorespiratory response to walk in multiple sclerosis patients,” *Respir. Med.*, vol. 98, no. 6, pp. 522–529, Jun. 2004.
- [24]M. D. Goldman, R. A. Marrie, and J. A. Cohen, “Evaluation of the six-minute walk in multiple sclerosis subjects and healthy controls,” *Mult. Scler.*, Oct. 2007.
- [25]“ActiGraph.” [Online]. Available: <http://actigraphcorp.com/>. [Accessed: 22-Jun-2016].
- [26]J. F. Kurtzke, “Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS),” *Neurology*, vol. 33, no. 11, pp. 1444–1452, Nov. 1983.
- [27]“Measuring the impact of MS on walking ability: the 12-Item MS Walking Scale (MSWS-12). - PubMed - NCBI.” [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/12525714>. [Accessed: 22-Jun-2016].
- [28]R. J. Mills, C. A. Young, J. F. Pallant, and A. Tennant, “Rasch analysis of the Modified Fatigue Impact Scale (MFIS) in multiple sclerosis,” *J. Neurol. Neurosurg. Psychiatry*, vol. 81, no. 9, pp. 1049–1051, Sep. 2010.
- [29]S. Abrahams, S. M. Fie, J. Patricios, M. Posthumus, and A. V. September, “Risk factors for sports concussion: an evidence-based systematic review,” *Br. J. Sports Med.*, vol. 48, no. 2, pp. 91–97, Jan. 2014.

- [30]“Concussion - Mayo Clinic.” [Online]. Available: <http://www.mayoclinic.org/diseases-conditions/concussion/basics/definition/con-20019272>. [Accessed: 03-Jun-2016].
- [31]E. L. Breedlove, M. Robinson, T. M. Talavage, K. E. Morigaki, U. Yoruk, K. O’Keefe, J. King, L. J. Leverenz, J. W. Gilger, and E. A. Nauman, “Biomechanical correlates of symptomatic and asymptomatic neurophysiological impairment in high school football,” *J. Biomech.*, vol. 45, no. 7, pp. 1265–1272, Apr. 2012.
- [32]“PLOS ONE: Consequences of Repeated Blood-Brain Barrier Disruption in Football Players.” [Online]. Available: <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0056805>. [Accessed: 21-Jun-2016].
- [33]C. D. Solutions and C. D. Solutions, “X2 BIOSYSTEMS,” *X2 BIOSYSTEMS*. [Online]. Available: <http://www.x2biosystems.com/>. [Accessed: 23-Jun-2016].
- [34]R. Jadischke, D. C. Viano, N. Dau, A. I. King, and J. McCarthy, “On the accuracy of the Head Impact Telemetry (HIT) System used in football helmets,” *J. Biomech.*, vol. 46, no. 13, pp. 2310–2315, Sep. 2013.
- [35]P. Russom and others, “Big data analytics,” *TDWI Best Pract. Rep. Fourth Quart.*, pp. 1–35, 2011.
- [36]R. Cooley, B. Mobasher, and J. Srivastava, “Data Preparation for Mining World Wide Web Browsing Patterns,” *Knowl. Inf. Syst.*, vol. 1, no. 1, pp. 5–32, Jul. 2013.
- [37]E. McLellan, K. M. MacQueen, and J. L. Neidig, “Beyond the Qualitative Interview: Data Preparation and Transcription,” *Field Methods*, vol. 15, no. 1, pp. 63–84, Feb. 2003.
- [38]M. Sekine, T. Tamura, M. Akay, T. Fujimoto, T. Togawa, and Y. Fukui, “Discrimination of walking patterns using wavelet-based fractal analysis,” *IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc.*, vol. 10, no. 3, pp. 188–196, Sep. 2002.
- [39]M. Mathie, “Monitoring and Interpreting Human Movement Patterns Using a Triaxial Accelerometer,” *ResearchGate*, Jan. 2003.

- [40]U. Maurer, A. Rowe, A. Smailagic, and D. Siewiorek, "Location and Activity Recognition Using eWatch: A Wearable Sensor Platform," in *Ambient Intelligence in Everyday Life*, Y. Cai and J. Abascal, Eds. Springer Berlin Heidelberg, 2006, pp. 86–102.
- [41]J. Baek, G. Lee, W. Park, and B.-J. Yun, "Accelerometer Signal Processing for User Activity Detection," in *Knowledge-Based Intelligent Information and Engineering Systems*, M. G. Negoita, R. J. Howlett, and L. C. Jain, Eds. Springer Berlin Heidelberg, 2004, pp. 610–617.
- [42]"The prediction of speed and incline in outdoor running in humans using accelerometry. - PubMed - NCBI." [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/10416569>. [Accessed: 06-Jul-2016].
- [43]K. Aminian, P. Robert, E. Jéquier, and Y. Schutz, "Incline, speed, and distance assessment during unconstrained walking," *Med. Sci. Sports Exerc.*, vol. 27, no. 2, pp. 226–234, Feb. 1995.
- [44]F. Foerster and J. Fahrenberg, "Motion pattern and posture: correctly assessed by calibrated accelerometers," *Behav. Res. Methods Instrum. Comput. J. Psychon. Soc. Inc*, vol. 32, no. 3, pp. 450–457, Aug. 2000.
- [45]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 Trans. Biomed. Eng.*, vol. 56, no. 3, pp. 871–879, Mar. 2009.
- [46]T. Huynh and B. Schiele, "Analyzing Features for Activity Recognition," in *Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-aware Services: Usages and Technologies*, New York, NY, USA, 2005, pp. 159–163.
- [47]M. Sekine, T. Tamura, T. Fujimoto, and Y. Fukui, "Classification of walking pattern using acceleration waveform in elderly people," in *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2000*, 2000, vol. 2, pp. 1356–1359 vol.2.
- [48]M. N. Nyan, F. E. H. Tay, K. H. W. Seah, and Y. Y. Sitoh, "Classification of gait patterns in the time-frequency domain," *J. Biomech.*, vol. 39, no. 14, pp. 2647–2656, 2006.

- [49]N. Wang, E. Ambikairajah, N. H. Lovell, and B. G. Celler, “Accelerometry Based Classification of Walking Patterns Using Time-frequency Analysis,” in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2007, pp. 4899–4902.
- [50]J. C. Lucero, K. G. Munhall, V. L. Gracco, and J. O. Ramsay, “On the Registration of Time and the Patterning of Speech Movements,” *J. Speech Lang. Hear. Res.*, vol. 40, no. 5, p. 1111, Oct. 1997.
- [51]E. Keogh, L. Wei, X. Xi, M. Vlachos, S.-H. Lee, and P. Protopapas, “Supporting Exact Indexing of Arbitrarily Rotated Shapes and Periodic Time Series Under Euclidean and Warping Distance Measures,” *VLDB J.*, vol. 18, no. 3, pp. 611–630, Jun. 2009.
- [52]A. D. Wilson and A. F. Bobick, “Parametric hidden Markov models for gesture recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 9, pp. 884–900, Sep. 1999.
- [53]M. Raptis, K. Wnuk, and S. Soatto, “Spike train driven dynamical models for human actions,” in *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 2077–2084.
- [54]A. Bissacco and S. Soatto, “Classifying Human Dynamics Without Contact Forces,” in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06)*, 2006, vol. 2, pp. 1678–1685.
- [55]J. Gong, P. Asare, Y. Qi, and J. Lach, “Piecewise Linear Dynamical Model for Action Clustering from Real-World Deployments of Inertial Body Sensors,” *IEEE Trans. Affect. Comput.*, vol. PP, no. 99, pp. 1–1, 2016.
- [56]M. Ermes, J. Parkka, J. Mantjarvi, and I. Korhonen, “Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions,” *IEEE Trans. Inf. Technol. Biomed. Publ. IEEE Eng. Med. Biol. Soc.*, vol. 12, no. 1, pp. 20–26, Jan. 2008.
- [57]J. Parkka, M. Ermes, P. Korpipaa, J. Mantjarvi, J. Peltola, and I. Korhonen, “Activity classification using realistic data from wearable sensors,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 119–128, Jan. 2006.
- [58]J. Liu and T. E. Lockhart, “Development and Evaluation of a Prior-to-Impact Fall Event Detection Algorithm,” *IEEE Trans. Biomed. Eng.*, vol. 61, no. 7, pp. 2135–2140, Jul. 2014.

- [59]L. Palmerini, L. Rocchi, S. Mellone, F. Valzania, and L. Chiari, "Feature Selection for Accelerometer-Based Posture Analysis in Parkinson's Disease," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 3, pp. 481–490, May 2011.
- [60]P. Reberšek, D. Novak, J. Podobnik, and M. Munih, "Intention detection during gait initiation using supervised learning," in *2011 11th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2011, pp. 34–39.
- [61]R. Ganea, P.-Y. Jeannet, A. Paraschiv-Ionescu, N. M. Goemans, C. Piot, M. V. den Hauwe, and K. Aminian, "Gait Assessment in Children With Duchenne Muscular Dystrophy During Long-Distance Walking," *J. Child Neurol.*, vol. 27, no. 1, pp. 30–38, Jan. 2012.
- [62]J. C. Platt, "Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods," in *ADVANCES IN LARGE MARGIN CLASSIFIERS*, 1999, pp. 61–74.
- [63]G. Uslu and S. Baydere, "Support Vector Machine based activity detection," in *Signal Processing and Communications Applications Conference (SIU), 2013 21st*, 2013, pp. 1–4.
- [64]A. Ataya and P. Jallon, "Amelioration of physical activity estimation from accelerometer sensors using prior knowledge," in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, 2012, pp. 954–958.
- [65]W. C. Cheng and D. M. Jhan, "Triaxial Accelerometer-Based Fall Detection Method Using a Self-Constructing Cascade-AdaBoost-SVM Classifier," *IEEE J. Biomed. Health Inform.*, vol. 17, no. 2, pp. 411–419, Mar. 2013.
- [66]S. Patel, K. Lorincz, R. Hughes, N. Huggins, J. Growdon, D. Standaert, M. Akay, J. Dy, M. Welsh, and P. Bonato, "Monitoring Motor Fluctuations in Patients With Parkinsons Disease Using Wearable Sensors," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 6, pp. 864–873, Nov. 2009.
- [67]D. Kumar, J. Gubbi, B. Yan, and M. Palaniswami, "Motor recovery monitoring in post acute stroke patients using wireless accelerometer and cross-correlation," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, pp. 6703–6706.

- [68]N. Shibuya, B. T. Nukala, A. I. Rodriguez, J. Tsay, T. Q. Nguyen, S. Zupancic, and D. Y. C. Lie, “A real-time fall detection system using a wearable gait analysis sensor and a Support Vector Machine (SVM) classifier,” in *2015 Eighth International Conference on Mobile Computing and Ubiquitous Networking (ICMU)*, 2015, pp. 66–67.
- [69]N. C. Krishnan and S. Panchanathan, “Analysis of low resolution accelerometer data for continuous human activity recognition,” in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008, pp. 3337–3340.
- [70]M. N. S. Zainudin, M. N. Sulaiman, N. Mustapha, and T. Perumal, “Activity recognition based on accelerometer sensor using combinational classifiers,” in *2015 IEEE Confernece on Open Systems (ICOS)*, 2015, pp. 68–73.
- [71]J. Knighten, S. McMillan, T. Chambers, and J. Payton, “Recognizing social gestures with a wrist-worn smartband,” in *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, 2015, pp. 544–549.
- [72]F. Masse, R. Gonzenbach, A. Paraschiv-Ionescu, A. Luft, and K. Aminian, “Wearable Barometric Pressure Sensor to Improve Postural Transition Recognition of Mobility-Impaired Stroke Patients,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. PP, no. 99, pp. 1–1, 2016.

List of Publications

- [P1] Engelhard MM, **Dandu SR**, Patek SD, Lach JC, and Goldman MD, "Quantifying Six-Minute Walk Induced Gait Deterioration with Inertial Sensors in Multiple Sclerosis Subjects," *Gait & Posture* 2016
- [P2] **Dandu, S.R.**, Engelhard, M., Goldman, M., Lach, J., "Determining Physiological Significance of Inertial Gait Features in Multiple Sclerosis." *Wearable and Implantable Body Sensor Networks (BSN), 2016 IEEE 13th International Conference on*. IEEE, 2016.
- [P3] Engelhard, M.M., **Dandu, S.R.**, Lach, J.C., Goldman, M.D. and Patek, S.D., "Toward detection and monitoring of gait pathology using inertial sensors under rotation, scale, and offset invariant dynamic time warping." *Proceedings of the 10th EAI International Conference on Body Area Networks*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2015.
- [P4] Gong, J., **Dandu, S. R.**, Lach, J., et al. "Home wireless sensing system for monitoring nighttime agitation and incontinence in patients with Alzheimer's disease." *Proceedings of the conference on Wireless Health*. ACM, 2015.