OXFORD BROOKES UNIVERSITY

Development of Machine Learning Models to Detect Dynamic Disturbances in Human Gait

A thesis submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Computing and Mathematical Sciences

by

Shadi Eltanani

Committee in charge:

Dr Tjeerd V. olde Scheper, Chair
Professor Helen Dawes
Dr Peter Ball
Professor Angelo Cangelosi

2022

The thesis of Shadi Eltanani is approved, and it is acceptable in quality and form for publication on microfilm:

_____

_____

_____

_____

_____

Chair

Oxford Brookes University

2022

DEDICATION

To the soul of my father, Yusuf who always wished me the best success in my study and life

To my lovely mother, Amal

To my Wife, Hadil

and

to my dearest brothers and sisters

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Firstly, I express my gratitude to God for providing me with countless blessings, patience, and strength to successfully accomplish my goals.

I am thankful to my supervisor, Dr. Tjeerd V. olde Scheper, for his unwavering support, guidance, and extensive knowledge during my PhD research. His exceptional care, attention, and motivation have impacted me both professionally and personally, and I aim to carry forward his teachings and leadership in my future academic endeavors.

I extend my sincere appreciation to Prof. Helen Dawes and Dr. Patrick Esser for their valuable assistance and fpr providing me with the clinical data to conduct my research, without which my research would not have been possible. It is an honor to have Prof. Helen Dawes as my second supervisor.

I would like to thank my colleagues and friends at the School of Engineering, Computing, and Mathematics at Oxford Brookes University for their support and making my PhD journey enjoyable. My gratitude goes to Prof. Nigel Crook, Mrs. Gordana Collier, Dr. Matthias Rolf, Prof. Michael Todinov, Mr. Peter Whittle, and other staff members.

I am grateful to Oxford Brookes University for awarding me a fully-funded studentship to pursue my PhD and providing me with the opportunity to work as an Associate Lecturer to gain teaching experience alongside my research.

My family deserves my deepest appreciation and lifelong gratitude. My parents always emphasised the importance of education and worked tirelessly to ensure we received the best possible education. Completing my PhD was a childhood ambition, and I am thankful for the unwavering support, encouragement, and trust of my siblings. I also thank my wife, Hadil, for her care and support throughout my PhD research.

VITA

2004-2009	B. S. in Electrical Engineering, Islamic University of Gaza, Gaza Strip, Palestine
2014-2015	M. Sc. in Telecommunication Engineering, Staffordshire University, Stoke-on-Trent, United Kingdom
2019-2022	Ph. D. in Computer Science, Oxford Brookes University, Oxford, United Kingdom

PUBLICATIONS

Shadi Eltanani, Tjeerd Olde Scheper and Helen Dawes. A Novel Criticality Analysis Technique for Detecting Dynamic Disturbances in Human Gait. *Computers* 2022, 11, 120. https://doi.org/10.3390/computers11080120.

Shadi Eltanani, Tjeerd Olde Scheper and Helen Dawes, K-Nearest Neighbour Algorithm: Proposed Solution for Human Gait Data Classification, 2021 Symposium on Computers and Communications (ISCC): 26^{th} *IEEE Symposium on Computers and Communications* - Workshop on ICT Solutions for eHealth (ICTS4eHealth) (ICTS4eHealth2021).

Shadi Eltanani, Tjeerd Olde Scheper and Helen Dawes. What Role Criticality Analysis Play in Human Age Estimation?, (In-progress to appear in the *IEEE Transactions on Pattern Analysis and Machine Intelligence*).

Shadi Eltanani, Tjeerd Olde Scheper and Helen Dawes. Detecting Human Gait Metabolism Disorders Based on A Novel Criticality Analysis Methodology, (In-progress to appear in the IEEE Transactions on Biomedical Engineering).

ABSTRACT OF THE THESIS

Development of Machine Learning Models to Detect Dynamic Disturbances in Human Gait

by

Shadi Eltanani
Doctor of Philosophy in Computing and Mathematical Sciences
Oxford Brookes University, 2022
Dr Tjeerd V. olde Scheper, Chair

Machine learning has transformed the medical field by automating tasks and achieving objectives that are closer to human cognitive capabilities. Gait is a series of intricate interactions for humans, and identifying impaired gait is critical for effective decision-making in clinical practice. However, analysing the complex biological system that governs gait can be challenging. This research proposes a novel Criticality Analysis (CA) methodology to extract dynamic interactions in human gait and represent multivariate data in a nonlinear space. The proposed methodology characterises each data sample with a unique orbit, resulting from perturbations of a critical system composed of nonlinear controlled oscillators. The scale-free network of orbits is a quantitative measure of non-scale-free interacting sets of patterns, which reveal organised features of the structure of dynamic properties interconnected with human gait. This thesis focuses on implementing robust machine learning algorithms for effective detection and classification of complex dynamic patterns in human gait. The CA method maps gait features into a nonlinear representation, which is then used for training and testing categorisation algorithms. The proposed models utilise the Kernel property of the Support Vector Machines (SVM) classifier to identify

high-order interactions between multiple gait data variables that may be challenging for traditional statistics. The algorithm was applied to three real datasets, and the SVM models designed using the CA method achieved an accuracy of 88.27% on average, compared to the K-Nearest Neighbours (KNN) approach's accuracy of 67.7%. The proposed SVM models use the receiver operating characteristics (ROC) and the area under the ROC metrics to evaluate their overall performance. The results of this research suggest that the proposed SVM models, with the support of the CA method, can perform as a robust and reliable classification tool for detecting dynamic disturbances of biological data patterns. This provides tremendous opportunities for clinical diagnosis and rehabilitation.

Chapter 1

Introduction

1.1 Motivation

Human gait is a complex and nonlinear process that involves the integration of many different systems in the body, including the musculoskeletal, neurological, and cardiovascular systems [1]. Gait is a fundamental part of human locomotion and is essential for daily activities such as walking, running, and jumping. Factors such as body weight, height, age, gender, and physical fitness level influence gait [2]. Gait analysis is a tool used to measure and evaluate the biomechanical parameters of gait, such as step length, stride length, step width, cadence, and velocity [3]. This analysis helps in the diagnosis and treatment of various gait-related disorders, including Parkinson's disease, cerebral palsy, stroke, and spinal cord injuries [208].

However, capturing and interpreting neurophysiological signals, such as gait for humans, is crucial for medical diagnosis. Human gait represents nonlinear, nonstationary, and non-equilibrium biological processes that require advanced technology or techniques to provide healthcare professionals with a useful clinical value of the intricate sequential spatiotemporal interacting states of the motor system function [4]. Medical professionals often ask their patients to walk for several steps and monitor

the pattern of their walk to obtain a rough estimate of the behaviour of human gait for better diagnosis. However, the interpretation of the biological information obtained from gait analysis can be challenging.

One of the challenges in studying human gait is the nonlinear dynamics involved. Gait is a highly complex and nonlinear system, with many different factors influencing the movement of the body. Criticality analysis is a mathematical technique that can be used to study the nonlinear dynamics of gait patterns. Criticality analysis involves analysing the behaviour of a system at or near a critical point, where small changes in the input can lead to large changes in the output. In the context of gait analysis, criticality analysis can help to identify the stable and controlled orbits of the nonlinear dynamics of gait patterns. By identifying these orbits, it is possible to understand the underlying mechanisms that govern gait and develop strategies for improving gait performance. The importance of criticality analysis extends beyond gait analysis and into the broader field of machine learning and AI. Criticality analysis can be used to study the behaviour of complex systems, such as neural networks, and to identify the critical points where small changes in input can lead to large changes in output. This can help to improve the performance and stability of machine learning algorithms and enable the development of more sophisticated AI systems. Thus, a combination of advanced gait analysis techniques, such as criticality analysis, can help healthcare professionals to interpret the intricate sequential spatiotemporal interacting states of the motor system function and improve diagnosis and treatment of various gait-related disorders.

In this thesis, a novel approach called Criticality Analysis (CA) was developed to create a distinct nonlinear representation space from biological data, specifically for human gait analysis. The CA method can reveal intricate hidden patterns that are not discernible through traditional analysis methods [34]. The discovered information has the potential to aid in predicting gait-related injuries in older adults, identifying walking irregularities in individuals, and elucidating the fundamental mechanisms

of healthy and pathological biological system functions. The effectiveness of the CA approach was demonstrated by its ability to successfully extract and represent human gait information from measured time-series gait sensory data, as evidenced in previous studies [50]. The output of the CA technique was subsequently used in developing machine learning classification models.

1.2 Criticality Analysis Methodology

Criticality analysis (CA) is a method used in the field of physics to study the behaviour of complex systems near the critical point, where small changes in the system can lead to large-scale fluctuations [6]. In recent years, CA has been applied to the study of human gait to better understand the dynamics of the gait pattern and how it responds to perturbations [50]. The human gait is a complex, nonlinear system that involves the coordination of multiple muscle groups and joints. The dynamics of the gait pattern can be affected by a range of factors, such as changes in speed, terrain, or external perturbations [7]. CA can be used to study how the gait pattern responds to these perturbations and how it approaches the critical point, where small changes can have a large effect on the dynamics of the system [50]. In gait analysis, CA has been used to study the fluctuations in stride time and other gait parameters, as well as to identify the presence of long-range correlations in the gait pattern. Long-range correlations indicate that the fluctuations in the gait pattern are not independent but are instead correlated over long periods of time, indicating a degree of complexity in the system [8] [9]. CA has also been used to study the effects of aging and neurological disorders on the gait pattern. For example, studies have found that older adults and individuals with Parkinson's disease exhibit a reduced degree of criticality in their gait pattern, indicating a loss of complexity in the system [10] [11]. These findings suggest that CA may be a useful tool for early detection and monitoring of neurological disorders.

In biological systems, criticality arises from the local nonlinear interactions of small components, leading to complex interactions that enable the system to change states and operate at the edge of chaos. However, in some cases, the system may degenerate into unstable or chaotic states. To address this, the rate control of chaos (RCC) mechanism has been proposed, which allows for dynamic control of the system even when chaotically disrupted [13]. Mathematical modeling of such systems involves a set of variables that describe the operation of variable quantities and derive system properties. The RCC approach employs rate equations to describe how these variable quantities change over time, with a control term that is dominated by nonlinear variable components and can affect the stability of the overall system. The rate control function is then derived from the contributing variables to maintain a stable control component for the rate of change equation. The CA methodology based on the RCC concept involves outputting the variable quantities in the rate equations that dominantly show nonlinear relationships and generate nonlinear representation spaces from observations of the biological system [13].

1.3 Modelling Human Gait: Exploring the Spatiotemporal Oscillation Behaviour of Locomotion System

Since the human gait displays structured spatiotemporal oscillatory behavior, there has been no consensus within the scientific community regarding the optimisation of a mathematical model to represent the chaotic rhythmic behavior of the locomotion system based on the bioenzymatic changes in the biological system.

The study of the human locomotion system has attracted the attention of researchers from many scientific disciplines, such as engineering, physics, biology, medicine, and others [14]. The majority of research on gait has focused on dynamic

systems techniques, which are mostly based on well-known physical system ideas, including nonlinear dynamics, thermodynamics, and biological systems [15] [16].

Some approaches in the literature have studied the structured spatiotemporal behavior of human gait, taking into consideration the physical oscillators of dynamic systems such as mass-spring systems, pendula, or hybrid spring-pendulum systems [17]. When running or walking, the stance phase of the gait cycle has been represented as a mass-spring system and an inverted pendulum, respectively. Models of a pendulum and a hybrid spring-pendulum system have been used to describe the swing phase of gait. A spring-mass system is an oscillatory system with a mass attached to the free end of a spring. Using the force, displacement, stiffness, mass, and frequency properties of this system, human gait can be modelled [17]. The lower extremities resemble a spring, and the center of mass of the body serves as a point mass that acts downward on the neuromusculoskeletal spring [18]. The spring-like lower extremity exhibits the ability to self-assemble into a spring-like structure when forced to move at different frequencies, with different masses, and on different surfaces [19]. The system makes an effort to optimise how it interacts with the environment in order to provide the proper stiffness and satisfy various task objectives. It has been demonstrated that the complexity of the biological spring changes from a linear spring at or above optimal movement frequencies to a nonlinear spring below such frequencies [20]. The relationship between the mass-spring system and the stance phase of the gait cycle was believed to have great promise for identifying gait abnormalities and determining the best diagnostic procedures and forms of treatment. The output of this model includes a number of possible characteristics, namely frequency, amplitude, and velocity [18]. Additionally, the indirect relationship between force and displacement that is derived from this model improves the clinical evaluation of biological stiffness [21].

The inverted pendulum model, which assumes that the lower extremities and trunk are rigid segments that rotate over the fixed foot in the sagittal plane around

the talocrural joint axis, has been used by researchers to study the stance phase of gait during walking. Kram *et al.* [22] demonstrated the transition of gait between walk and run cycles using the inverted pendulum model to investigate the storage and recovery of mechanical energy, foot posture, and the stability of the frontal plane. The inverted pendulum model is expanded to incorporate active neuromuscular variables, which constitute a dynamic system. The active control of gait derives from the interplay of the hip and foot muscles with the gravitational factor [23]. At intermediate gait velocities, the inverted pendulum system exhibits a unique minimum energy state when the stance leg recovers the most energy [24]. During fast walking and running, the model exhibits a dynamic shift towards a more spring-like behavior during the stance phase, which limits its performance.

During the swing phase of the gait cycle, the motion of the leg has been optimised to resemble a simple pendulum. This basic pendulum model has a variety of configurations, including simple and compound pendulums, as well as a hybrid force-driven harmonic oscillator. Among these configurations, the pendulum model includes a solid lower extremity suspended from the pelvis. However, this model may be improved to resemble a compound pendulum with several connected segments. Mochon *et al.*[25] merged the operation of the double pendulum system, which stands in for the swing leg, with the inverted pendulum model, which represents the stance leg. Despite the fact that much research work has used pendulum dynamics to analyze the swing phase, Whittlesey *et al.*[26] criticised the idea that the swing phase of gait should be viewed as an unforced pendulum and interpreted as a passive action. In addition, Holt *et al.* [27] adopted the force-driven harmonic oscillator system's dynamic principle to describe human gait after discovering a substantial relationship between stride frequency, resonance, and energy minima. Additionally, their research demonstrated the use of varied stiffness to regulate the lower extremities' temporal behavior. The hybrid pendulum mass-spring system is the central component of the force-driven harmonic oscillator model, which accounts for both the active and

passive connective tissues of the locomotor system, as well as the stiffness of the vibratory system. The length and mass of the lower extremities, which determine their inertial characteristics, are fed into the force-driven harmonic oscillator model as inputs. The harmonic oscillator is considered to be in a situation of dynamic equilibrium when the quantity of periodic energy provided by muscle is equal to the amount of energy required to maintain the oscillation, and the system eventually exhibits observable behavior. There is a resonant frequency at the point of equilibrium that can preserve the harmonic oscillation of the system, allowing biological systems to be attracted to such preferable states.

Furthermore, Holt *et al.* [27] conducted research that precisely estimated the stride frequency of both children and adults using the force-driven harmonic oscillator model, and confirmed the presence of resonance states throughout the walking phase of the gait cycle. The biological system has the ability to detect and self-select the resonant frequency and other frequencies at a minimum level of energy during a variety of activities such as walking, running, lifting, and upper body ergometry. However, the force-driven harmonic oscillator model predicts the stride frequency at which energy consumption would be at a minimum based on the stability boundary of closed trajectories in the phase space. The locomotor system, which functions similarly to the force-driven harmonic oscillator, has characteristics of a dynamic system with preferred movement frequencies that are subject to energy limitations.

The aforementioned dynamical systems have been considered for modeling human gait based on how the various components of the musculoskeletal system are connected to each other to enable body movement. However, these models only took into account the two stages of gait: walking and running, and overlooked diseased gait, which may hide a variety of unknown unhealthy conditions. Despite these efforts, the scientific community has not reached a consensus on the research of the locomotor system based on bioenzymatic changes occurring in the biological system to not only characterise the dynamics of normal gait but also detect gait abnormal-

ities in humans.

In this thesis, this challenge was overcome by using a bioenzymatic dynamical model that incorporates the concept of rate control of chaos to limit the exponential growth of the system and enable it to respond in accordance with various changes in the biological system processes over a range of time scales by evolving, accelerating, and decelerating and showing a power-law relationship in the direction of movement. The model also has the adaptability to adopt new movement patterns in response to modifications in system conditions. Furthermore, the model can capture the actual dynamic patterns of human gait, generate nonlinear representation spaces of these perturbation inputs, and understand the biological self-organisation of the complex biological system when the biological system is exposed to perturbing forces, whether they are sourced from inside or outside the biological system. Additionally, the model can persist and remain stable even in the presence of disturbances, and show variability in its phase space patterns when the effect of perturbations becomes minimal or negligible. The dynamic system must have all of the aforementioned capabilities to produce optimal gait patterns in which every part of the complex system is operating toward a safe and energy-efficient state. Therefore, it is critical to have a dynamic model that can incorporate experimental gait data into its theoretical characteristics to detect dynamic disturbances affecting the locomotor system in humans. The details of this model are thoroughly discussed in Chapter 3.

1.4 Incorporation of Biomechanical Models for Gait Disturbances Detection

When the influence of biological enzymes oscillates, the system generates successive oscillations that allow it to behave chaotically, thereby disrupting vital organism functions such as locomotion. This opens up new horizons for thinking about how to

control this arising chaos, which acts as a source of perturbation to the main biological system, so that the global stability state function can be maintained. Therefore, it is imperative to understand how biological systems function in terms of a biochemical mathematical model. However, unfortunately, some simplified or reduced models represented by a set of ordinary differential equations in the time domain do not sufficiently comprise the required complexity to characterise how the control mechanism, responsible for translating the network behavior into spatial outcomes in a biological system, maintains global stable and controlled states.

Biologically inspired by biochemical dynamic models, Hugues Berry [29] demonstrated a general bienzymatic cyclic system using two autocatalytic loops to support and give structure to cells and tissues in a biological system. This helps them communicate with each other and enhance their movement and functions. The model adopts the kinetic rates of enzymes and is described by four ordinary differential equations. This model is believed to provide oscillatory behaviors of complex biological processes, including human gait, and is therefore optimised based on the rate control of chaos concept to control the disturbances (as chaos) of human gait. Technical details of this model are included in Chapter 3.

Moreover, Decroley *et al.* [28] demonstrated a biochemical model that operates similarly to the Berry model [29]. They studied the dynamics of two allosteric enzymes, where the enzyme itself is attached to small regulators (molecules) that change the shape of the enzyme and affect its catalytic activity. The autocatalytic loops of the two enzymes in the Berry model [29] originate from the synthesis of the enzyme by the cell itself, while the autocatalytic loops of the two allosteric enzymes in the other model arise directly from the regulated molecules bound in the separate cavity of the enzyme. Although each model has its own characteristics, both provide complex dynamics (oscillatory behaviors).

1.5 RCC Method Based on Theoretical Dynamic Model

The significance of merging experimental data into theoretical dynamic models to analyse the behavior of disturbances that perturb the human locomotor system has been discussed. Experimental measurements, commonly known as noisy perturbing sources, are frequently used to find system parameters and provide system inputs for theoretical modeling. To validate the theoretical modeling method, some empirically measured parameters might be employed. A theoretical model, on the other hand, may be used to guide the development of experimental methodologies that can be utilised for routine clinical examinations. Furthermore, employing experimental data assists in the estimation of non-measurable quantities as well as the interpretation of experimental outcomes. Spatiotemporal information and moment inertia data of human gait are employed in experimental measurements to establish system features required for theoretical dynamic modeling.

Regardless of the representative biological model used, the RCC approach [208] is considered a robust bio-inspired metabolic feedback control mechanism that allows small external perturbations to grow exponentially in the biological system, enabling it to stabilise any of the unstable periodic states that are inherently involved. The unstable periodic chaotic response, for example, can be suitably stabilised into a controlled periodic of oscillations by perturbing the operating conditions of the biological system itself. Controlling the spatiotemporal changes or disturbances of human gait patterns on a smaller scale requires tuning the biological system parameters in both time and space. One of the best advantages of using the RCC concept in such a dynamic model is that it can be applied without prior knowledge of a specific mechanism of the biological system but requires some instantaneous knowledge of some of the variables of the perturbed system, unlike other control methods that require an accurate model to describe the biological system to retain it in a stabilising mode.

In this way, it is always possible, depending on the degree of applied disturbances, to guide chaotic gait disturbances into different nonlinear oscillatory controlled stable orbits. The interrelation between the biological system perturbations and the amplitude of controlled chaotic orbits shows power-law and exponential relations. This also provides a structured representation at the edge of chaos of various biological properties, such as those interconnected with the motor system of human gait, which can remotely reflect a deep understanding of the underlying complex processes and their corresponding functionalities. This is exactly what the proposed methodology of the CA does. The CA technique allows for the nonlinear representation of various human gait features that have perturbed the biological system and accommodates them in a reduced lower-dimensional space.

1.6 Thesis Goals

Recent research has focused on dynamical modelling approaches, but stabilising the chaotic behavior of these methods using another chaotic source, such as disturbing inputs, to maintain stability and generate distinctive dynamical orbits based on gait oscillations has received little attention. This prompts several research questions to be addressed in this thesis:

- **Research Question 1:** *Is it possible to identify the non-linear dynamics of human gait using the criticality analysis technique?*
- **Research Question 2:** *How effective are the supervised machine learning models, specifically the Support Vector Machine (SVM) classifiers, that have been developed into practice for identifying and categorising patterns of human gait disturbance?*

The primary objective of this thesis is to contribute to filling gaps in the research field of chaos control theory by creating a series of processes that utilise existing work

based on the criticality analysis technique. The purpose is to improve the detection of disturbances related to human gait and create innovative stable patterns of nonlinear gait dynamics in a reduced dimensional space.

The results of the criticality analysis system will be used to train supervised machine learning models for classification and detection. To generate the gait patterns that act as disturbance sources for the criticality analysis system, the CARDIGAN, MoRES, and Museum datasets will be employed. To validate the effectiveness of the criticality analysis system, two supervised machine learning algorithms, the K-Nearest Neighbour (KNN) and the Support Vector Machine (SVM), have been proposed. The KNN has only been implemented for training the gait features of the MoRES dataset due to its low model accuracy. Meanwhile, the SVM has been designed on all three datasets as it has shown promising performance in classifying the extracted nonlinear gait patterns. To achieve generalisation performance, specific model parameters were adapted to produce an average accuracy rate of 90% when trained on the three datasets.

1.7 Thesis Contributions

The contribution of the thesis is significant in advancing the understanding of human gait analysis using criticality analysis (CA) technique and supervised machine learning models. The research provides a thorough analysis of three datasets of real-time human gait data, which is a critical aspect in the development of accurate gait recognition and monitoring systems. The use of the support vector machine (SVM) classifier to identify nonlinear spatiotemporal patterns and detect disturbances in human gait is a novel approach that adds to the existing literature. The performance evaluation of the generated SVM models using relevant metrics such as the confusion matrix, receiver operating characteristics (ROC) curve, and the area under the ROC curve provides a comprehensive analysis of the models effectiveness. Furthermore, the

implementation of the K-Nearest Neighbour (KNN) machine learning model to test the validity of the data generated from the CA of the MoRES dataset is an innovative step towards identifying gait patterns among various participants. Overall, this thesis provides a valuable contribution to the field of human gait analysis and has the potential to enhance the development of effective gait recognition and monitoring systems.

1.8 Thesis Overview

The remainder of the thesis is organised in the following chapters as follows:

Chapter 2 describes an overview of gait analysis and its application using advanced technology, such as smart motion capture systems, smart homes, and virtual reality environments. It then reviews the use of machine learning (ML) for analysing gait, including gait phase recognition, event prediction, detection of abnormal gait patterns, and monitoring neurological diseases, as well as gender recognition. The chapter also highlights limitations in previously reported work in this area. Finally, the chapter reviews specific applications of support vector machines (SVM) for gait analysis.

Chapter 3 provides an overview of the biochemical mathematical model for the CA that incorporates the RCC method as a fundamental design principle for modelling disturbances in the locomotion system. The model comprises four ordinary differential equations (ODEs) that describe the rate of enzymatic reactions in the biological system. Along with the set of ODEs, two control exponential terms have been established as RCC functions, which will be merged into the original ODEs to regulate the release of enzymes and uphold the stability of the model. Finally, the model's output is characterised by two variables that depict the controlled gait

disturbances.

Chapter 4 outlines a comparative analysis of two supervised machine learning algorithms, Support Vector Machines (SVM) and K-Nearest Neighbour s (KNN), for human gait analysis. Technical details for each technique, along with a comprehensive explanation of the implementation process, are provided. It also addresses the implementation of the proposed SVM algorithm processes and how the SVM model employs the Kernel Trick feature to overcome the nonlinearity issue in human gait data. Finally, the chapter emphasises the critical factors to consider when evaluating the model's performance.

Chapter 5 investigates the application of the proposed SVM method on the CARDIGAN dataset, which provides insights into the gait patterns of both healthy individuals and those suffering from obesity-related health disorders. Criticality analysis features were derived from the RCC-optimised Berry model [29] and segmented into healthy control and obese groups of various individuals. The gait measurements for each participant were recorded over a six-week trial period, and to gauge the efficacy of the proposed models, a battery of tests such as the confusion matrix, receiver operating characteristics (ROC), area under the ROC curve, and mean square error (MSE) were performed. The simulation results demonstrate that the Kernelised SVM approach is a robust classifier tool, accurately distinguishing the unique dynamics of gait disruptions. Notably, the material in this chapter will be submitted for publication in the IEEE Transactions on Biomedical Engineering, and the thesis author is the primary author of the this material.

Chapter 6 details the implementation of the proposed SVM algorithm on the MoRES dataset, which comprises both individuals with normal gait patterns and those exhibiting abnormal patterns. The criticality analysis features of the dataset

were obtained from the RCC-optimised Berry model [29] and classified into normal and abnormal groups across five individuals. To evaluate the efficacy of the developed models, the confusion matrix, receiver operating characteristics (ROC), area under the ROC curve, and mean square error (MSE) tests were employed. Additionally, the KNN technique was utilised to train the dataset features, with the cross-validation test serving to validate the model's performance. Based on the simulation results, the Kernelised SVM method exhibited greater accuracy than the KNN model.

The majority of this chapter features previously published material, which appeared in the MDPI Computers journal as part of a special issue titled "Advances of Machine and Deep Learning in the Health Domain," and in the IEEE Symposium on Computers and Communications, Workshop on ICT Solutions for eHealth (ICTS4eHealth) (ICTS4eHealth2021) as a conference paper. The author of this thesis is the principal author of the material.

Chapter 7 presents a detailed analysis of the implementation of a proposed SVM algorithm on the Museum dataset. The Museum dataset contains gait patterns from two gender groups across five different age categories. The criticality analysis features of this dataset were derived from the RCC-optimised Berry model [29], with unique features extracted for each age category. The proposed models were evaluated for their robustness using tests such as the confusion matrix, receiver operating characteristics (ROC), area under the ROC curve, and mean square error (MSE). The simulation results indicate that the Kernelised SVM approach generates accurate results, making it a reliable tool for classifying unique dynamics of gait disruptions.

The contents of this chapter largely overlap with the material that will appear in an upcoming publication in the IEEE Transactions on Pattern Analysis and Machine Intelligence, of which the thesis author is the principal contributor.

Chapter 8 provides a summary of the accomplishments of the thesis and offers suggestions for future research directions.

Chapter 2

Literature Survey on Gait Analysis

2.1 Introduction

Gait analysis is a method used to study human walking in order to understand and diagnose a wide range of conditions and disorders. The study of gait is important as it can provide insight into an individual's overall health and functional abilities. Gait analysis is commonly used to assess patients with conditions such as Parkinson's disease, stroke, and osteoarthritis, as well as to study the effects of aging and rehabilitation on gait [12]. In recent years, there have been significant advancements in the technologies and techniques used for gait analysis. The state-of-the-art in gait analysis includes the use of motion capture systems, wearable sensors, and computational models. Furthermore, machine learning techniques, particularly Support Vector Machines, have been increasingly used in gait analysis with demonstrated potential in detecting and analysing gait disorders, predicting fall risks, and so on. The purpose of this literature review is to provide an overview of the state of the art in gait analysis, to review the use of machine learning for analysing gait and to highlight the limitations in previously reported work, and to review specific applications of Support Vector Machines for gait analysis.

2.2 Overview to Machine Learning

Artificial intelligence (AI) and machine learning (ML) are now being researched numerously in medicine, with the potential to automate human activities and potentially accomplish tasks beyond human capabilities [58].

Artificial intelligence (AI) is the processing of data by machines to achieve goals that normally require human cognitive function. Recently, AI has processed massive amounts of data and offered the opportunity to detect patterns that may be difficult to classify. It also provided unique and novel data representation techniques to handle different types of complex data patterns. Machine learning (ML) is an AI-subfield that uses algorithms to train and learn patterns emprically from data. ML broadens the domain of the conventional rules of statistics due to the ability it has to map the multivariate and nonlinear high-dimensional data patterns that may appear challenging to conventional methods of statistics.

Deep Learning (DL), which has recently appeared as a robust ML method, enables efficient processing of dense datasets that require a huge amount of computational power to facilitate better decisions on complex data. Recent advances in ML and DL in a wide range of fields, including natural language processing, computer vision, engineering, medicine, and other scientific sectors, have raised significant awareness across a wide range of disciplines [59].

The use of AI in relation to human gait analysis is not a novel notion. However, the development of automation systems with well-designed ML algorithms and hardware capabilites makes the role of AI in gait analysis attractive to medicine and public health in terms of early disease detection and improving disorder prognosis prediction [60].

With the support of state-of-the-art examples that are related to human gait analysis and automated recognition, a general overview of the technical concepts of artificial intelligence and machine learning will be presented, as well as the types of

machine learning techniques will be introduced. Moreover, this chapter will restrict the discussion only to two supervised machine learning approaches that have been used in the implementation analysis of the three gait datasets analysed in Chapter 5, 6, and 7, respectively. The mathematical intuition behind each technique will be described. Prior studies on how machine learning research has impacted the direction of research in human gait classification and detection will be explored. Furthermore, this chapter will discuss the implementation procedures of the proposed SVM algorithm and how the SVM model adopts the Kernel Trick feature [46] to solve the nonlinearity problem of human gait data. Finally, this chapter will emphasise the critical factors to consider while evaluating the model's performance.

2.3 Types of Machine Learning Algorithms

In order for AI to analyse data, input data must be arranged as feature vectors. The features are measurable attributes of the input data, which can be formatted into feature vectors in order to be mathematically described and facilitate the computation process of ML models [44]. The features can be a form of a variety of types of data and are mostly extracted from the raw data, such as sensory variables of human gait that are collected by the inertia moment unit (IMU). In conventional ML methods that do not use advanced approaches to data representation, the relevant data features can be selected manually or by other dimensionality reduction methods. Meanwhile, the developed CA approach is capable of automatically representing the entire or relevant features from the original (collected or stored) data.

Machine learning may be categorised into four major groups based on model properties and input data [61]:

1. **Supervised Machine Learning:**

   The main role of supervised machine learning is to train a model that associates input data, which is represented by features, with known labelled outputs of

interest. Predicting labels on new data is referred to as classification, which can either be binary (for instance, discriminating between normal and abnormal human gait) or multiclass (for example, distinguishing between the type of gait; normal versus gender-based versus unique age group, etc.), whereas predicting a continuous resultant is referred to as regression (such as, predicting the likelihood of having healthy walking behaviour after serious gait injury) [62].

The process of supervised machine learning is based on providing the training data, which is composed of feature vectors and output labels, to a machine learning classifier of interest. Then, the ML classifier creates a ML model by learning through continuous iterations to correlate the input feature data to the output labels. Once trained, the testing process starts, where the ML model makes predictions based on new data.

There have been various supervised learning approaches developed throughout the years that may identify correlations between input data and observation outputs in a linear or nonlinear fashion. Beyond model selection, the most significant aspect of model success in supervised machine learning is appropriate feature engineering selection to reflect on the overall model's performance. There are several typical algorithms which are related to supervised learning, such as (1) *Logistic Regression* [63], as it does classification rather than regression. However, it does not automatically efficiently handle interaction or nonlinearity of data. (2) *Naïve Bayes* [64], which is based on Baye's theorem of conditional probabilities. It is fast and only accommodates a limited number of data samples, but may be difficult to interpret and presume independance between data variables. (3) *Random Forest* [65], which integrates the results of multiple decision trees to produce a single outcome. Even with a high number of variables, it is characterised by its capability of detecting interactions between data variables. Also, it is resilient to overfitting and noisy data. The

other two algorithms (*K-Nearest Neighbour* and *Support Vector Machine*) [61], which are used in this research for classification purposes, are described in detail in Section 4.2 and 4.3 respectively.

Due to the difficulty of deciding what type of supervised ML model would perform best, the model itself undergoes development and optimisation procedures, including selecting the learning algorithm of interest, optimal feature set, and most importantly, fine-tuning the hyperparameters of the selected learning method. Several methods exist for evaluating model performance during the model development process [66]. Classification algorithms generate a continuous output, instead of a binary or discrete outcome. The Receiver Operating Characteristics (ROC) curve is utilised as a metric to evaluate the area under the curve (AUC), depending on the ML model outcome [67]. The confusion matrix is an alternative performance measure metric that can be used at specific operating points on the receiver operating characteristic curve to compute the model's accuracy, sensitivity (sometimes referred to as recall), specificity, precision (also referred to as positive predictive value), negative predictive value, and the F_1-score [68]. On the other hand, the mean square error (MSE) or root mean square error (RMSE) are commonly used to evaluate the performance of regression models [69].

Data partitioning as a method of model optimization must be well-engineered, depending on the functionality of the applied model (whether classification or regression). As an unbiased data partitioning technique, the $K-$fold cross-validation method is used to assess the effect of various ML model parameters during model training iteratively over several K values in order to produce an optimal model [70]. This method is commonly used in KNN machine learning algorithm. In contrast to the cross-validation approach, which increases computation time, there is another type of data partitioning strategy that divides a dataset into three parts: a training set, a testing set, and a hold-out validation

set [71]. The training set is used to fit the model, while the validation test evaluates the model adjustments. Model improvements can be achieved on the testing set to create an unbiased optimal model.

2. **Unsupervised Machine Learning:** Unsupervised machine learning techniques are relatively uncommon in comparison with their supervised learning counterparts. It usually measures patterns within input data that do not have identifiable labels of interest [72]. Many methods have been developed for unsupervised learning, such as: (1) *Clustering Analysis* [73], which is frequently used to find subgroups in large amounts of complex dataset. (2) *Hierarchical Agglomerative Clustering* [74], which creates small clusters of data whose patterns are similar to each other, and then joins these clusters together to form large clusters of similar patterns. (3) *K-means Clustering* [75] is an alternative to hierarchical clustering in which the number of K clusters within the data is pre-defined, allowing data samples close to the cluster's centroid to have the identity of that cluster. The authors in [76] used Sagittal plane gait features as input to the K-means clustering approach to identify homogeneous groups of individuals with cerebral palsy impairment who walk with excessive knee flexion at initial contact as an example of this approach. (4) *Dimensionality Reduction* [77], which used to rank the dimension of input data into lower dimensions to facilitate data clustering. The criticality analysis methodology developed in this study based on the RCC approach can be considered a type of dimensionality reduction technique as it maps the original data patterns into a new lower-dimension space. The criticality analysis method can transform unlabeled data patterns into new identifiable or labelled features. In terms of the challenges of unsupervised machine learning, it requires a massive amount of computational power and thus excessive time to decide patterns within data, and as a result, it may produce nebulous features that may not be relevant to the patterns of interest. Also, choosing the number of clusters in various

clustering analysis methods provides a high degree of ambiguity as it must be premeasured and also depends on the data patterns. For example, although the authors in [78] applied the K-means clustering analysis to classify flexible flatfeet based on 3D foot kinematics during walking for a number of patients, they ran an interative test to select the best or optimal number of clusters and chose two clusters out of six that showed a minimal reduction in mean square error rate.

3. **Semi-supervised Machine Learning:** This kind of learning combines elements of both supervised and unsupervised learning in which some data patterns can be labelled while others do not contain labels [79]. It uses the knowledge contained in a small amount of labelled data to produce a labelled outcome from a large amount of unlabeled data patterns. If the training procedure revealed that there were no relevant labels that could be used for class labelling, the model may be overstressed and experience time-complexity issues. As an example of this approach, the authors in [80] demonstrated a personalised classifer to detect human motion based on this semi-supervised learning concept.

4. **Reinforcement Learning:** It is sometimes referred to as "learning with criticism" since it gives computers and agents the capacity to autonomously make the optimal judgments within their environment [81]. It varies from other supervised machine learning approaches in that there are no accurate input or output sets, which means that when the algorithm knows the answer is incorrect, it makes no modifications to rectify it.

2.4 State of the Art on Gait Analysis

Cutting-edge gait analysis employs diverse technologies such as motion capture systems, wearable sensors, and computational models to enable informed decision-

making across various domains, including healthcare, rehabilitation, sports, and security.

2.4.1 Smart Motion Capture Systems

Smart motion capture systems are an essential technology used in gait analysis. They consist of a series of cameras and sensing markers that are used to track the movement of the body and to facilitate health monitoring and control, providing highly accurate data on gait kinematics [82]. They are typically used in controlled laboratory environments, and have been used to study a wide range of gait-related conditions. A variety of wearable systems, including shoes [83], socks [84], kneepads and anklets [85], insoles [86], smartphones [87], and smartwatches [88], are being developed to gather gait data for passive mental health assessment and for use in health, sports, security, and entertainment applications. Additionally, implantable medical devices and wearable robotics such as prosthetics and orthotics are also being developed. An example of this is the Ubiquitous Healthcare for Elderly (UbiHeld) system, which uses gait data from a smartphone and a Kinect camera to monitor the health, location, and other activities of elderly individuals at home [89]. The creation and advancement of smart gait devices poses a variety of difficulties, such as the need for a balance between elasticity and ergonomic design, as well as aesthetic appeal, compactness, and flexibility [90]. Additionally, these devices must be able to withstand frequent washing and maintain their robustness and reliability over time. Person identification is a crucial aspect of these devices, requiring a balance between personalised user experience and user privacy. Ensuring seamless, ongoing user authentication is crucial to protect against unauthorised access to medical records [91]. Additionally, these devices lack additional features that are essential for their intended use, including water resistance, durability, and the capability to connect with other smart devices and environments [83]. Several investigations have mainly cen-

tred on the utilisation of sensing technology to fulfil these specifications. Wearable devices utilise a variety of sensors including Inertial Measurement Units (IMUs), capacitive sensors [92], Kinetic Energy Harvesters (KEH) [93], solar cells [94], resistive sensors and stretchable conductive micro-fluids [86], and Triboelectric Nanogenerators (TENGs) [95]. These sensors are engineered to be compact, lightweight, and resilient, while simultaneously delivering precise and dependable data. Research has shown the effectiveness of using AI, sensors, and related technologies in predicting movement intentions and recognizing human activities through the use of smart gait devices. As examples, Yang and Yin (2021) [90] employed a Long Short-Term Memory (LSTM) method in conjunction with a Convolutional Autoencoder (CAE) to estimate joint torque using soft pneumatic sensors and 3D Inertial Measurement Units (IMUs) in smart shoes. Similarly, Xu et al. (2021) [91] utilised an attention-based LSTM for gait recognition while preserving the privacy of users through the implementation of KEH PrivGait, a wearable device equipped with Kinematic and Environmental Hybrid sensors. Sandhu et al. (2021) [94] used Random Forest (RF) for human activity recognition in a wrist-worn device, known as SolAR, which is self-powered through solar energy. Also, Zhang et al. (2020) [95] applied a 1D Convolutional Neural Network (CNN) for gait and human activity recognition in smart textile-based socks for long-term monitoring. Lan et al. (2020) [92] employed Naive Bayes (NB), Random Forest (RF), Decision Tree (DT), and K-Nearest Neighbours (KNN) for human activity recognition in a self-powered shoe with embedded Capacitive Sensing technology. Gao et al. (2020) [96] utilised Recurrent Neural Networks (RNN) for imitation learning in real-time prosthetic control with built-in motion sensors in a transfemoral prosthesis. Zhang et al. (2020) [97] used Reinforcement Learning (RL) for assist-as-needed control in robot-assisted gait training with motion sensors in an SAFE orthosis. Llorente-Vidrio et al. (2020) [98] applied a Deep Dual Neural Network (DDNN) for the classification of Electromyography (EMG) signals to activate an event-driven controller with electromyography (EMG) sensors in a

mobile lower limb active orthosis.

2.4.2 Smart Homes

Another technology that is commonly used in gait analysis is smart homes, such as houses or residential environments that have been outfitted with technology that allows for the monitoring and analysis of human gait patterns. This technology typically includes a variety of sensors, cameras, and other devices that are strategically placed throughout the home to track and record individuals' movement and gait patterns [99]. The data collected by these sensors and cameras is then analysed by specialised software to provide insights into a person's gait, including information on gait kinematics, speed, and symmetry. This information can be used to identify and track gait-related conditions such as diseases or disorders, as well as to monitor progress in rehabilitation and gait training programs. Smart homes for human gait analysis can also include devices like smart shoes, smart socks, and other wearable technology that can track gait and other movement data. The scientific community has extensively explored the use of AI algorithms in smart home and wearable technologies. As an illustration, Borelli et al. (2019) [100] implemented AI algorithms into smart objects such as a wall light for indoor localization, an armchair for sitting posture monitoring, and a belt for capturing movement information. The focus of their research was on the techniques, structure, planning, and creation of these intelligent devices, with the wall light additionally providing input to a fall detection algorithm, where the wall panel and mobile devices acted as the user interface. According to Hsu et al. (2017) [101] study, 3D gesture recognition had a 95.3% accuracy using Probabilistic Neural Network (PNN) and a 10-fold Cross Validation (CV). The study also focused on pedestrian navigation, home safety and fire detection using a wearable IMU on the wrist and feet, and environmental sensors. The smart home testbed included a web camera and multisensory circuit module for these tasks. In

Cook et al. (2015) study [99], various AI algorithms such as DT, NB, RF, SVM, Adaptive DT, and Adaptive RF were tried out, with Adaptive DT providing the best classification accuracy. Their study used a CASAS smart home and wearable sensors for in-home health monitoring for early detection of changes associated with PD and Mild Cognitive Impairment (MCI), and evaluation of treatment. Principal Component Analysis (PCA) was used to reduce features and k-Means clustering and random resampling were used to add features in smaller datasets of individual activities.

2.4.3 Advanced Technology Using Virtual and Animation Environments

Another technology for analysing human gait is that uses virtual and animation environments to analyse and understand human walking patterns. This technology typically employs virtual reality, computer animation, and motion capture techniques to simulate and record human gait movements [102]. This can be used for medical and rehabilitation purposes, such as analysing and diagnosing gait disorders, or for research and development in the field of biomechanics. This technology can also be used for ergonomic evaluations, for example, for evaluating the design of footwear or prosthetic devices [103]. Multiple AI algorithms are utilised for character control and motion reconstruction in video games and virtual reality applications. In recent years, various AI algorithms have been employed for motion reconstruction and character control in video games and VR applications. For example, a study by Feigl et al. (2020) [104] utilised head-mounted accelerometer data to reconstruct motion using Threshold Based Method (THR), Pearson Correlation-based Method (COR), SVM, and Bidirectional LSTM (BiLSTM). The results of the study found that the COR method had the highest accuracy for real-time VR applications with low delay. Another study by Bergamin et al. (2019) [105] used unstructured motion data from

a motion capture system to create real-time physics-based character control for video games using Data-Driven Responsive Control (DReCon), which is a motion matching and deep RL technique. In a similar vein, Peng et al. (2018b) [106] used simulated character models and YouTube video clips to learn dynamic physics-based character controllers using OpenPose/Human Mesh Recovery (HMR) and DRL. Peng et al. (2018a) [107] also used a character model and kinematic reference motion from video clips to create physics-based character controllers using DeepMimic, another DRL technique. Additionally, Huang et al. (2018) [108] used a Skinned Multi-Person Linear (SMPL) body model and BiLSTM with 6 IMUs to reconstruct 3D human poses from a sparse set of IMUs. Holden et al. (2016) [109] used the Carnegie Mellon University (CMU) Motion Capture Database to learn an unsupervised human motion manifold using CAE. Huang et al. (2015) [110] used a Surface Motion Graphs (SMG) and part-based Laplacian deformation with three 4D Performance Capture (4DPC) datasets to create a data-driven approach for animating 4DPC character models. Ding and Fan (2015) [111] used multilayer Joint Gait-Pose Manifolds (JGPMs) and topologically constrained Gaussian Process (GP) Latent Variable Models (GPLVMs) with the CMU Motion Capture Database and simulated data to model human gait. Finally, Alvarez-Alvarez et al. (2012) [112] used Fuzzy Finite State Machines (FFSM) with automatic learning of the fuzzy knowledge base by Genetic Algorithm (GA) to model human gait using an accelerometer attached to a belt. Overall, these studies demonstrate the ongoing effort to develop sophisticated AI algorithms for motion reconstruction and character control in video games and VR applications.

2.5 Machine Learning for Gait Analysis: A Critical Review of Techniques and Limitations

Recently, the utilisation of Machine Learning (ML) techniques in gait analysis has seen a significant increase for the purpose of classifying gait patterns and predicting gait-related outcomes. One of the key benefits of ML techniques is their ability to automatically extract relevant features from raw spatiotemporal gait data, thus enhancing the accuracy of gait analysis. While there are several data preprocessing techniques employed, such as Ground Reaction Force (GRF) filtering, time derivative, time normalisation, data reduction, weight normalisation, and data scaling. Burdack et al. (2020) [113] found that only GRF filtering and supervised data reduction methods, such as PCA, improved the performance of ML classifiers. Additionally, the study revealed that Random Forest (RF) was more robust in feature reduction when compared to Support Vector Machines (SVM), Multi-Layer Perceptron (MLP), and Convolutional Neural Networks (CNN). However, the research falls short in taking into account the representation of all data features in a comprehensible manner, which is crucial for achieving the highest possible accuracy, as is the case with the CA technique discussed in this thesis. ML techniques play a vital role in various gait analysis applications, such as prediction of gait events and classification of gait phases, health, and gender recognition.

2.5.1 Recognition of Gait Phase and Event Prediction

The way in which a human moves, known as gait, is a continuous sequence of movements or cycles. Each gait cycle refers to the movement of a single limb from heel-strike of one foot to heel-strike of the same foot again Gage et al. (1995) [114]. The two primary gait phases are the stance phase and the swing phase. Depending on the intended use of the gait analysis, identifying these two phases alone may be

sufficient. This simplification allows for a more straightforward and cost-effective gait analysis, particularly in wearable systems Di Nardo et al. (2020) [115]. However, a more common four-phase cycle includes initial contact, mid-stance, pre-swing, and swing Jiang et al. (2018) [116]. The integration of Artificial Intelligence (AI) in these studies is essential for real-time gait analysis, which is highly valuable in various control devices such as orthotics and prosthetics, rehabilitation monitoring, and fall detection systems for aging-in-place applications. There are various studies on gait phase recognition and gait event prediction. As an illustration, Vaith et al. (2020) [117] used LSTM-Net and DENSE-Net with 12 subjects and 7 IMUs, 2 IMU pressure insoles for offline AL to reduce labelling cost and online gait phase classification. Pérez-Ibarra et al. (2020) [118] used a Hybrid Simulated Annealing (SA) and Genetic Algorithm (GA) approach with 3 subjects, 1 HC, 2 impaired gaits, and an IMU at the back of the heel for online gait event detection. Di Nardo et al. (2020) [115] employed a MLP with 23 subjects and 1 electrogoniometer per leg for gait phase classification, 2 phases (stance/swing). Morbidoni et al. (2019) [119] used MLP with 23 healthy subjects, Surface Electromyography (sEMG) and barographic data, natural walking conditions for gait phase classification, 2 phases: stance/swing and gait event prediction, Heel-Strike (HS) and Toe-Off (TO). Jiang et al. (2018) [116] used LDA with 9 healthy subjects, TW, 8 pressure sensors in an ankle-worn band for wearable gait phase recognition system. Farah et al. (2017) [120] used DT, RF, MLP, and SVM with 31 subjects and an inertial sensor at the thigh for gait event detection. Mannini et al. (2014) [121] used HMM with Short-Time Viterbi (STV) with 9 healthy subjects, Treadmill Walking (TW), and 1 IMU gyro at the instep of the left foot for online gait event detection.

2.5.2 Detection of Abnormal Gait Patterns

The application of Machine Learning (ML) in gait analysis for health is wide-ranging and includes the detection of abnormal gait patterns (fatigued, and neurological gait) caused by conditions or diseases, sports performance management, fall detection, and recognition of activities. Several studies on fatigued gait were conducted. For instance, Russell et al. (2021) [122] used CNN with 97.8% accuracy and data collected by an accelerometer worn around the chest and a GPS watch for location tracking, to detect variations in terrain and fatigue. Baghdadi et al. (2021) [123] used Multivariate Hierarchical Time Series Clustering Algorithm (MHTSCA) with DTW as a dissimilarity measure and data collected by an IMU worn at the right ankle, with 15 subjects to detect fatigue development over time. Sedighi Maman et al. (2020) [124] used RF with Best Subset Selection (BSS) and 85.5% accuracy, with data collected by one sensor in the torso, with 15 subjects to develop a 4-phase fatigue management framework in the workplace (detection, identification, diagnosis, and recovery). Karvekar (2019) [125] used 2-class SVM with 91% accuracy, 24 subjects, and a smartphone attached to the shank to detect fatigue: baseline, low, medium, and strong fatigued states, 3-class SVM with 76% accuracy, and 4-class SVM with 61% accuracy. Baghdadi et al. (2018) [126] used SVM with 90% accuracy, data collected by one IMU in the ankle, and 20 subjects to detect fatigue after Manual Material Handling (MMH) tasks. Janssen et al. (2011) [127] used SVM and SOM with PCA with 98.1% accuracy, data collected by 9 subjects' GRFs, to classify inter and intra-personal gait before, during, and after leg exhaustion.

2.5.3 Monitoring of Neurological Diseases

Other studies in the field of Machine Learning (ML) have investigated the use of ML techniques for the diagnosis and monitoring of neurological diseases, with the goal of improving detection and therapy. There have been numerous studies that

have employed machine learning techniques to identify pathological gait patterns. For instance, Lu et al. (2020) [128] utilised a support vector machine with principal component analysis on a dataset collected from a Kinect camera with image rectification, resulting in an 88.89% accuracy in the detection of depression. Similarly, Khodabandehloo et al. (2021) [129] applied HealthXAI Classification and Regression Trees (CART) on Partial Center for Advanced Studies in Adaptive Systems (CASAS) dataset, including individuals with Parkinson's disease and mild cognitive impairment, to determine a numerical score and explanation of cognitive decline in the elderly. Iosa et al. (2021) [130] applied an artificial neural network (ANN) on a dataset collected from an inertial measurement unit to develop a stroke prognostic tool that predicts the ability of an individual to return to work, achieving an accuracy of 93.9%. Zhu et al. (2020) [131] used RF with IAFSA on a dataset of 3 patients with knee replacement to predict knee joint impairment. Zhou et al. (2020) [132] used kernel PCA with SVM, RF, and ANN on a dataset of 239 subjects, including young and old healthy individuals and those with geriatric conditions, achieving an accuracy of 90%. Zeng et al. (2020) [133] used an RBF neural network with Deterministic Learning (DL) for kinematic modelling and classification of Chronic Unilateral ACL deficiency. Pepa et al. (2020) [134] using a novel FL on smartphone data for real-time, interpretable Freezing of Gait (FoG) detection. Kaur et al. (2020) [135] used LR, SVM, and RF on GRFs from instrumented treadmill for classification of Multiple Sclerosi (MS) mild and moderate. Guo et al. (2019) [136] used SVM and BiLSTM on a light-weight telepresence robot equipped with a single RGB-D camera to classify normal, in-toeing, out-toeing, and drop-foot gait with no additional sensing feedback. Sato et al. (2019) [137] used ST-ACF DTW and KNN with OpenPose on a CASIA-B dataset of frontal videos of two Parkinson's disease (PD) patients to quantify normal and Parkinsonian gait features from home movies. Fang et al. (2019) [138] used RF on a dataset of 95 graduate students, 52 with depression and 43 healthy controls, to analyse depression through gait patterns.

Acosta-Escalante et al. (2018) [139] used Logitboost and RF on a dataset of 14 Huntington's disease (HD) patients and 7 healthy controls, using smart phones affixed to both ankles, to classify HD gait patterns. Pulido-Valdeolivas et al. (2018) [140] used RF with DTW based on an optokinetic IGA system to monitor the progression of HSP (Hereditary Spastic Paraplegia) and personalise therapies. Wan et al. (2018b) [141] used DMLP (Deep Multi-Layer Perceptron) and a dataset of 50 individuals wearing a phone on the waist to analyse speech and movement data captured by a smartphone to estimate the severity of Parkinson's disease. Similarly, Arifoglu and Bouchachia (2017) [142] used LSTM on a public dataset collected through environmental sensors to perform human activity recognition and activity of daily living for elderly people with dementia. Dolatabadi et al. (2017) [143] used Gaussian Process Latent Variable Models (GPLVM) and KNN-DTW to discriminate between healthy and pathological gait patterns because of stroke or ABI, with an F1-score of greater than 0.94. Shetty and Rao (2016) [144] used SVM with Gaussian Radial Basis Function (RBF) to distinguish PD gait from HD, ALS, and HC, achieving an accuracy of 83.3%.

2.5.4 Gender Recognition

Gender recognition based on human gait is a method of identifying the gender of an individual based on characteristics of their walk. This can be done using computer vision and machine learning techniques, which have been widely used in recent years for a variety of applications including surveillance, human-computer interaction, and biometrics [152]. The human gait is a complex phenomenon that is influenced by various factors such as body structure, muscle strength, and coordination. There are specific gait characteristics that can be used to differentiate between men and women, such as the length of steps, pace, posture, and the position of joints relative to certain planes, known as swing energy [153]. Specific techniques, such as the his-

togram of gradient method [154], can be used to extract features from video, depth, or sensor data obtained from cameras. These features can then be converted into 1D temporal descriptors for use in training machine learning models, using data from 3D IMU sensors. The primary difficulty in identifying gender through gait analysis is the variability in walking speed and other related factors such as clothing, footwear, and carrying objects [155]. This, combined with the limitations of the size of the training set and the complexity of the model, can greatly impact the gait parameters and make it difficult to obtain accurate features, resulting in poor classification performance [156]. Therefore, it is crucial to be aware of the limitations of this method and not rely solely on it without additional information and evidence, as the recognition may not always be reliable and may not always produce correct results. Additionally, it is imperative to take into account the ethical and privacy implications of using this technology as it could potentially be used to discriminate or profile individuals based on their gender [157]. Therefore, further research is necessary to improve the performance and reliability of gender recognition based on human gait while addressing the ethical and privacy concerns associated with its use. A number of studies have been conducted in recent years to evaluate the performance of gender recognition based on human gait, using various datasets and methods. The results have been mixed, with some studies reporting high accuracy rates, while others have reported lower rates. This can be attributed to the complexity and variability of the data, as well as the limitations of the models and the methods used. As an illustration, Castro et al. (2017) [158] proposed an end-to-end approach using a CNN, which utilises gait signature representation of individuals extracted from low-resolution video using optical flow maps to achieve high accuracy in recognising both gender and identity. Zhang S. et al., (2019) [159] demonstrated that utilising a multi-task convolutional neural network (CNN) architecture, in which the network learns multiple attributes simultaneously, can significantly improve the overall accuracy of the model. Furthermore, Kwon et al. (2021) [160] reported an accuracy of 100% using a group of four machine

learning techniques (KNN, SVM, NB, and DT) on UPCVgaitK1 and UPCVgaitK2 dataset for gait recognition. Various researchers such as Lu et al., (2014) [161] in the field of ML have also focused on recognising gender from gait data using a variety of techniques such as the combination of AP clustering and Sparse Reconstruction-based Metric Learning (SPML) with an accuracy of 87.6%. Despite the advantages of ML techniques, there are also limitations to their use in gait analysis [162]. One limitation of previously reported work in this area is that many studies have focused on small datasets or have used unrealistic or simplified gait models. This has limited the generalizability of the results and the ability to apply the findings to real-world situations. Additionally, while many studies have reported high accuracy rates for ML-based gait analysis, these results may not be generalizable to other datasets or to real-world scenarios. Another limitation is that ML algorithm use for Gait analysis is often data heavy which leads to a lack of robustness to changes in the environment, making it difficult to apply to real world scenarios [163]. The dataset used for training and testing the model should be diverse and inclusive to improve the performance of the model in real world situations.

2.6 A Review of Support Vector Machine-based Approaches in Human Gait Analysis with a Focus on Criticality Analysis

The field of human gait analysis has seen significant advancements in recent years, yet the application of criticality analysis and rate control of chaos remains an under-explored area of research, particularly in human gait analysis. However, the utilisation of Support Vector Machines (SVMs) as a tool for gait analysis has been well-established in the literature. Although, most of the Machine Learning (ML) methods and techniques applied to human gait analysis have focused on utilising

gait features or patterns without fully understanding their biological significance. These studies often employ ML approaches to gait data without considering the underlying complexities of human gait. But the use of criticality analysis presents an opportunity to unlock the understanding of these complexities by creating spatiotemporal patterns of gait patterns in a more interpretable space. This approach allows for a deeper understanding of the dynamics of human gait. In this section, we aim to review the existing literature on the use of SVMs in gait analysis, and its potential applications in understanding the dynamics of human gait, particularly in the light of criticality analysis. The objective of this review is to provide a thorough examination of the state-of-the-art in this field and to serve as a foundation for future research endeavours. One of the studies, which is conducted by Wonjin et al. (2021) [164] on gait analysis for detection of initial characteristics of gait disorders is a valuable contribution to the field of gait analysis and neurological disorder diagnosis. The use of the SVM method to classify abnormal gait from a walking person, as well as the integration of gait features extracted from the individual's walking movement using the Kinect depth camera, are both strong points of the study. The 96.52% classification accuracy achieved in the study is also noteworthy and suggests that the proposed method has a high level of performance. However, it is important to note that this study has several limitations. The sample size of the study is relatively small, which can limit the generalisability of the results to a larger population. Additionally, the study only uses one type of depth camera and it would be beneficial to test their proposed method using other types of sensors to ensure the robustness of their proposed method. Furthermore, the study does not provide any information on the cost and practicality of their proposed method. The use of a depth camera and the required computational resources may make their proposed method impractical for use in a clinical setting. It would be useful to include an analysis of the cost-effectiveness of the proposed method in future studies. Another study conducted by Iris et al. (2010) [165] presented to detect and characterise gait abnormalities in individuals

with Parkinson's disease (PD) using wireless inertial sensor system measurements. The study utilises the physical features of pitch, roll, and yaw rotations of the foot during walking and applies the Principal Component Analysis (PCA) to select features and the SVM method to create a classification model. The study found that the model performed with over 93% sensitivity and specificity and 97.7% precision in the binary classification task of detecting the presence of PD. This suggests that the wireless inertial sensor system was able to successfully detect the presence of PD based on physical features of gait and identify the specific features that characterise parkinsonian gait. Furthermore, the use of a cost-sensitive learner in the study reflected the different costs associated with misclassifying PD and control subjects, which led to 100% specificity and precision while maintaining sensitivity of close to 89The study also performed a multi-class classification task of characterising parkinsonian gait by distinguishing among PD with significant gait disturbance, PD with no significant gait disturbance, and control subjects. This resulted in 91.7% class recall for control subjects and 84.6% precision for PD subjects with significant gait disturbance. Furthermore, Nukala et al. (2015) [166] and Shibuya et al. (2015) [167] used wireless gait analysis sensors (WGAS) to perform automatic fall detection and have similar sensor designs, including a tri-axial accelerometer, 2 gyroscopes, and a micro-controller. Both studies also used two supervised machine learning techniques such as Back Propagation Artificial Neural Network (BP-ANN) and SVM for fall classification. However, there were some differences in the methodology used in these studies. Nukala et al. (2015) [166] used a total of 322 tests on young volunteers and achieved an overall accuracy of 98.2% and 98.7% with BP-ANN. On the other hand, Shibuya et al. (2015) [167] utilised SVM classifier to detect falls in real-time and achieved high fall classification accuracy (98.8% and 98.7% for different sensor positions) and high overall specificity (99.5%) and sensitivity (97.0%). However, the study conducted by Nukala et al. (2015) [166] did not provide information about the specific parameters used for the algorithm, which would have been useful to under-

stand the results and compare them to other studies. Additionally, the study only reported the accuracy of the algorithm, without providing information about other performance metrics such as precision, recall, or F1-score. On the other hand, the study by Shibuya et al. (2015) [167] reported high fall classification accuracy (98.8% and 98.7% for different sensor positions) and high overall specificity (99.5%) and sensitivity (97.0%). The study also provided information about the specific parameters used for the algorithm, which is useful to understand the results and compare them to other studies. Additionally, the study examined the performance of the algorithm in real-time, which is important in practical applications. In a similar vein, Huang et al. (2018) [168] presented a novel approach to gait analysis by using audio data instead of visual data. This is a significant advancement in the field as it addresses the limitations of traditional gait analysis methods, which rely on visual data and are affected by changes in clothing, visibility, and angles. The use of audio data in gait analysis is an innovative approach as it is not affected by the same limitations as visual data. The technique is based on the time differences between steps rather than frequency-based features, which are affected by changes in footwear and floor surfaces. This allows for a more robust and accurate analysis of gait. The authors used SVM technique for classification. The results showed high classification rates and excellent discriminative abilities, indicating that the method is effective in identifying individuals based on their gait. Additionally, Hayashi et al. (2015) [169] study introduced to improve the diagnostic accuracy of lumbar spinal canal stenosis (LSS) by using gait analysis as a classification method. The study group consisted of 13 healthy individuals, 11 patients with L4 radiculopathy, and 22 patients with L5 radiculopathy. The authors used video recordings and a development program to analyse gait characteristics, and an SVM to classify L4 and L5 radiculopathy. The study found that knee extension at initial contact was slightly greater in the L4 group and a one-peak waveform pattern with the disappearance of the second peak was present in 45.5% of the L5 group. The total classification accuracy was

80.4% using the SVM, with the highest accuracy in the control group and the lowest in the L4 group. The authors concluded that their walking motion analysis system was able to identify useful factors for differentiating between healthy individuals and patients with L4 and L5 radiculopathy with a high accuracy rate. The study has several strengths, such as the use of a well-established method for gait analysis and the use of a development program to analyse gait characteristics. Additionally, the use of an SVM to classify L4 and L5 radiculopathy is a valid method and the results of the study demonstrated that this method can be useful in identifying patients with LSS. However, the study also has some limitations. Firstly, the sample size is small, which may limit the generalisability of the findings. Secondly, the study only analysed patients with L4 and L5 radiculopathy, and it is unclear if the findings would be generalisable to other types of radiculopathy or other types of spinal stenosis. Additionally, the study only used healthy individuals as the control group and it would be beneficial to also include patients with other conditions to compare the gait characteristics. Similarly, Yoo et al. (2005) [170] proposed an automated system for classifying gender by analysing human gait patterns. Their study provided valuable insights into the application of gait analysis for gender identification. The three-stage system proposed in the study is well-structured, with the detection and extraction of the moving human body and its contour from image sequences, extraction of human gait signature using joint angles and body points, and motion analysis and feature extraction for classifying gender, providing a clear and logical progression in the analysis. The use of 2D stick figures to represent the gait signature was found to be an effective approach, as it allows for the clear visualisation of the gait patterns and allows for easy comparison between subjects. Additionally, the use of an SVM classifier to classify gender is a well-established method in machine learning, and its application in this study demonstrates the robustness and reliability of the system. The experiments performed on a large database showed a high performance of 96% for classifying gender for 100 subjects, indicating that the system is highly accurate

in identifying gender based on gait patterns. This is an important finding, as it has potential applications in areas such as security and surveillance, where automated gender identification is needed. However, the study has some limitations, such as the lack of diversity in the sample population, which could limit the generalisability of the findings. Additionally, the study did not address the potential ethical implications of using gait analysis for gender identification, such as privacy concerns and potential biases in the system. Besides, the research work presented by Begg et al. (2003) [171] is a valuable contribution to the field of gait analysis and falls prevention in the older population. The use of artificial intelligence techniques, specifically Neural Networks (NN) and SVM classifiers, for automatic identification of young and old gait types is a novel approach that has the potential to aid in early identification of at-risk gait for falls prevention. The study design is well-executed, with 12 young and 12 elderly participants being recorded and analysed using a motion analysis system and a force platform. The extraction of 24 gait parameters for training and testing the NN and SVM systems is a thorough approach that allows for a robust analysis of the data. The results of the study indicated that the SVM system has a better performance than NN with 91.7% and 83.3% respectively in distinguishing between young and elderly gait patterns. This suggested that the SVM is a more effective tool for identifying at-risk gait patterns in the older population. Additionally, the classification ability of SVM was found to be unaffected by the choice of kernel functions, further highlighting the potential of SVM for applications in gait identification for falls-risk minimisation in the elderly. However, the study is limited by the small sample size of participants, which may not be representative of the older population as a whole. Additionally, the study only included participants of a specific age range, and further research is needed to evaluate the applicability of their results to older adults outside of this range. Likewise, the work of Si et al. (2019) [172] on the development of a wearable sensing system for studying gait dynamics is a valuable contribution to the field of bio-mechanics and human movement analysis. The proposed system, con-

sisting of sensing shoes worn by a tester, effectively captures data on various types of movements, including standing, jumping, and walking. The use of five features extracted from foot pressure signals for motion analysis is a novel approach, and the application of SVM and fractal analysis for gait recognition is well thought-out. The testing results of the system, with an overall accuracy of 93.57% using radial basis function kernel function, were impressive and demonstrate the strong potential for gait identification using this method. The limitations of the study include the small sample size, lack of diversity in the population tested, which limits the generalisability of the results, and the lack of discussion on potential applications and how the system can be used to improve human movement understanding. Moreover, Kamruzzaman et al. (2006) [173] presented an interesting and valuable study on the use of SVM to classify children with cerebral palsy (CP) using gait parameters. The study makes use of two gait parameters, stride length and cadence, which are known to be affected in children with CP. The results of the study showed that the SVM classifier has an overall accuracy of 96.8% when normalised by leg length and age. This is a high accuracy rate and suggests that the SVM is a useful tool for identifying CP in children. One of the strengths of the study is that it compares the performance of different classifiers, including polynomial and radial basis kernel. The study found that these two classifiers performed comparably and outperformed the others. This suggests that the choice of kernel is not a significant factor in the performance of the SVM. This is an important finding as it can inform the selection of kernel in future studies. However, there are also some limitations to their study. Firstly, the sample size of the study is relatively small, with only 30 children with CP and 30 healthy children. This limits the generalisability of the results. Secondly, the study only used two gait parameters, stride length and cadence, which may not be sufficient to fully classify children with CP.

In addition, Horst et al. (2016) [174] conducted an examination to understand the nature of intrinsic inter-session variability in gait patterns. The use of 8 healthy

subjects who performed 15 gait trials at a self-selected speed on 8 different days within 2 weeks allows for a thorough analysis of the separable characteristics of gait patterns between and within individuals in repeated measurement sessions. The use of ground reaction forces and lower body kinematics as a means of quantifying each trial is a sound choice, as it allows for a detailed examination of the movement patterns of the subjects. The results of the study are analysed using an SVM classifier and the coefficient of multiple correlation approach, which are both appropriate methods for the type of data being analysed. The results showed that there are a remarkable amount of individual characteristics in human gait and that gait patterns can be assumed not to be constant over time, but rather exhibit discernible daily changes within previously stated good repeatability. This is an important finding, as it suggests that gait patterns are not always stable and may change depending on various factors such as physical or emotional state. The use of SVM results in an error-free assignment of gait patterns to the corresponding individual, which demonstrates the robustness of the method. Additionally, the classification rates of 97.3% and 59.5% for the eight-day classification of lower body joint angles and ground reaction forces, respectively, highlights the ability of the method to distinguish day-specific characteristics within the range of individual gait patterns. The potential implications of this study are that it may provide a basis for more individual and situational diagnoses or therapy for gait patterns. This could lead to more effective treatments for patients with gait disorders, as well as the ability to monitor changes in gait patterns over time.

2.7 An Overview of K-Nearest Neighbour (KNN) Approaches for Analysing Human Gait

The K-Nearest Neighbour (KNN) approach is a robust supervised machine learning technique that leverages data learning to effectively handle classification problems, achieving superior classification performance. KNN's strength as a reliable gait classifier lies in its ability to create strong prediction models. By computing the Euclidean distance in a Euclidean space and considering the majority vote of Neighbour ing labels, KNN determines the best k nearest Neighbour s. Previous research has explored various applications of the KNN method in gait classification. For example, in [188], researchers employed the KNN methodology to identify neurological disorders in the brain through distinct human gait features. Additionally, the KNN classification procedure was utilised to recognise individuals walking at the same speed in [189]. In [190], researchers proposed a KNN-based-rule to identify healthy gait patterns in Parkinson's disease patients. In [189], the KNN concept was applied to construct a secure authentication system that could recognise human cognitive processes based on the distance between right and left gait skeletal joints. Similarly, in [191], the KNN approach was used to classify biometric-related gait characteristics based on data collected from a wearable sensor device. The study utilised the best-matched gait signature metrics as a classification measure to identify individuals. Furthermore, in [192], the KNN classification rule was employed to categorise human activities based on their sitting, standing, running, and walking configurations. Likewise, in [193], Micro-Doppler data of gait characteristics were used for classification purposes. The authors in [194] discussed the use of the KNN technique to classify pedestrian motion based on the location of the inertial measurement units (IMUs) device on various portions of human joints, such as fixed hands, swinging hands, pockets, and backpacks. In addition, the study in [195] used a KNN classifier to recognise a variety of turning actions of a lower limb prosthesis utilising wearable

sensor data. One limitation of the KNN approach for gait analysis is its sensitivity to noise and outliers in the data. Since KNN relies on the distances between Neighbour ing data points to make predictions, noisy or outlier data points can greatly affect the accuracy of the classification. Another limitation is that the choice of the number of nearest Neighbour s, k, can have a significant impact on the performance of the algorithm. Choosing a value of k that is too small can result in overfitting, while choosing a value that is too large can result in underfitting. Finally, KNN can also be computationally expensive, especially when dealing with large datasets or high-dimensional feature spaces.

2.8 Summary

In conclusion, gait analysis is a multifaceted field that incorporates various technologies, techniques, and methods to study human walking. The state-of-the-art in gait analysis involves the use of a variety of technologies, including motion capture systems, wearable sensors, and computational models. These tools have allowed for a deeper understanding of human gait and its underlying mechanisms, as well as the identification of gait abnormalities in a wide range of conditions and populations. Advanced technology such as robotics, deep learning and machine learning have been used to enhance the gait analysis. However, it is important to note that the technology and techniques used in gait analysis are constantly advancing and new approaches are continually being developed. Therefore, it is crucial for researchers and practitioners to stay up-to-date with the latest developments in order to improve the diagnosis, treatment, and monitoring of gait-related conditions.

Chapter 3

Mathematical Modelling of Human Gait for Criticality Analysis

3.1 Introduction

Scientists have been challenged for decades to define the order and function of biological systems. The functional order of the biological system, which typically represents self-organised dynamical states of living matter, underpins the vast majority of biological processes. These states include spatial, temporal, and spatiotemporal structures that may be found in both living and non-living materials. Several nonlinear phenomena evolve in space and time, including patterns that depict static functions, oscillatory modes that exhibit rhythmic processes (including human gait behavior), and travelling and spiralling waves. From a mathematical point of view, biological functions must be understood with respect to its dynamic properties. Biological systems show some degrees of stability for some states, however in most common scenarios the biological system exhibits high degree of chaos behaviour. Chaos lies in between regularity and randomness, in which the system exposes unpredictable explanations when the system remains far away from its equilibrium state

due to nonlinear interactions between biological sub-related functions. The complex functions of a biological system necessitate the development of theoretical mathematical models capable of describing the nonlinear effects of biological interactions and understanding the motions of small atoms and the concentration of biochemical reactions in biological cells to provide a comprehensive understanding of the biological system's interaction dynamics. The primary goal of this chapter is to quantitatively characterise the spatiotemporal patterns of human gait when the biological system is activated by a source of perturbations, either internally or externally. It is also worth noting that chaos is only one aspect of the intriguing behavior that nonlinear systems may exhibit. In particular, disturbance or turbulence must be investigated when exposed to a biological system to observe how stable the system can survive.

Human motor systems, which are a subset of the global biological system, are physiologically complicated, and no mathematical models exist to accurately characterise their dynamics in both time and space dimensions. When a normal gait is disrupted, the entire dynamic state of the biological system drifts chaotically to a more critical state due to an excessive rise in biological system kinetic energy, and the behavior may exhibit near power-law and exponential interactions.

The interactions of the biological system functions within these disruptions produce a very complicated phenomenon that can be measured in the same way as the concentration of metabolic reactions in the biological system. Fluctuations in the dynamics of the biological system can be regulated, but they must also be analytically modelled. The RCC method, which has recently been demonstrated to be a robust technique for stabilising a dynamic system's nonlinear chaotic disturbance, regulates the evolution rate of a nonlinear system so that the exponential growth of such an unstable chaotic oscillator is controlled into stable trajectories. The control is set to utilise the rate of growth of some of the variables in proportion to their total embedded phase space. This is then passed into an exponential control function, which allows the variable's rate of change to be regulated or controlled to accelerate or

decelerate. When no control is applied or the system is not changing exponentially, the proportionate rate of change equals unity.

Despite being a recently developed technique in previous research [208] [34] [50] [41], the CA based on the RCC method has been utilised in this thesis due to its robustness in detecting disturbances in human gait. This novel technique is able to control disturbances that could otherwise affect other biological functions in the biological system. Its effectiveness in addressing these issues has made it a valuable tool in analysing gait patterns, which is essential in identifying and treating various disorders. It is inspired by the traditional biochemical enzyme control concept that adjusts the reaction control [29]. Controlling the reaction rate of biochemical reactions on the basis of local information allows the biological system to operate under a broad spectrum of certain conditions. This can certainly be extended further to control spatiotemporal chaos to achieve stability, such as the underlying dynamics of human gait. The stability of the overall biological system trajectories must meet the Lyapunov stability, such that the biological system can return to its equilibrium region of its phase space-controlled dynamics. The RCC can restore the perturbed biological system to its normal stable state by applying localised control to some of its variables; however, it does not completely suppress the biological system's underlying nonlinear behaviour, as it still has some nonlinear properties, allowing it to respond to perturbations and be less chaotic. Therefore, the RCC method does not entirely eliminate the chaotic properties of the underlying nonlinear system, but it applies limited localised control to the system to maintain an apparently stable system. The controlled system retains many of the nonlinear system's properties, including the ability to respond nonlinearly to weakly chaotic perturbations.

This chapter will discuss some important background information related to the concepts of nonlinear dynamic systems, supported by their mathematical formulations. Different types of dynamical systems will also be explored. Furthermore, the oscillatory behaviour of dynamic systems will be highlighted. The baseline charac-

teristics of the biochemical model derived by Berry model [29] will be presented. The optimised version of the Berry model [29], which is used to describe the spatiotemporal disturbances affecting the human gait, based on the rate of control of chaos, will be discussed. Simulation analysis will follow to support the proof-of-concept.

3.2 Overview to Dynamic Systems

Dynamical systems methods have been widely used in the field of mathematical biology to model various phenomena at different scales, ranging from the molecular level to the population level [175]. At the cellular level, these methods can be used to model metabolic pathways and cellular signalling networks, among other processes. In the context of behavioural gait phenomena, the use of dynamical systems methods can be relevant in several ways. For example, the coordination of the muscles involved in walking can be modelled using coupled oscillators, where each oscillator represents the activity of a specific muscle or group of muscles. These models can capture the phase relationships between the oscillators and how they change over time to produce coordinated movement [176]. Dynamical systems methods can also be used to study the stability and adaptability of gait patterns under different conditions, such as changes in speed, terrain, or external perturbations [177]. For instance, a model of human walking can be constructed using a system of differential equations that describe the dynamics of the various muscle groups, as well as the interactions between the body and the environment. This model can be used to predict how changes in the parameters of the system (e.g., muscle activation levels, joint angles, etc.) affect the stability and adaptability of the gait pattern [178]. Furthermore, dynamical systems methods can be used to analyse and interpret data from experiments on gait patterns [179]. For example, time series data on the joint angles and muscle activations during walking can be analysed using techniques such as phase space reconstruction, attractor reconstruction, and Lyapunov exponents [180]. These

techniques can provide insights into the underlying dynamics of the gait pattern, as well as the degree of variability and stability of the pattern over time. The use of dynamical systems methods for modelling cell-level metabolic phenomena can be relevant to the study of behavioural gait phenomena by providing a framework for understanding the underlying dynamics of the gait pattern, as well as its stability and adaptability under different conditions.

A dynamic system can be represented by a set of ordinary differential equations (ODEs) that describe the evolving state of the system forward in time. Time can be either a continuous or a discrete integer-valued variable. A dynamical system is said to be deterministic if it has a single future state for a definite current state and stochastic if the future state is determined by a probability distribution of potential states. Chaos theory is the study of certain deterministic dynamical systems that are extremely sensitive to initial conditions [30]. A chaotic system is a deterministic dynamical system with sensitivity to its initial conditions. Mathematically, this can be expressed as a set of ODEs:

$$\begin{cases} \dfrac{dx^{(1)}}{dt} = F_1\big(x^{(1)}, x^{(2)}, x^{(3)}, ..., x^{(N)}\big) \\ \dfrac{dx^{(2)}}{dt} = F_1\big(x^{(1)}, x^{(2)}, x^{(3)}, ..., x^{(N)}\big) \\ \qquad\qquad \vdots \\ \dfrac{dx^{(N)}}{dt} = F_1\big(x^{(1)}, x^{(2)}, x^{(3)}, ..., x^{(N)}\big), \end{cases} \tag{3.1}$$

which can also be given in a vector form as:

$$\frac{\mathbf{x}(t)}{dt} = F[\mathbf{x}(t)] \tag{3.2}$$

where $\mathbf{x}$ is an N–dimensional vector. The future system state $\mathbf{x}(t)$ for $t > 0$ can be

determined for any initial condition of the system at $\mathbf{x}(0)$. The *phase space* is the solution of the dynamic system $F[\mathbf{x}(t)]$ represented by $(x^{(1)}, x^{(2)}, x^{(3)}, ..., x^{(N)})$.

The states of a chaotic system are often thought to exist in an N-dimensional region known as *phase space* [31]. A chaotic system evolves in its phase space over time according to the system variables that drive the dynamics. A *trajectory* or *orbit* is the path taken by the system over time, and an *attractor* is the area of phase space where the trajectories settle down as time approaches infinity [32].

The behaviour of the dynamic system $F[\mathbf{x}(t)]$ in terms of stability (equilibrium), instability, or chaos can be measured when the dynamic system $F[\mathbf{x}(t)] = 0$. If there is a constant solution that satisfies $F[\mathbf{x}(t)] = 0$, then the system is at equilibrium or stable state [33]. When the dynamic system repeats itself after a specific time interval, such that $F[\mathbf{x}(t)]$=$F[\mathbf{x}(t + nT)]$, then the dynamic system exhibits periodic obits or cycles. When the dynamic system shows an unstable behaviour that is neither stable (or in equilibrium) nor periodic, then the sytem is in chaos state. This also means that the system is sensitive to small perturbations of its initial conditions, allowing the sytem to diverage from its equilibrium point [33].

3.3 Incorporation of Biomechanical Berry Model

The biochemical enzyme control of the Berry model [29] is recognisably applicable for several types of control needed for biological control processes, such as the motor system of the human gait. The RCC method has mainly been used to control many different models, including Rossler, Lorenz, and Grey-Scott models [34] [35] to analytically model biological patterns in critical systems, such as the release process of insulin.

The original enzymatic system consists of two cyclically ordered enzyme processes that interconvert chemicals A and B through antagonist enzymes E_1 and E_2. The impact of B on the activity of both enzymes results in the formation of two auto-

catalytic loops. B accelerates the rate of its formation, acting as a positive feedback loop, as well as its consumption, acting as a positive forward loop. This fundamental design idea is shown in the context of extracellular matrix degradation balance. Proteinases (p) breakdown insoluble extracellular matrix proteins (m) into soluble proteolysis fragments (f), which can then be cross-linked by transglutaminase (g) or another intermolecular protein-cross-linking enzyme to create new insoluble extracellular matrix proteins m. Proteinases (p) and transglutaminase (g) act as two opposing enzymes that interconvert extracellular matrix proteins (m) and soluble proteolysis fragments (f). The impact of f on the generation of p and g by the surrounding cells causes the autocatalytic loops. Cells use many membrane receptors that uniquely detect extracellular matrix elements to assess the features of the surrounding extracellular matrix in vivo [36].

These receptors respond to physical and chemical changes in the extracellular matrix by regulating signal transduction cascades within the cells [37]. In a first autocatalytic loop, f increases proteinase amount, which raises f formation rate. Likewise, the amount of extracellular matrix surrounding the cells can modulate transglutaminase expression and secretion [38]. Assuming that cells can up-regulate g in response to f amount, a second autocatalytic loop starts. Other biologically relevant mechanisms must be considered in addition to the fundamental bienzymatic arrangement. Surrounding cells create extracellular matrix proteins m at a constant rate r_{im} [39]. Except for p autoproteolysis, which is characterised as a simple second-order rate, the equations describing this system are based on the standard Michaelis-Menten formalism [36]. The autocatalytic loops do not directly impact the enzyme activity but instead regulate p and g cell production. Therefore, the bienzymatic cyclic model with two autocatalytic loops can mathematically be presented as:

$$\frac{dm}{dt} = k_g \frac{fg}{K_G + f} - \frac{mp}{1+m} + r_{im} \tag{3.3}$$

$$\frac{df}{dt} = -k_g \frac{fg}{K_G + f} + \frac{mp}{1+m} - \frac{fp}{1+f} \tag{3.4}$$

$$\frac{dp}{dt} = \gamma \frac{f^n}{{K_R}^n + f^n} - k_a p^2 \tag{3.5}$$

$$\frac{dg}{dt} = \beta \frac{f^l}{{K_S}^l + f^l} - k_{deg} \frac{gp}{K_{deg} + g} \tag{3.6}$$

The m variation rate is described by equation (3.3), where the first term represents the synthesis of m from both f and g. The terms k_g and K_G correspond to the catalytic and Michaelis constants, respectively. Also, the second term describes the consumption of m by the proteolysis p. The final term, r_{im}, is a constant that controls the production of m by other cells.

equation (3.4) describes the f variation rate, where the first and second terms are identical to their counterparts in equation (3.3), and the last term describes f proteolysis by p.

The first term in equation (3.5) refers to the first autocatalytic loop, where γ represents its level. The term K_R is the threshold constant, while the hill number n controls the first autocatalytic loop. The second term of equation (3.5) describes p autoproteolysis, where k_a is a constant.

Similarly, the second autocatalytic loop is represented by equation (3.6), where β describes its level. The term K_S is the threshold constant, while the hill number l controls the second autocatalytic loop. Transglutaminase g proteolysis is modelled as a Michaelis-Menten rate in the second term, where k_{deg} and K_{deg} correspond to the catalytic and Michaelis constants, respectively.

It is worth noting that m and f concentrations evolve as a result of existing enzymes, whereas p and g are introduced into the system through cell synthesis.

3.4 The Extended Model for Human Gait

The nonlinear biochemical enzyme control model discussed in Section 3.3 is being extended to describe human gait function and the effect of its disturbance to the overall biological system [40]. Using the RCC method, the model, described by equations (3.10)–(3.13), has been proven to be controllable. The model applied a control to allow the stabilisation of the external perturbations to the motor system by adjusting the amount of enzyme used to determine the concentration of variable f. This model illustrates the regulation process of two enzymes: proteinase p, which converts the matrix into filaments, and transglutaminase g, which returns the filaments to the matrix. The process results in the creation of extracellular matrix m from soluble filaments f.

Proteinase p converts the matrix into filaments, while transglutaminase g returns the filaments to the matrix. The extracellular matrix m is produced at a constant rate by neighbouring cells r_{im}, and each protein degrades in catalytic processes is proportional to the concentration of proteinase p. The bifurcation parameter r_{im} is an external turbulent input that causes the biological system to lose stability and attain a chaotic state. The RCC is defined by the soluble filaments f supplied in equation (3.7) and the rate of change of enzyme production p and g stated in equations (3.8) and (3.9). The output of this controlled model can be presented as either a time series of the primary changing parameters f and m or a phase space plot. In the latter, f is plotted on the (x-axis) and m is plotted on the (y-axis).

$$\delta_f = \frac{f}{f + \eta_f} \tag{3.7}$$

$$\Theta_p(\delta_f) = f_p e^{(}\xi_p \delta_f) \tag{3.8}$$

$$\Theta_g(\delta_f) = f_g e^{(}\xi_g \delta_f) \tag{3.9}$$

$$\frac{dm}{dt} = k_g \frac{fg}{K_G + f} - \frac{mp}{1+m} + r_{im} \tag{3.10}$$

$$\frac{df}{dt} = -k_g \frac{fg}{K_G + f} + \frac{mp}{1+m} - \frac{fp}{1+f} \tag{3.11}$$

$$\frac{dp}{dt} = \Theta_p(\delta_f)\gamma \frac{f^n}{{K_R}^n + f^n} - k_a p^2 \tag{3.12}$$

$$\frac{dg}{dt} = \Theta_g(\delta_f)\beta \frac{f^l}{{K_S}^l + f^l} - k_{deg} \frac{gp}{K_{deg} + g} \tag{3.13}$$

The extended Berry model [29] parameters are as follows [40]: $\gamma = 0.026$, $\beta = 0.00075$, $K_R = 4.5$, $K_S = 1$, $K_G = 0.1$, $K_{deg} = 1.1$, $k_g = k_{deg} = 0.05$, $k_a = \frac{k_{deg}}{K_{deg}} = 0.0455$, and the Hill numbers $n = l = 4$. The r_{im} bifurcation parameter exhibits a wide spectrum of dynamic behaviours, including periodic stable limited cycles, bistability, and chaos. This parameter remains constant for all oscillators within the chaotic domain. In this extended model, an external input is used as a perturbation to this r_{im} parameter, as shown in equation (3.14). This parameter links different oscillators together by using a relative scale contribution from all other oscillators. Several parameters of the RCC extended model throughout the experiment simulations, as shown in equations (3.7)–(3.9) ($f_p = f_g = 1$, $\xi_p = \xi_g = -1$, and $\eta_f = 2$), remained constant despite the fact that their values could change, allowing the local oscillator to change its oscillatory orbits.

$${r_{im}}^i = \sum_{j=1, j\neq i}^{n} w_j m_j + \varepsilon \tag{3.14}$$

The connectivity strength between various oscillators is represented by w_j, which can be 0.00011, 0.00012, or 0.00025. The ε parameter represents the uniform Gaussian distribution of the external perturbations applied to each oscillator and is scaled

over the domain $[-1, 1]$. Moreover, the perturbation is graphically observed over a range of evolution steps. This allows the system to explore various perturbation values resulting in different oscillatory cycles. To determine the connectivity strength value $w_j = 0.0002$, the chaotic domain of the underlying oscillators was used. For each oscillator, this value may change to affect only the dynamics, not the overall stability.

The incorporated RCC control method of the extended Berry model [29] has been shown to stabilise the local nonlinear spatiotemporal patterns of the human gait. The RCC method performs effectively when the response of the underlying system is chaotic. This method enhances the stability of nonlinear systems into stabilised periodic limited cycles according to the local dynamic behaviour of each dynamic oscillator. The network of nonlinear models consists of 16 oscillators, such that each individual oscillator can adjust its local dynamics to accommodate the external perturbations from its adjacent Neighbour s. The total model is simulated by EuNeurone software [41] using Fehlberg-RK as an Ordinary Differential Equation (ODE) fixed step integration method. The total unweighted dynamics of M and F as in equations (3.15) and (3.16) is measured by the net sum of the individual oscillators, which can be seen by a remote observer even though the individual oscillators themselves are invisible.

$$M = \sum_{i=1}^{n} m_i \tag{3.15}$$

$$F = \sum_{i=1}^{n} f_i \tag{3.16}$$

3.5 Topology of Phase Space

The phase space manifold is the space that contains all of the system states. Future states of a deterministic dynamical system can be determined by using current and previous states of knowledge. However, the system equations may be too sophisticated for some applications, such as the interpretation of human gait dynamics. Despite this complexity, the phase space topology contains significant information about the dynamic system and can be described mathematically as:

$$\mathbf{x(t)} = \mathbf{(x^{(1)}, x^{(2)}, x^{(3)}, ..., x^{(N)})} \tag{3.17}$$

The phase space manifold in the extended model can be written as:

$$\mathbf{x(t)} = \mathbf{(F, M)} \tag{3.18}$$

3.6 Extended Model Simulation Results

The original and extended models were simulated to demonstrate how the RCC-incorporated method improves the stability of the RCC-free model. At the given values of the other system parameters, the initial conditions of $m(0) = f(0) = p(0) = g(0) = 0.8$ were used. The simulation used 94×10^5 data points with a 0.5 time step size. The phase plot of the original model depicted in Figure 3.1 shows the projection of system trajectories over time in $f - m$ plane. It also shows that the generated orbits fluctuate between small and large amplitudes on the trajectories. As a result, future state time-variation plots shown in Figures 3.2 and 3.3, exhibit abrupt and high peaks interrupted by intervals of relative quiescence, represented by smaller amplitude oscillations.

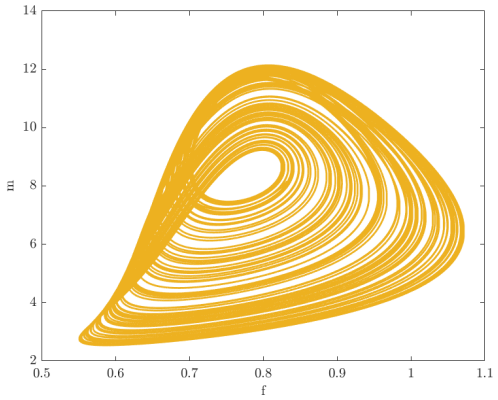


Figure 3.1: Phase plot projection in the $f-m$ plane of the uncontrolled Berry model.

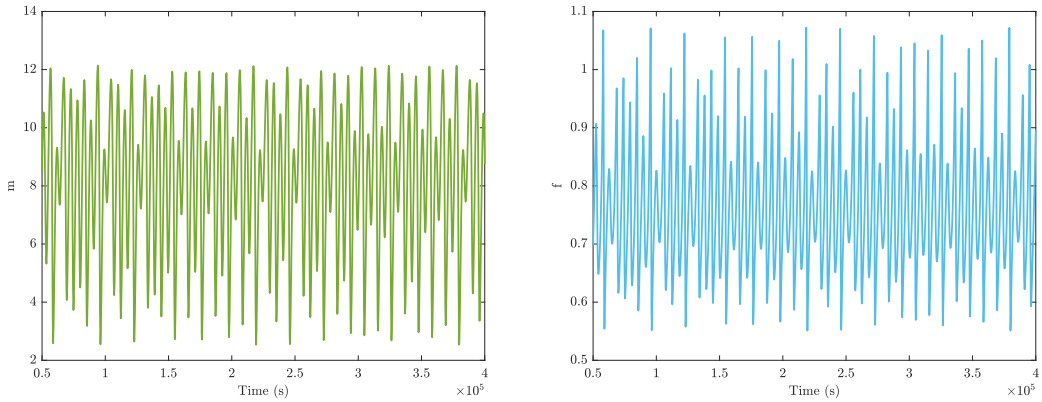


Figure 3.2: Time-variations of m and f along the uncontrolled Berry model.

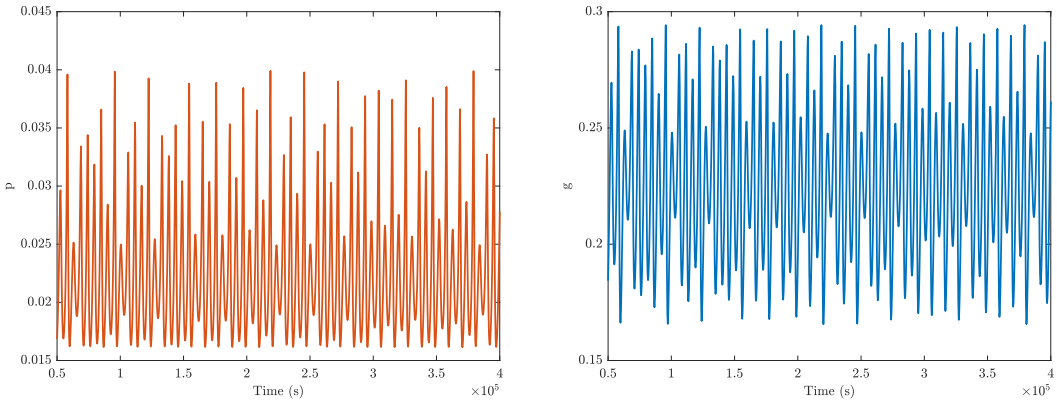
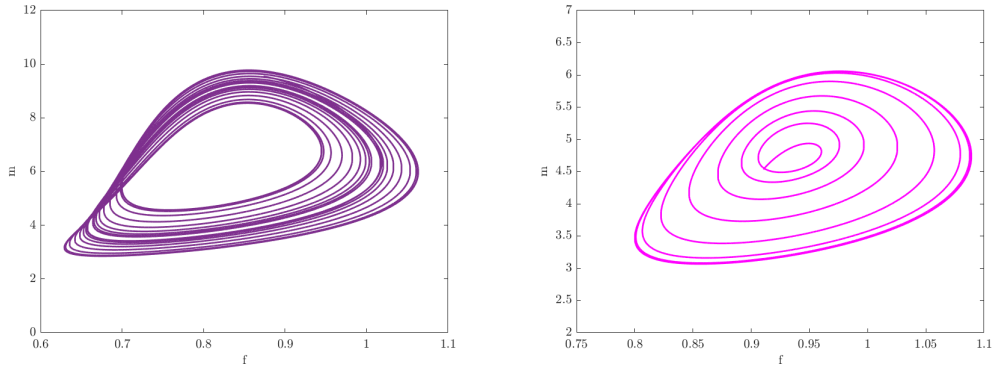


Figure 3.3: Time-variations of p and g along the uncontrolled Berry model.

When the RCC approach is applied to the original model shown in equations (3.10)–(3.13), the extended model becomes more stable in more structured trajectories at $\xi_p = -1$ as illustrated in Figure 3.4a, with a minor drop in amplitude when compared to the trajectories of the RCC-free model. The future state time-variation plots of the extended model shown in Figures 3.5 and 3.6, it exhibits on average consistent oscillations around 9.2561×10^4. The phase plot of the extended model at $\xi_p = -3$, on the other hand, exhibits sparse trajectories on the $f - m$ plane, but it stabilises on average 10.743×10^4 slower than when $\xi_p = -1$. The values of ξ_p and ξ_g are -1 for the stable trajectories and -3 for the sparse trajectories.



(a) At $\xi_p = -1$ (b) At $\xi_p = -3$

Figure 3.4: Phase plot projection in the $f - m$ plane of the controlled Berry model at $\xi_p = -1$ and $\xi_p = -3$ respectively.

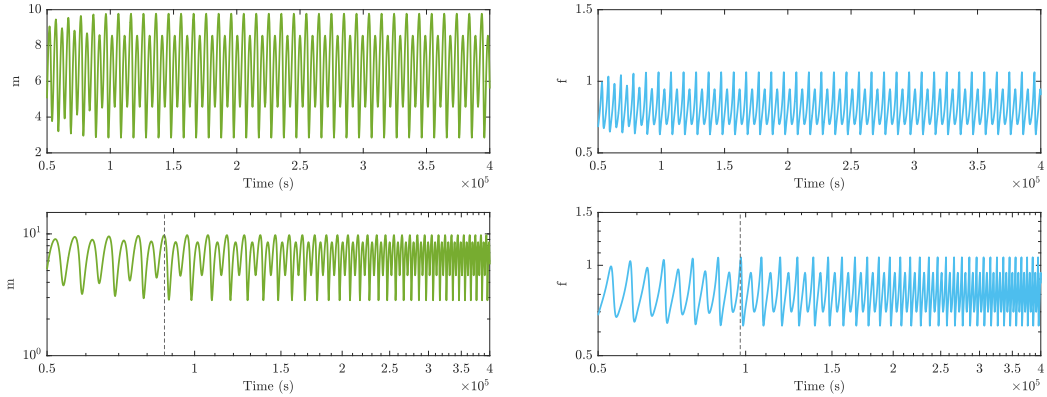


Figure 3.5: Time-variations and their corresonding Log-log plots of m and f along the controlled Berry model at $\xi_p = -1$.

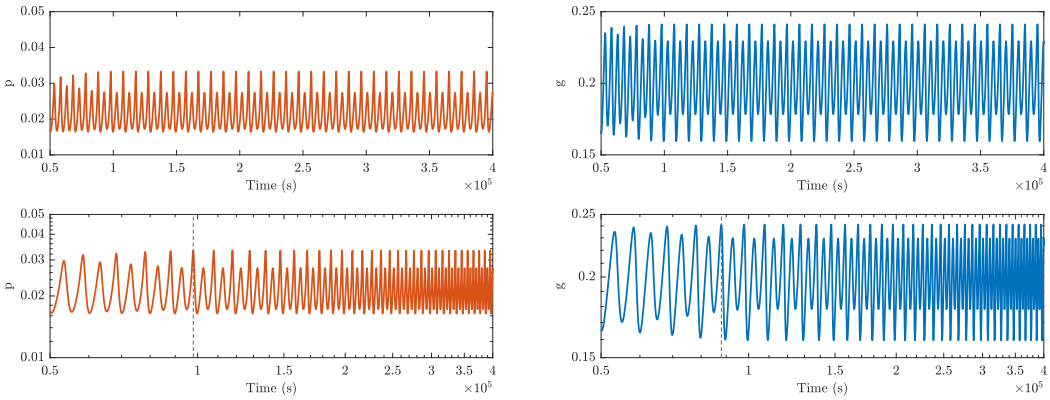


Figure 3.6: Time-variations and their corresonding Log-log plots of p and g along the controlled Berry model at $\xi_p = -1$.

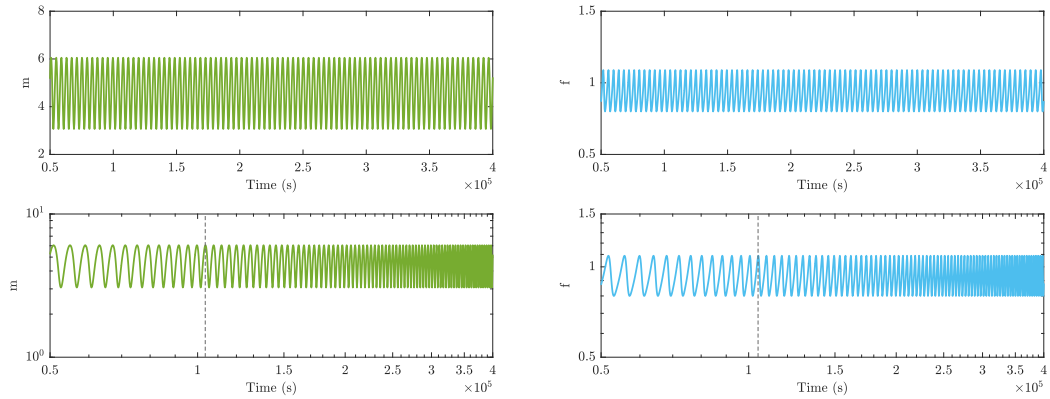


Figure 3.7: Time-variations and their corresonding Log-log plots of m and f along the controlled Berry model at $\xi_p = -3$.

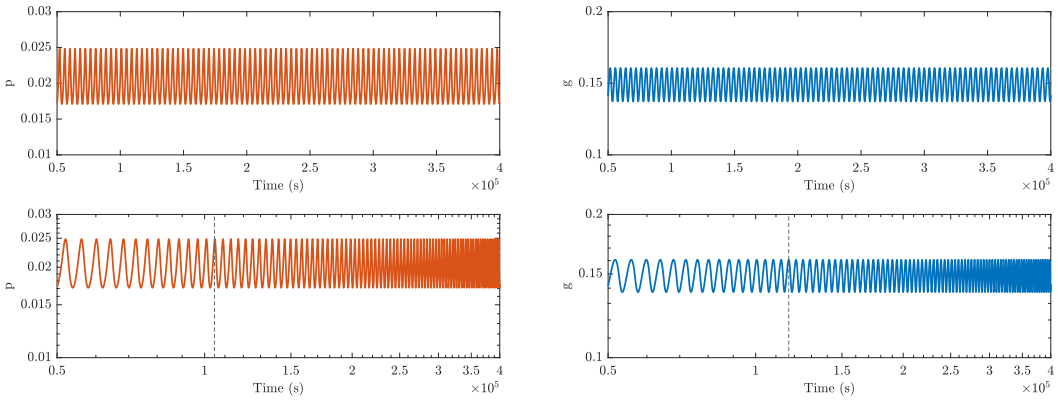


Figure 3.8: Time-variations and their corresponding Log-log plots of p and g along the controlled Berry model at $\xi_p = -3$.

The phase space trajectories of the extended model in the $f - m$ plane indicate the prominent characteristics that explain the behavior of the dynamic system and its sensitivity to small perturbations. Therefore, the vector space of these manifolds was used as an input feature for the proposed SVM supervised machine learning algorithm.

3.7 Summary

In this chapter, fundamental concepts of nonlinear dynamic systems and their mathematical formulation are introduced. Additionally, a mathematical model is presented, which explains how rate control of chaos can regulate disturbances that affect the biological system, particularly the gait system in humans. The model operates at specific parameters that maintain the stability of the motor system, with the connectivity strength that connects multiple oscillators being the most crucial parameter. A slight deviation or shift in the connectivity value can cause the system to operate in a chaotic mode, limiting the overall system's stability and control.

Chapter 4

Comparative Analysis of Support Vector Machine (SVM) and K-Nearest Neighbour s (KNN) Algorithms for Human Gait

4.1 Introduction

The focus on the two machine learning (ML) methods, K-Nearest Neighbour (KNN) and Support Vector Machines (SVM), for gait analysis is primarily due to their ability to handle high-dimensional data and their effectiveness in classification tasks [181]. Gait analysis often involves processing large amounts of data from multiple sources, such as accelerometers, gyroscopes, and force sensors [182]. The resulting data can be high-dimensional, making it challenging to extract relevant features and perform classification tasks. KNN and SVM are both well-suited for handling high-dimensional data and have been shown to perform well in gait analysis tasks [181]. KNN is a non-parametric method that classifies data points based on

their proximity to a set of training examples. It has been used in gait analysis to classify gait patterns based on features extracted from accelerometer and gyroscope data. KNN is relatively simple to implement and can be effective in cases where the decision boundary between classes is highly nonlinear [183]. SVM, on the other hand, is a supervised learning method that finds a hyperplane in a high-dimensional space that maximally separates the classes of data points [184]. It has been used in gait analysis to classify gait patterns based on features extracted from force plate data, electromyography (EMG) signals, and other sensor data [185]. SVM can handle non-linearly separable data by using kernel functions, which transform the data into a higher-dimensional space where it can be separated by a hyperplane [186]. While there are many other ML methods that can be used for gait analysis, KNN and SVM are among the most widely used and have been shown to perform well in many applications. Other methods such as Artificial Neural Networks (ANNs), Random Forests (RF), and Principal Component Analysis (PCA) have also been used in gait analysis, but their performance may depend on the specific application and dataset [187].

4.2 K-Nearest Neighbour Algorithm

4.2.1 Overview of KNN

The KNN is a non-parametric supervised machine learning classification technique. The principle of KNN relies on computing the Euclidean distance between the test (unknown data points) and the training data samples. Let $\mathbf{x} \in \mathbb{R}^{n \times d}=(x_1, ..., x_n)$ be the matrix of features, where n is the number of training samples and d is the number of features. For a given an arbitrary point in the unknown samples set x_o, the Euclidean distance in the feature plane $\mathbb{R}^p$, where $p = 2$ is a real number, can be formulated as:

$$d_i = \|\mathbf{x_r} - \mathbf{x_o}\|_p = \left(\sum_{i=1}^{n} |x_i - x_o|^p \right)^{\frac{1}{p}} \tag{4.1}$$

To classify a number of features into M classes, then the outcome of classified entities can be presented as Ω=$\{\Omega_1, ..., \Omega_m\}$, where $1 \leq m \leq d$. Choosing the k training samples with the minimum distance to the unknown data point x_o, the KNN algorithm calculates the number of Neighbours assigned to each data class $l \in \mathbb{R}^{1\times d}$=$(l_1, ..., l_d)$ existing in the training set $\mathbf{S}_r$=$\{(x_1, l_1), ..., (x_n, l_d)\}$, where $\mathbf{x}_r \in \mathbb{R}^{n\times 1}$=$(x_1, ..., x_n)$ is the training example associated with $\mathbf{S}_r$. Each member in $\mathbf{S}_r$ corresponds to a class label in Ω. The process is fundamentally based on estimating the conditional probability for each class as an empirical fraction. This can mathematically be given by:

$$\mathbf{P}_r = p\Big[m(l) \in l \mid \mathbf{x} = x_o\Big] = \frac{1}{k} \sum_{i \in \mathcal{N}(l, \mathbf{S}_r)} \mathbb{I}(\mathbf{x}_r \in l) \tag{4.2}$$

where $\mathcal{N}(l, \mathbf{S}_r)$ are the indices of the k nearest data samples to l in the training set $\mathbf{S}_r$. $\mathbb{I}(.)$ is an indicator function expressed as:

$$\mathbb{I}(w) = \begin{cases} 1, & \text{if } w \text{ is True} \\ 0, & \text{otherwise} \end{cases} \tag{4.3}$$

4.3 Support Vector Machine Algorithm

4.3.1 Overview of SVM

The Support Vector Machine (SVM) represents a powerful and versatile tool in the realm of machine learning. Its ability to handle non-linearly separable data and

perform both binary classification and regression estimation tasks has made it a widely adopted technique in various fields such as disease diagnosis, image classification, and facial recognition [42] [43] [44]. The utilisation of a kernel technique in the SVM algorithm allows for the formation of a non-linear class decision boundary, which is crucial in tackling complex classification problems. The primary objective of an SVM classifier is the identification of an optimal separating hyperplane (OSH) that maximises the margin between the different classes. The kernel technique transforms the input data into a high-dimensional space, and subsequently, an OSH is created to classify the various data labels in the transformed feature space. As a result, the linear OSH leads to the formulation of a non-linear boundary in the original data input space. The data vectors closest to the OSH in the transformed space are referred to as support vectors, as they contain vital information regarding the OSH. The choice of kernel function is crucial in SVM as it greatly affects the classification accuracy. Using the right kernel can enhance the SVM's performance, so it is important to carefully consider the nature of the data and the task at hand when selecting a kernel.

4.3.2 Basic Theory of SVM

Given a training dataset, $\Omega = \{(\boldsymbol{x_i}, y_i)\}\, i = 1^N$, independently drawn from a probability distribution on $(\mathcal{X}, \mathcal{Y})$, where $\mathcal{X} \in \mathbb{R}^m$ represents the input features and $\mathcal{Y} \in \{-1, +1\}$ represents the classification output, the Support Vector Machine (SVM) model can achieve optimal separation of linearly separable patterns in m-dimensional space by generating a decision function, $\boldsymbol{w^T}\phi(\boldsymbol{xi}) + b = 0$, through the minimisation of an appropriate trade-off between the structural empirical risk and the model complexity of its optimisation problems. The optimal adjustable weight vector, $\boldsymbol{w}_\circ$, and the optimal bias, $b_\circ$, of the decision function are defined when the feature vectors $\boldsymbol{x_i}$ are maximised. For any two arbitrary classes, $\{-1, +1\}$, the SVM

finds two parallel hyperplanes that correctly classify all training data points and either maximises the distance, $\frac{2}{\|\boldsymbol{w}\|}$, or minimises the margin, $\frac{1}{2}\|\boldsymbol{w}\|^2$, between them. For each linearly separable case, the SVM standard classification optimisation problem can be mathematically expressed as:

$$\begin{aligned} \min_{\boldsymbol{w},b} \quad & \frac{1}{2}\|\boldsymbol{w}\|^2 \\ \text{subject to} \quad & y_i(\boldsymbol{w}^{\boldsymbol{T}}\phi(\boldsymbol{x_i}) + b) \geq 1 \end{aligned} \tag{4.4}$$

The SVM classifier, represented by the equation $y_i = \text{sign}(\boldsymbol{w}^{\boldsymbol{T}}\phi(\boldsymbol{x_i}) + b)$, is utilised to assign each class to either side of the hyperplane. The minimum value of $(\boldsymbol{w}^{\boldsymbol{T}}\phi(\boldsymbol{x_i}) + b)$ is 1, and the maximum value is -1, when the conditions $y_i = +1$ and $y_i = -1$ are satisfied, respectively. In practical applications, it is common for data points to fall within the margin space or beyond the decision boundary, which poses a significant challenge for accurate classification. To address this issue, the use of a soft margin approach, utilising slack variables ξ_i, has been proposed as a practical solution for handling non-separable data. The primal form of the SVM as defined in 4.4 can be re-formulated accordingly.

$$\begin{aligned} \min_{\boldsymbol{w},b,\xi} \quad & \frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{n}\xi_i \\ \text{subject to} \quad & y_i(\boldsymbol{w}^{\boldsymbol{T}}\phi(\boldsymbol{x_i}) + b) \geq 1 - \xi_i \\ & \xi_i \geq 0 \end{aligned} \tag{4.5}$$

In the context of SVM, the regularisation parameter, denoted by C, serves to balance the trade-off between the margin and the loss. As outlined in equation 4.5, when $y_i(\boldsymbol{w}^{\boldsymbol{T}}\phi(\boldsymbol{x_i}) + b) \geq 1$, the value of ξ_i is set to zero, as previously stated in equation 4.4. However, if $y_i(\boldsymbol{w}^{\boldsymbol{T}}\phi(\boldsymbol{x_i}) + b) < 1$, then ξ_i takes on a positive value that

satisfies $y_i(\boldsymbol{w}^{\boldsymbol{T}}\phi(\boldsymbol{x_i}) + b) = 1 - \xi_i$. The SVM employs the Hinge loss function as a means of assessing the empirical risk of the training data points in accordance with the inequality $y_i(\boldsymbol{w}^{\boldsymbol{T}}\phi(\boldsymbol{x_i}) + b) < 1$. It should be noted that data points located in proximity to the boundary of the separating hyperplane may exhibit noise, which has the potential to skew the resultant separating hyperplane. To mitigate this issue and minimise model complexity, the SVM incorporates a regularisation term into its optimization problem. The standard form of the SVM Definition with slack variable, as represented in equation 4.5, can be represented as follows:

$$\min_{\boldsymbol{w},b} \quad \frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{n}\theta\Big(y_i(\boldsymbol{w}^{\boldsymbol{T}}\phi(\boldsymbol{x_i}) + b)\Big) \tag{4.6}$$

Utilising the Hinge loss function, represented by the variable θ,

$$\theta(\alpha) = (1-\alpha) = \begin{cases} 1-\alpha, & 1-\alpha > 0 \\ 0 & \text{otherwise} \end{cases}$$

The optimisation problem presented in equation 4.5, when combined with equation 4.6, can be mapped to a constrained optimisation problem with linear constraints and a global minimum. The methodology for evaluating the optimal values of $\boldsymbol{w}_\circ$ and $b_\circ$ can be found in the references [42] [45]. For the purpose of simplicity, the bias or offset parameter b can be disregarded and the output of the prediction function f can be parameterised by $\boldsymbol{w}$ as $f(x) = \langle \boldsymbol{w}, \phi(\boldsymbol{x_i})\rangle$. Thus, the SVM model outlined in equation 4.6 aims to solve the optimization problem as follows:

$$\min_{\boldsymbol{w}\in\mathbb{H}} \quad \frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{n}\theta(y_i, \langle \boldsymbol{w}^{\boldsymbol{T}}, \phi(\boldsymbol{x_i})\rangle) \tag{4.7}$$

In the context of a Reproducing Kernel Hilbert space (RKHS) $\mathbb{H}$ induced by a

kernel function $\kappa(\phi(\boldsymbol{x}), \phi(\boldsymbol{z}))$ and a feature mapping function $\phi : \mathbb{R}^m \mapsto \mathbb{H}$, the SVM model of equation 4.7 presents challenges in terms of efficiency when addressing nonlinear problems. This is due to the fact that the function $\phi(\cdot)$ is often high-dimensional and potentially infinite in terms of its mapping capabilities. However, through the utilisation of the representer theorem [46], it can be shown that there exists a vector $\beta^* \in \mathbb{R}^m$ such that the solution of equation 4.7 holds $\boldsymbol{w^*} = \sum_{i=1}^{m} \beta^* i \phi(\boldsymbol{xi})$. By substituting $\boldsymbol{w} = \sum_{i=1}^{m} \beta_i \phi(\boldsymbol{x_i})$ into equation 4.7, the optimisation problem can be presented as a finite-dimensional equivalent as follows:

$$\min_{\beta \in \mathbb{R}^m} \quad \frac{1}{2}\|\boldsymbol{w}\|^2 + C \sum_{i=1}^{n} \theta(y_i, \boldsymbol{K}_i \beta_i) \tag{4.8}$$

In the equation presented, the kernel matrix, denoted as $\boldsymbol{K}$, satisfies the relationship $\boldsymbol{K}_{i,j} = \kappa(\boldsymbol{x}_i, \boldsymbol{x}_j)$, where $\boldsymbol{K}_i$ represents the k-th row of $\boldsymbol{K}$. It has been established in literature, specifically in [45], that the coefficient β_i is subject to a bound, as it satisfies the inequality $0 \leqslant \beta_i \leqslant C$.

The Gaussian kernel function, also known as the radial basis function (RBF), is a commonly utilised technique in the capacity control and regularisation of radial basis function networks. This function, as outlined in [47], takes the form $\boldsymbol{K}_{i,j} = \kappa(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(\frac{-\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}{2\sigma^2}\right)$, where the function space is based on the norm in the RKHS. It has been observed that as the data points $\boldsymbol{x}_i$ move away from the center $\boldsymbol{x}_j$, the function monotonically decreases. The width parameter σ plays a crucial role in controlling the rate at which the RBF function decreases, and it has been established that this parameter is inversely proportional to its norm, as stated in [48].

4.4 Proposed Support Vector Machine Algorithm

The proposed Support Vector Machine (SVM) model capitalises on the utilisation of the Kernel property outlined in Section 4.8, owing to its capability of converting this nonlinear dynamic gait data into a novel feature space, thereby streamlining the classification process. The controlled CA model extracts salient features, represented by f of equation 3.11 on the x-axis and m of equation 3.10 on the y-axis, that serve as inputs for the SVM algorithm, enabling the assessment of classification performance under varying settings. The proposed SVM classifier underwent rigorous training and testing utilising these extracted features, specifically for both the healthy control and obese groups. The primary objective of this experiment was to evaluate the capacity of the SVM to discern gait patterns between the two groups. Algorithm 1 outlines the implementation procedures of the proposed SVM algorithm.

Algorithm 1 Pseudocode of SVM Implementation

1: **Input:** Given a training set $\Omega = (\boldsymbol{x_i}, y_i) : \boldsymbol{x_i} \in \mathbb{R}^m,\ y_i \in \{-1, +1\}$, $i = \{1, 2, ..., N\}$ and testing dataset $\boldsymbol{x} \in \mathbb{R}^m$.

2: **Output:** Predict subjects label for testing data $\boldsymbol{x}$.

3: Select a regularisation parameter C such that $C > 0$ and choose an appropriate kernel width control variable σ for validation.

4: Compute Gaussian Kernel RBF $\kappa(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(\frac{-\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}{2\sigma^2}\right)$.

5: Solve the kernalised optimisation problem in equation (4.8) using CVX optimisation solver [51] in MATLAB.

6: Obtain optimal value of β in equation (4.8) and the bias b in equation (4.6).

7: Predict labels for testing data $\boldsymbol{x}$.

8: Obtain performance measure (confusion matrix, accuracy, Receiver Operating Characteristics (ROC) Curve.

4.4.1 Preparing the Machine Learning Dataset

The proposed technique trains the extracted CA gait features from the Berry model [29] and uses them as input data for the SVM approach. The CA dataset's input data is split into two sets: (1) *Training Set*, which is used to fit the SVM model by learning a set of quantifiable properties of the input data and the data outcome labels. This is followed by (2) *Testing Set*, which evaluates the trained machine learning model on new data. To compensate for any data deviation, both sets are normalised using the $z-$score, which scales the data features to have a '0' mean and a '1' standard deviation. The SVM standard form finds it challenging to handle nonlinear data in the input space due to the dynamics involved in human gait. Therefore, nonlinear data samples can be mapped from their original input space to a new higher-dimensional feature space using this method. To express the binary SVM model, the dot product is applied to all data points, which transforms the original feature space to a new one through the Kernel Trick using a Kernel function like the *Linear Kernel*, *Polynomial Kernel*, or *Gaussian Kernel* [201]. In this case, the Gaussian Kernel is used and defined as $\boldsymbol{K}_{i,j}$ in equation (4.8) due to its ability to handle nonlinear data effectively, which is based on computing the distance between data samples in the input space. When two points in the input space are close together, the angle between the vectors representing them in the Kernel feature space is small. After Kernelisation of the CA dataset, the developed convex optimization toolbox in Matlab is employed to solve the dual SVM issue stated in equation (4.8), ensuring misclassified data samples are within the SVM margin thresholds.

4.4.2 Performance Analysis of the Proposed SVM Model

Once the model has been optimised as shown in equation (4.8) and trained, the next phase involves testing the model on newly generated data and predicting la-

bels for the test data. The proposed SVM algorithm's performance is evaluated using three primary metrics: the Confusion Matrix, which includes four key measures: true positive, false positive, true negative, and false negative values. Additionally, precision, recall, F_1-score, true positive rate, false positive rate, specificity, and overall accuracy of the model are calculated using the four metrics of the confusion matrix. These metrics are analysed in-depth in subsequent chapters. The following code snippet illustrates how to implement the confusion matrix parameters in Matlab.

```
%Confusion Matrix Calculations

ytest; % Test set
ypred; % Prediction set
npos = sum(Y_test == 1); %Number of positive samples in the test set
nneg = sum(Y_test == -1); %Number of negative samples in the test set

TP = sum((ypred == 1) & (Y_test == 1)); % True positive
FP = sum((ypred == 1) & (Y_test == -1)); % False positive
FN = sum((ypred == -1) & (Y_test == 1)); % False negative
TN = sum((ypred == -1) & (Y_test == -1)); % True negative

precision = 0;
  if ( (TP + FP) > 0)
    precision = TP / (TP + FP); % Precision
  end

recall = 0;
  if ( (TP + FN) > 0 )
    recall = TP / (TP + FN); % Recall
  end

F1 = 0;
  if ( (precision + recall) > 0)
    F1 = 2 * precision * recall / (precision + recall); % F1-score
```

```
  end

TPR = TP/npos; % True positve rate

FPR = FP/nneg; % False positve rate

Specificity= TN/(TN + FP); % Specificity

ConfusionTable= [TP, FP; FN, TN]; % Confusion table

ModelAccuracy= (TP + TN)/(npos + nneg); % Accuracy of the SVM model
```

Moreover, the proposed SVM model is designed to focus specifically on computing the area under the ROC curve (AUC), which is a standard performance metric used to evaluate machine learning models by measuring the discriminative ability between two classes through the area under the receiver operating characteristic (ROC) curve. The AUC value ranges from 0.5 to 1, where 0.5 indicates random classification and 1 indicates perfect classification. Generally, AUC values can be categorised as poor performance ($\leq 70\%$), fair performance ($\geq 70\%$), good performance ($\geq 80\%$), and excellent performance ($\geq 90\%$). Furthermore, the following code snippet illustrates how to implement the ROC curve and estimate the AUC in Matlab.

```
%Confusion Matrix Calculations

ytest; % Test set
ypred; % Prediction set

ytest=ytest>0;
[ypred,ind] = sort(ypred);
ytest       = ytest(ind);

fpr = cumsum(ytest)/sum(ytest); % False positive rate
tpr = cumsum(1-ytest)/sum(1-ytest); % True positive rate
```

```
tpr = [0 ; tpr ; 1];
fpr = [0 ; fpr ; 1];
n = size(tpr, 1);
AROC = sum((fpr(2:n) - fpr(1:n-1)).*(tpr(2:n)+tpr(1:n-1)))/2; % Area under the ROC curve

plot(fpr,tpr); % Plot ROC curve and AROC measures the area under the ROC curve
```

4.5 Summary

This chapter provides a comparative analysis of two supervised machine learning algorithms, Support Vector Machines (SVM) and K-Nearest Neighbour s (KNN), for human gait analysis. The chapter outlines technical details for each technique, along with a comprehensive explanation of the implementation process. While KNN is implemented on one dataset, the SVM is employed in the implementation study of all three datasets explored in chapters 5, 6, and 7. Each strategy's mathematical foundation is explained in detail, along with a contemporary perspective on how machine learning research has influenced human gait recognition and detection. Furthermore, this chapter specifically addresses the implementation of the proposed SVM algorithm processes and how the SVM model employs the Kernel Trick feature to overcome the nonlinearity issue in human gait data. Finally, the chapter emphasises the critical factors to consider when evaluating the model's performance.

Chapter 5

Detecting Human Gait Metabolism Disorders Based on the Criticality Analysis System

5.1 Introduction

The human gait is determined by the interactions between the neurological, musculoskeletal, and cardiorespiratory systems, and can be affected by age, personality, mood, and environmental factors [202]. Safe walking requires healthy intellect and executive control. Gait abnormalities can lead to a loss of personal freedom, falls, injuries, and a significant decline in the quality of life [203]. The sudden onset of a gait issue may indicate cerebrovascular or other acute nervous system injuries, systemic disorders, or harmful drug effects, especially from polypharmacy containing sedatives [204]. Normal gait is characterised by rhythmical and seamless movement, with legs swinging freely and an upright body posture. However, disturbances to vital functions can cause partial or complete impairment to the biological system, leading to abnormal gait behavior. In addition to external factors causing physical

gait impairments, inherited metabolic disorders such as obesity, which may be caused by genetic conditions or excessive calorie consumption, can lead to metabolism and enzyme deficiency issues, resulting in imbalanced gait [205].

This chapter introduces the proposed Kernelised SVM algorithm, which is based on the developed criticality analysis technique of gait. The algorithm aims to build several binary classification models to identify individuals at risk of obesity among healthy control teenagers. A comprehensive experimental study is conducted, resulting in improved classification accuracy and overall model performance.

5.2 CARDIGAN Dataset

In an effort to support the Criticality Analysis of Diabetic Gait in Children (CARDIGAN) project, a provisional clinical study was conducted at Mexico Children's Hospital (MCH) between January 2019 and January 2020. Participants, who were at the early adolescence age and suffering from both Obesity and Diabetes, were offered the opportunity to participate in the study in accordance with the Newton Fund Grant Agreement. All participants provided informed consent prior to their involvement in the study.

The critical pathology characteristics of the participants were assessed by a medical professional prior to and during walking exercise. This assessment aimed to measure the participants' ability to walk independently as their primary means of mobility, as well as to determine the underlying causes of their health conditions and their level of motivation to follow the program during the clinical study. Additionally, a healthy control group, free from peripheral injury or other conditions that may impede their mobility without the use of assistive devices, was also recruited to participate in the trial.

For the clinical experiment, the participants, consisting of 50 individuals, were asked to walk back and forth along a 30-metre track on a flat surface over a period of

6 minutes, starting at a slow walking speed and gradually increasing to a fast walking speed. Critical changes in walking control, stride frequency, and length of steps were observed as the participants changed speed, effectively stressing their mobility. The experiment was repeated after 3 and 6 months as a follow-up.

The assessment was based on the use of a movement sensor, placed on the fourth lumbar vertebra located on the top left of the anatomical position of the lumber spine, known as the body Centre of Mass (CoM). The sensor was designed to be incredibly flexible, providing for mobility in many different planes including flexion, extension, side bending, and rotation. The single built-in Inertial Measurement Unit (IMU) sensory helped determine the motion, orientation, and heading of the body's motor function. Participants were also provided with a wrist-worn Axivity AX3 triaxial accelerometer, which measures linear acceleration from $^{+}/_{-}$ $2g$ to $^{+}/_{-}$ $16g$ range along three orthogonal axes known as z (upward and downward), y (left and right), and x (forward and backward), to record and track their physical activity. The dynamics of their walking activity were monitored using the Polar Team tracking system [207].

As part of the research work of this thesis, Oxford Brookes University received access to the CARDIGAN dataset for research purposes, which has been anonymised in accordance with the regulations outlined in the General Data Protection Regulation (GDPR). The data underwent official approval procedures by the University Research Ethics Committee (UREC) before being analysed. The data was divided into three groups: Healthy controls, individuals with Obesity, and those with Diabetes. However, because of a lack of diabetic data points, they were not included in the analysis. Only 40 data points were used for the analysis, with 20 coming from the healthy control group and the remaining from the Obesity group.

5.3 Methodology

The proposed methodology for classifying human gait disorders includes a framework consisting of several key components, including data collection, data processing, feature extraction, and the use of the SVM technique. This methodology is illustrated in Figure 5.1.

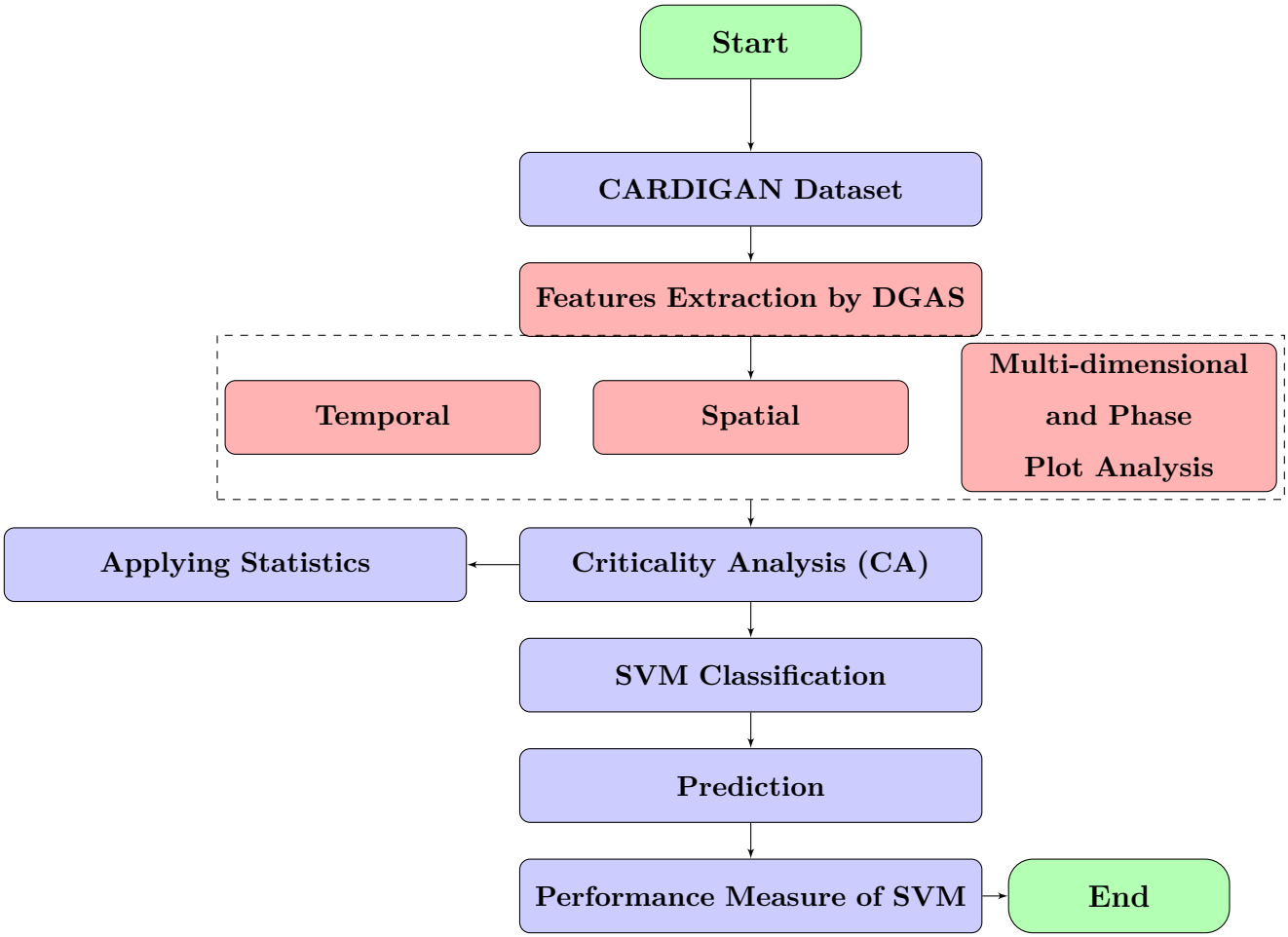


Figure 5.1: The Methodology of CARDIGAN Data Analysis

5.3.1 Data Analysis

The analysis of the CARDIGAN dataset was conducted using the DataGait Analysis Software (DGAS), which was commercially developed as a stand-alone software analysis package using LabVIEW2011 (National Instruments, Ireland) by the Movement Science Group at Oxford Brookes University. DGAS serves as a self-contained software analysis tool designed to extract gait parameters from sensor data obtained from an inertial measurement unit (IMU) containing accelerometers, gyroscopes, and magnetometers. To achieve this, DGAS employs quaternion rotation matrices and double integration to measure translational vertical CoM accelerations during walking and obtain a relative change in position. This process involves transposing the acceleration frame of the object's $z-$axis to the global system, and the quality of human gait parameters on a global scale is assessed by measuring upward CoM [206] [207].

5.3.2 Features Extraction

It is imperative to note that DGAS relies on the analysis of a single inertial measurement unit (IMU) that comprises of three-dimensional accelerometers, gyroscopes, and magnetometers, which accurately provide orientation and acceleration data. The single-sensor approach was deliberately chosen to facilitate the description of bodily movements in a quick and straightforward manner. The approach involves the attachment of an IMU over the projected center of mass of the human body, which is positioned above the fourth lumbar vertebra. This placement allows for easy accessibility and application of the required sensors for gait analysis, with the CoM aligning with the top of the Iliac crest. The use of the single sensor method described earlier enables the application of pendulum mathematics to analyse data and extract parameters related to walking. The DGAS software extracts a total of 17 features from the IMU sensor data, as summarised in Table 1. The gait features

that have been extracted are divided into three sub-groups:

- **Temporal and Spatial Parameters:** When describing someone's gait pattern, temporal and spatial parameters are commonly utilised. The timing aspect of gait is referred to as temporal parameters, while the distance aspect is denoted as spatial parameters.

  Temporal Gait Parameters:

  1. **Cadence** [**steps/min**] refers to the number of steps taken in a given time period, typically measured in steps per minute, with the adult norm being 113 steps per minute.
  2. **Step Time** [**ms**], measured in milliseconds, refers to the duration between the first contact of one foot to the first contact of the opposite foot, with the adult norm being 500ms.
  3. **Stride Time** [**ms**], also measured in milliseconds, refers to the duration between the first contact of one foot to the next contact of the same foot, with the adult norm being 1000ms or 1 second.

  Spatial Gait Parameters:

  1. **Step Length** [**m**] refers to the distance between the first contact of one foot and the initial contact of the other foot, with an adult average of 0.69m.
  2. **Stride Length** [**m**] indicates the distance between the initial contact of one foot and the next contact of the same foot, with an adult average of 1.38m.
  3. **Velocity** [**m/s**] indicates the distance traveled in a specific amount of time, with an adult average of 1.5m/s.

- **Multi-dimensional Gait Parameters:**

1. **Duty Factor Double Stance** [%] is the duration in the gait cycle when both feet touch the ground, referred to as the stance phase, which typically accounts for 38-40% of the cycle in adults.
2. **Duty Factor Single Stance** [%] refers to the time when one foot contacts the ground, also called the swing phase, and usually lasts for 60-62% of the cycle in adults.
3. **Froude Number** [**au**] is a dimensionless gait parameter used to determine the optimal walking speed, normalised by an individual's height or leg length.
4. **Walk Ratio** [**mm/steps/min**] is a gait ratio calculated by dividing the steplength by the cadence, and it is believed that this ratio should remain constant, regardless of speed.

- **Phase Plot Analysis:** The aim of this analysis is to investigate gait variability. The output of this analysis consists of four numerical values:

1. **Beta Angle** [**Degree (°)**] represents the alteration in step length as measured by the vertical movement of CoM.
2. **SDa** [**au**] is a non-standard unit that indicates the change in step frequency and step length.
3. **SDb** [**au**] is a non-standard unit that indicates the change in step frequency.
4. **Ratio** [**Dimensionless**] is the ratio between SDa and SDb.

5.3.3 Criticality Analysis as a Data Representation Method

The features extracted from the data, as shown in Table 1, are utilised to perturb the mathematical model represented by equations (3.10)–(3.13). This perturbation

Table 1: Extracted Gait Features of CA System

Gait Parameter	**Measurement**	**Unit**
Temporal	Step Time	[ms]
	Step Time (Left)	[ms]
	Step Time (Right)	[ms]
	Stride Time	[ms]
	Cadence	[steps/min]
Spatial	Step Length (Left)	[m]
	Step Length (Right)	[m]
	Stride Length	[m]
	Velocity	[m/s]
Multi-dimensional	Duty Factor Double Stance	[%]
	Duty Factor Single Stance	[%]
	Froude Number	[au]
	Walk Ratio	[mm/steps/min]
Phase Plot Analysis	Beta Angle	[Degree (°)]
	SDa	[au]
	SDb	[au]
	Ratio= SDa/SDb	[Dimensionless]

is achieved using the Criticality Analysis method, which converts the multivariate dynamic data into a nonlinear representation space, where each sample is characterised by a unique orbit. The orbit is a result of the perturbation of the underlying critical system, which is composed of a network of nonlinear controlled oscillators. This network is critical because any small input change leads to a state change in the orbit of the entire network, while remaining stable. The resulting orbit is a scale-free representation of the original data, which can be used to demonstrate that a specific set of attributes representing a patient's or control's gait is similar to the matching category of the sample, relative to other data members [208] [50]. The representation of the data samples in this way reduces the dimensionality of the CARDIGAN raw data, as presented in Table 1. The phase space plots of the CA data representation for each categorised group (Healthy Control and Obesity) of the participants during the 6 clinical weeks are displayed in Figures 5.2, 5.3, 5.4, 5.5, 5.6 and 5.7.

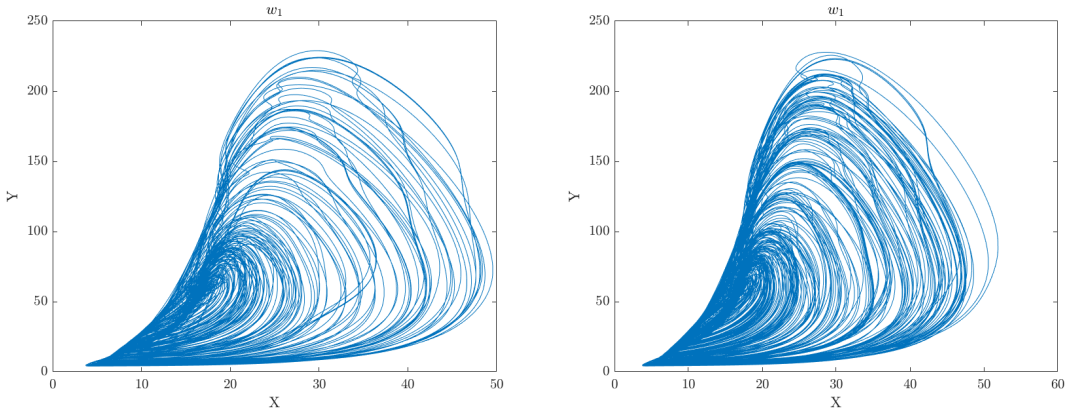


Figure 5.2: This figure shows the phase space plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_1. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

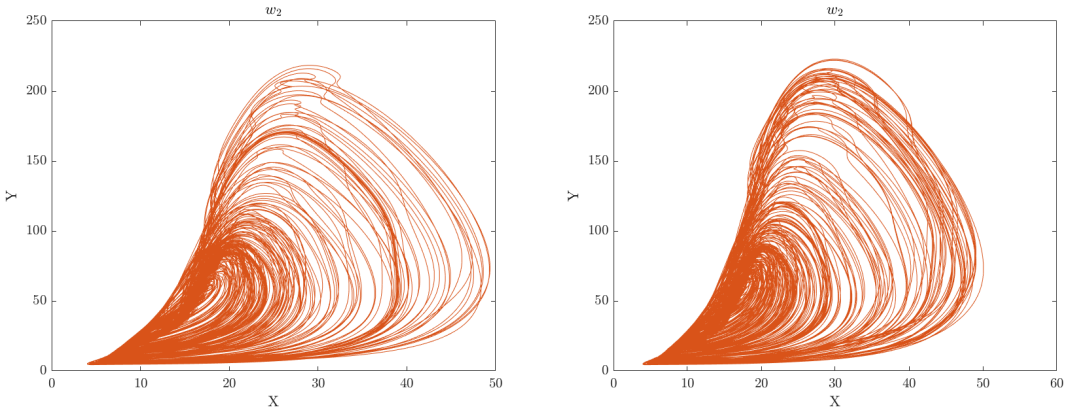


Figure 5.3: This figure shows the phase space plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_2. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

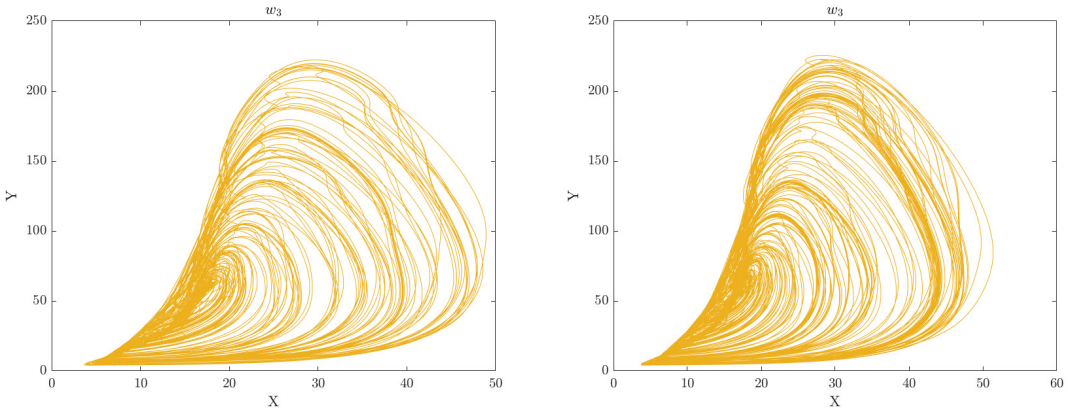


Figure 5.4: This figure shows the phase space plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_3. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

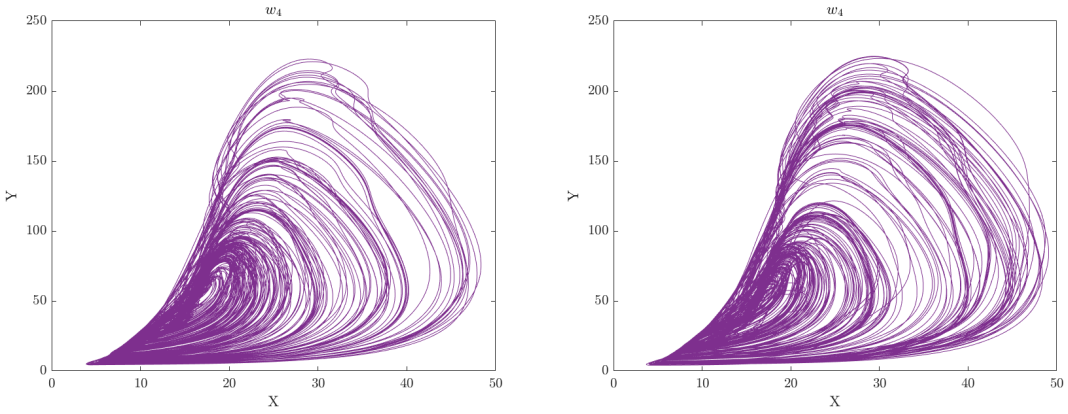


Figure 5.5: This figure shows the phase space plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_4. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

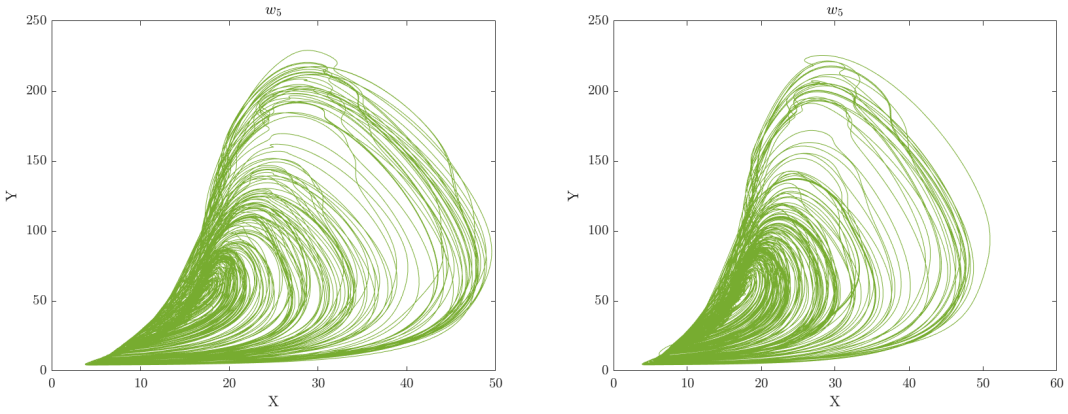


Figure 5.6: This figure shows the phase space plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_5. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

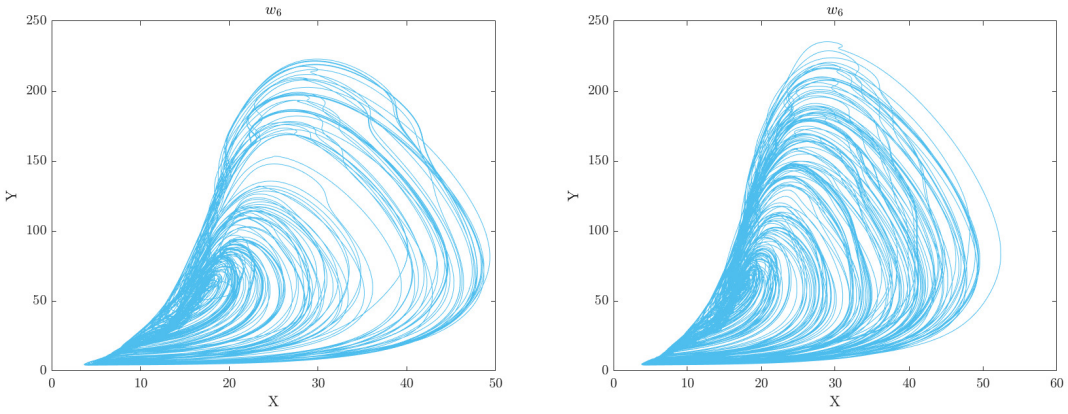


Figure 5.7: This figure shows the phase space plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_6. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

The dynamic representations of individual data samples can be utilised to compare between categories, individual data samples, and track the progression over time of the same individual. In this thesis, these properties are employed to illustrate how obese patients' gait progresses over the weeks of treatment by comparing it within their category and against normal healthy gait. Healthy gait exhibits a considerable amount of variability, which enables individuals to move without physical restrictions, as evidenced by their changing behavior during the trials. Obese children, on the other hand, experience more limitations in varying their gait, which is difficult to discern from the data due to its high nonlinearity. The analysis of the CA phase plots, presented in Figures 5.2, 5.3, 5.4, 5.5,5.6 and 5.7, demonstrates that obese individuals exhibit a gait pattern that is slower and more labored in comparison to healthy individuals. Additionally, their base of support is wider and steps are shorter. The additional weight carried by obese individuals leads to increased stress on joints and muscles, resulting in stiffness during walking. Furthermore, obesity

is associated with an increased risk of conditions such as joint pain and arthritis, which can contribute to stiffness in the gait. Conversely, healthy individuals display a gait pattern that is smooth and fluid with a narrower base of support and longer steps and have less stiffness in their walking. It is also important to note that the gait features depicted in these figures can have an impact on the network of coupled CA oscillators. The scaling of additive connectivity strength in response to external perturbations of gait can either enhance or hinder gait patterns. This is done to ensure that the input provided to each oscillator does not exceed a certain threshold and remains within the controlled domain.

5.4 Statistical Analysis of Criticality Analysis Data

5.4.1 Scatter Analysis

A scatter plot is a graphical representation of a set of data points, where each point represents an observation in a two-dimensional coordinate system. In the context of spatiotemporal gait analysis, a scatter analysis can be used to visualise the relationship between two variables that are measured during gait. For example, a scatter plot of the CA data can reveal whether there is a linear relationship between their two variables. Alternatively, the scatter plot may reveal a non-linear relationship, indicating that the relationship between gait variables is more complex. Scatter plot analysis can also be used to identify outliers or other patterns in the data, which may indicate abnormal gait patterns or other issues that require further investigation.

The interpretation of Figures 5.2 to 5.7 is presented in the distribution of CA data representation for each categorised individual orbit, as shown in Figures 5.8 to 5.13. The plots, labeled from $w_1 - w_6$, depict the variability of gait interaction events over time in the musculoskeletal system during the six-week clinical trials. As the

CA data shows a correlation between the two extracted features in 2D space, the Pearson's correlation coefficient ρ measures both the direction and strength of the dependency between them, with $0 \leq \rho \leq 1$.

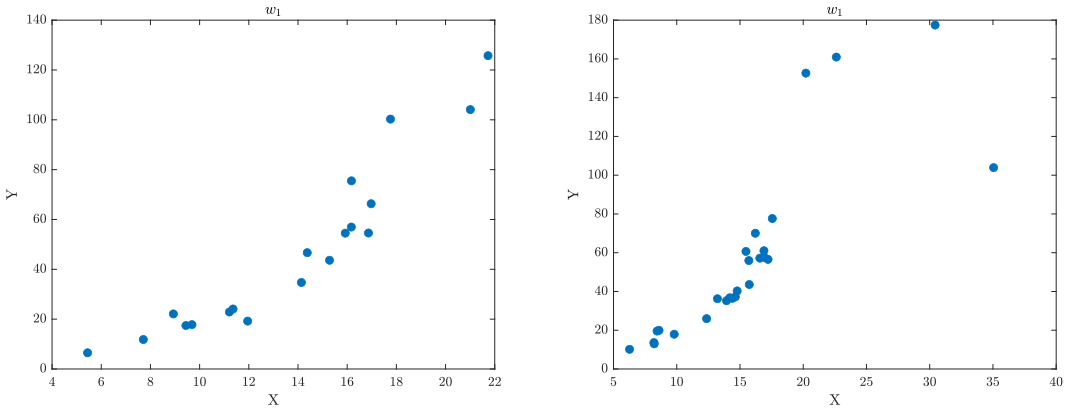


Figure 5.8: This figure shows the scatter plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_1. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

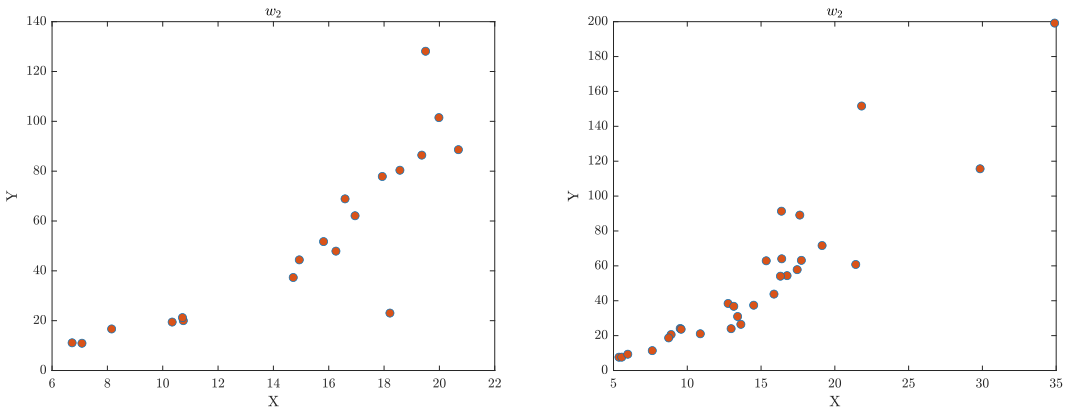


Figure 5.9: This figure shows the scatter plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_2. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

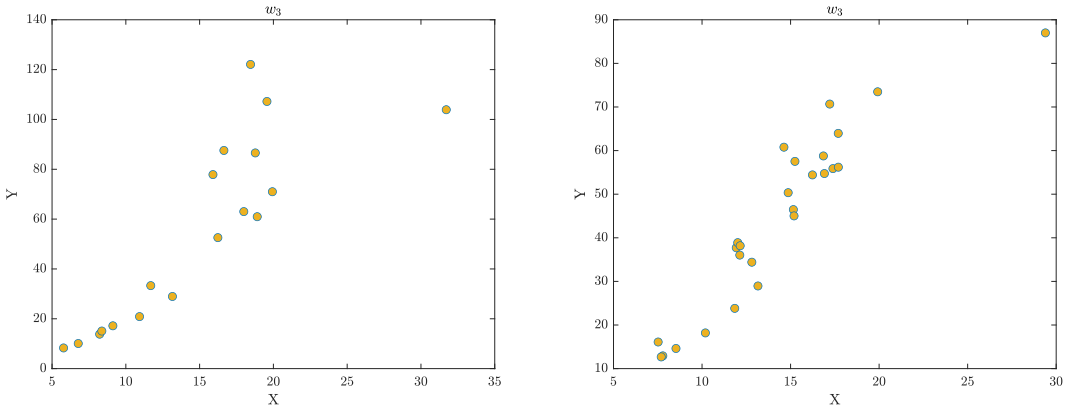


Figure 5.10: This figure shows the scatter plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_3. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

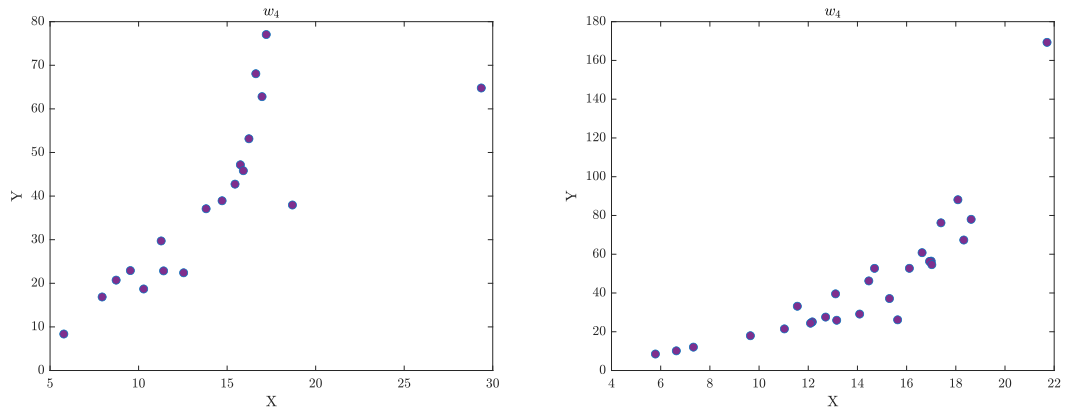


Figure 5.11: This figure shows the scatter plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_4. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

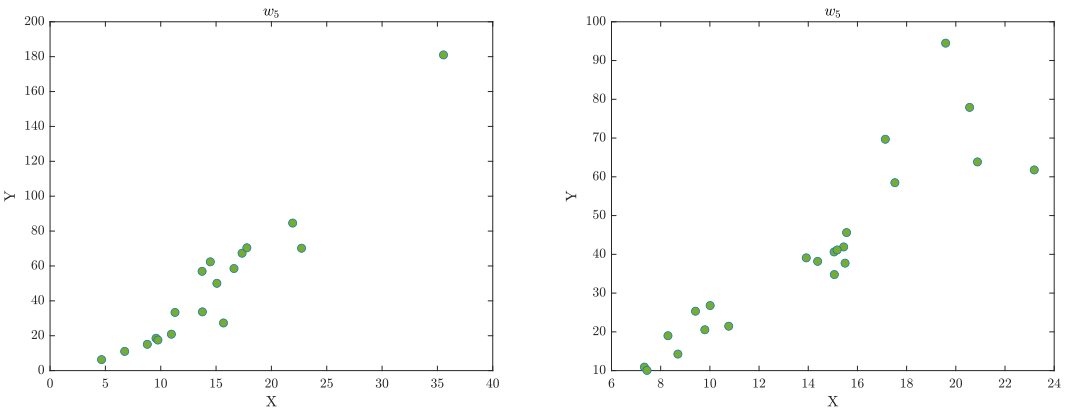


Figure 5.12: This figure shows the scatter plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_5. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

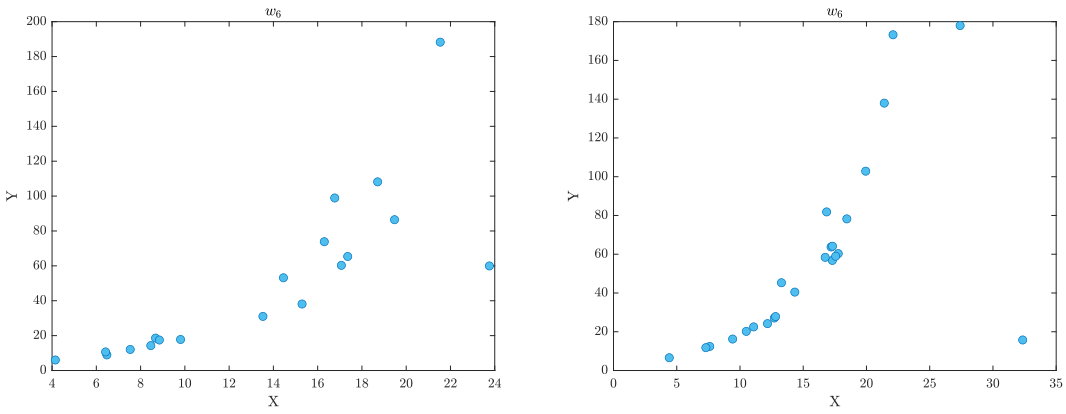


Figure 5.13: This figure shows the scatter plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_6. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

Table 5.2 displays the correlation coefficient (ρ) of six gait variables (w_1 to w_6) for two categories: Healthy Control and Obesity. The correlation coefficients for the Healthy Control group range from 0.79 to 0.95, with the highest correlation observed for w_5. The correlation coefficients for the Obesity group range from 0.64 to 0.92, with the highest correlation observed for w_3. Overall, the correlation coefficients suggest strong positive correlations between the gait variables for both groups, with slightly higher correlations observed for the Healthy Control group.

Table 5.2 provides a summary of the ρ values for each group of disorders over a 6-week period.

Table 5.2: The correlation coefficient ρ of w_1, w_2, w_3, w_4, w_5 and w_6.

Category	**Correlation Coefficient (ρ)**					
	$\boldsymbol{w_1}$	$\boldsymbol{w_2}$	$\boldsymbol{w_3}$	$\boldsymbol{w_4}$	$\boldsymbol{w_5}$	$\boldsymbol{w_6}$
Healthy Control	0.92	0.85	0.84	0.79	0.95	0.80
Obesity	0.82	0.91	0.92	0.83	0.90	0.64

Additionally, it should be highlighted that the distribution demonstrated in Figures 5.2–5.7 and Figures 5.8–5.13 quantifies the extent of variation in CA data and displays how the position of each data point changes over time relative to the others. Each data point on the $w_1 - w_6$ graphs for both healthy controls and obese individuals represents an individual participant, and each plot replicates the others over a six-week duration. To comprehend how gait behavior evolves for each participant within their respective disorder category over the course of the trial, Figures 5.14 and 5.15 show the confidence standard deviation region over the six-week period. Since the total number of cases for each disorder is identical, twenty plots labeled from p_1 to p_{20} are generated, each corresponding to a unique participant, illustrating the progression of their gait performance throughout the entire trial. Each graph distinctly shows a central line, which represents the average or mean of the $y-$axis features that corresponds to their projected features on the $x-$axis. This line indicates the shift in the position of an individual's gait, while the shaded region of the standard deviation is used to determine the magnitude of change in scale within the 95% confidence interval. Furthermore, it can be observed from both Figures 5.14 and 5.15 that the evolution of gait fluctuates over time, allowing for comparison between various disorder groups by visualising and mapping the development of gait against normal gait patterns.

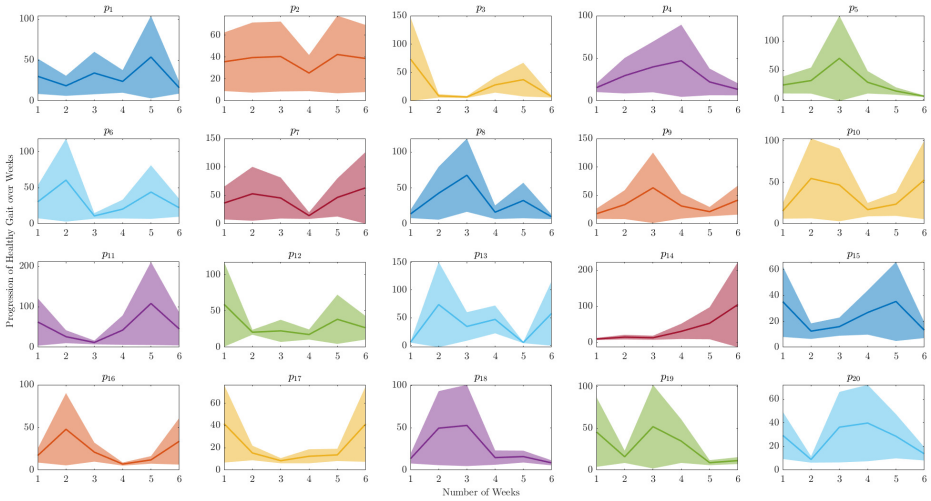


Figure 5.14: The progression of healthy control gait for each individual over a 6-week period.

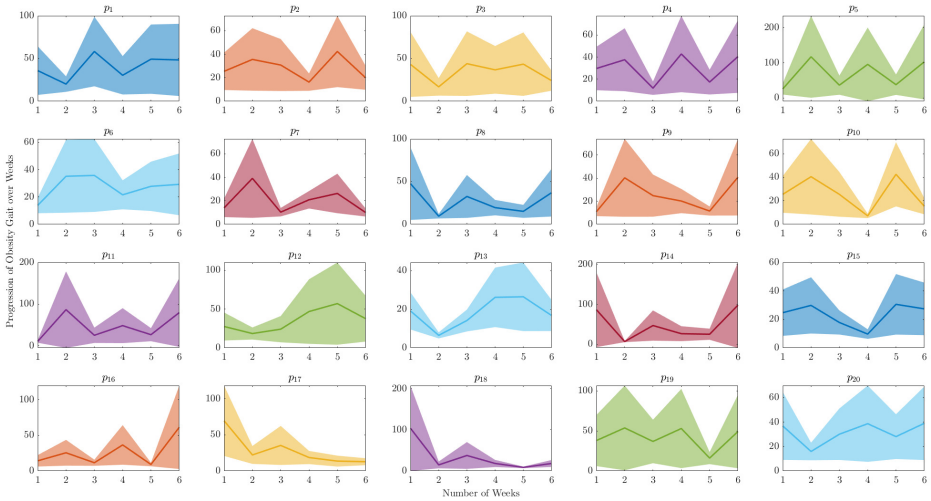


Figure 5.15: The progression of obesity control gait for each individual over a 6-week period.

5.4.2 Spatiotemporal Analysis

The Spatiotemporal Analysis of Gait Data features provide several insights into a person's gait. These features can be used to analyse and quantify various aspects of gait, including step length, step width, stride length, cadence, and velocity. By examining these features, one can identify abnormalities or asymmetries in gait that may be indicative of certain neurological or musculoskeletal conditions. For example, changes in step length or cadence may indicate the presence of diseases or other movement disorders, while changes in stride length or velocity may be indicative of muscle weakness or joint dysfunction. In addition, these features can also be used to track changes in gait over time and evaluate the effectiveness of interventions or treatments aimed at improving gait.

By comparing the gait disorder resulting from obesity to healthy controls using the CA data, Figures 5.16– 5.18 visually demonstrate that gait disorders significantly impact gait progression. The spatiotemporal gait analysis figures also present a visual display of how data is distributed across various categories, including individuals with gait obesity disorders and healthy controls. The figures typically display the median as a central line within a box that extends from the lower to the upper quartile of the data. The whiskers extend from these figures to indicate the range of the data, with outliers represented by individual points beyond the whiskers. The figures also allow for quick and easy comparison of the distribution of spatiotemporal gait data between different categories. They can provide insight into the location and degree of variation of specific gait parameters, such as step length or velocity, among different groups. Additionally, changes in these figures over time can be used to evaluate the effectiveness of interventions or treatments aimed at improving gait in patients with gait disorders.

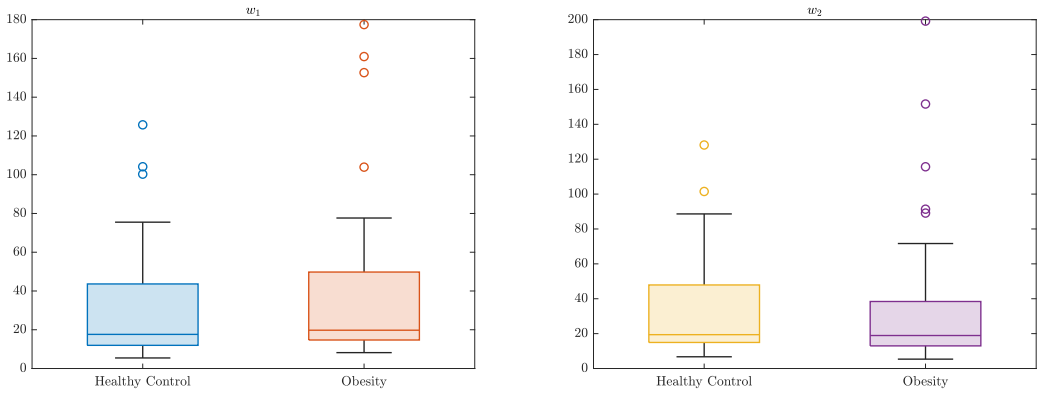


Figure 5.16: This figure shows the box plots for healthy control and obesity walk patterns that correspond to the clinical gait experiments conducted in w_1 and w_2.

Table 5.3 provides the median values of six features (w_1, w_2, w_3, w_4, w_5, and w_6) for two categories, healthy control, and obesity. The median value is a measure of central tendency that indicates the middle value of the data set. In this case, it represents the average value of the respective feature for each category. The table indicates that for features w_1, w_4, w_5, and w_6, the median value is lower for healthy control than obesity, while for features w_2 and w_3, it is higher for healthy control than obesity. This suggests that there are differences in spatiotemporal gait features between healthy control and obesity, and these differences could be used to distinguish between the two categories.

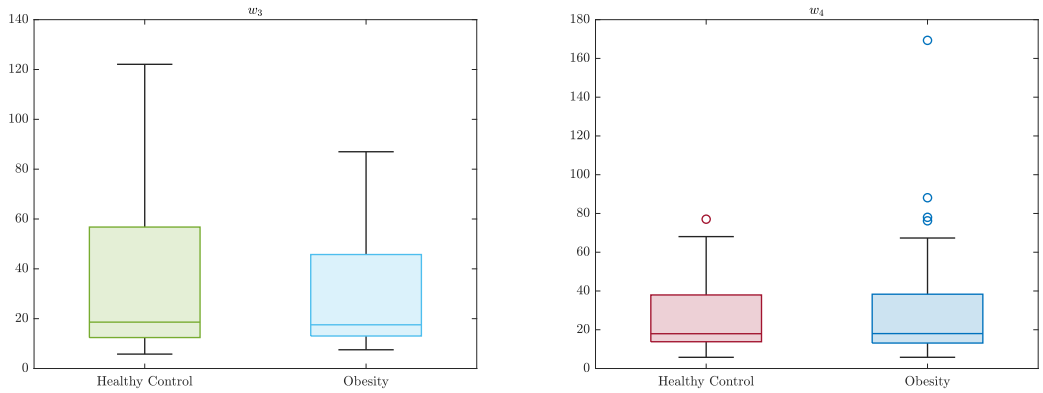


Figure 5.17: This figure shows the box plots for healthy control and obesity walk patterns that correspond to the clinical gait experiments conducted in w_3 and w_4.

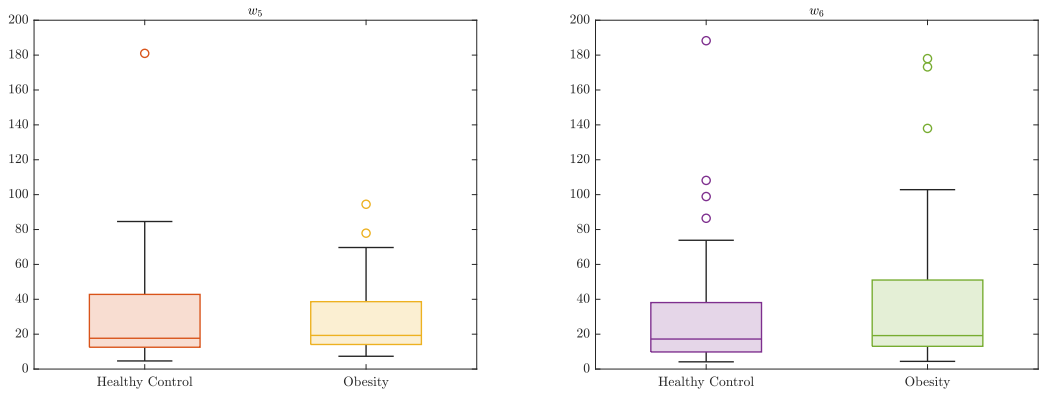


Figure 5.18: This figure shows the box plots for healthy control and obesity walk patterns that correspond to the clinical gait experiments conducted in w_5 and w_6.

Table 5.3: The median of w_1, w_2, w_3, w_4, w_5 and w_6.

Category	**Median**					
	$\boldsymbol{w_1}$	$\boldsymbol{w_2}$	$\boldsymbol{w_3}$	$\boldsymbol{w_4}$	$\boldsymbol{w_5}$	$\boldsymbol{w_6}$
Healthy Control	17.6	19.4	18.6	17.9	17.6	17.2
Obesity	19.7	18.9	17.5	18	19.2	19.1

5.4.3 Histogram Analysis

Figures 5.19–5.24 illustrate the histogram of individual features and their respective frequencies, providing unique statistical characteristics of the distribution of CA data. The distribution of the 20 obese participants showed a higher average mean compared to the other healthy control individuals. The right-skewed distribution of each category is evident from both Figures 5.19–5.24, with some data points acting as outliers on the right tail of the data stream due to increased variability, higher skewness, and the lower bound of data features being almost one-sided. These outliers occur due to unintentional changes in trial settings that affect the direction of distribution. Figure 5.19– 5.24 of w_3 stands out for its absence of outliers, with the average mean of obesity cases increasing by approximately 4.5% compared to the healthy control group.

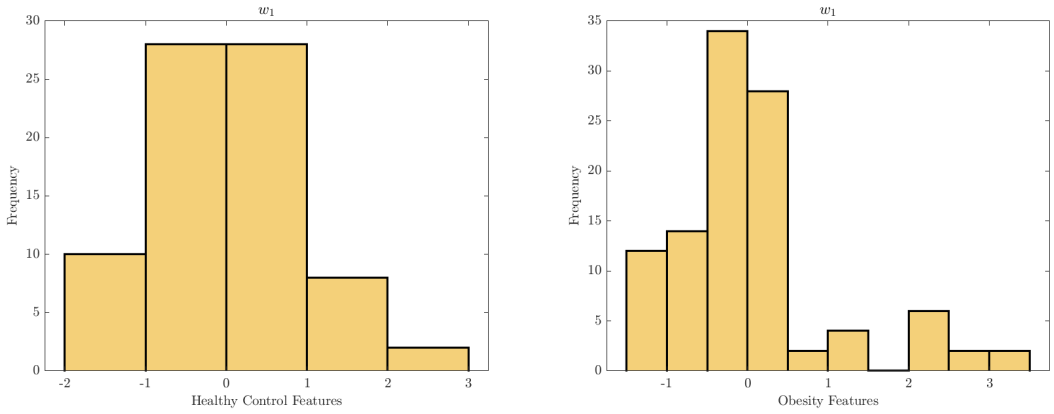


Figure 5.19: This figure shows the histogram plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_1. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

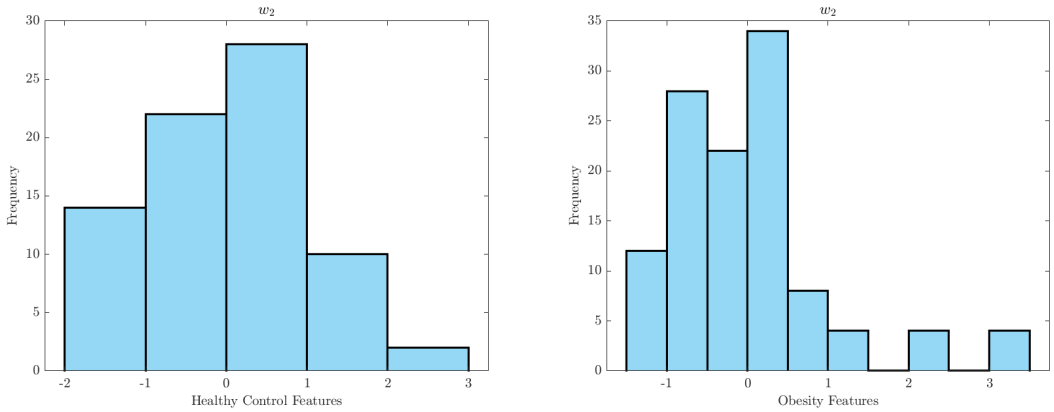


Figure 5.20: This figure shows the histogram plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_2. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

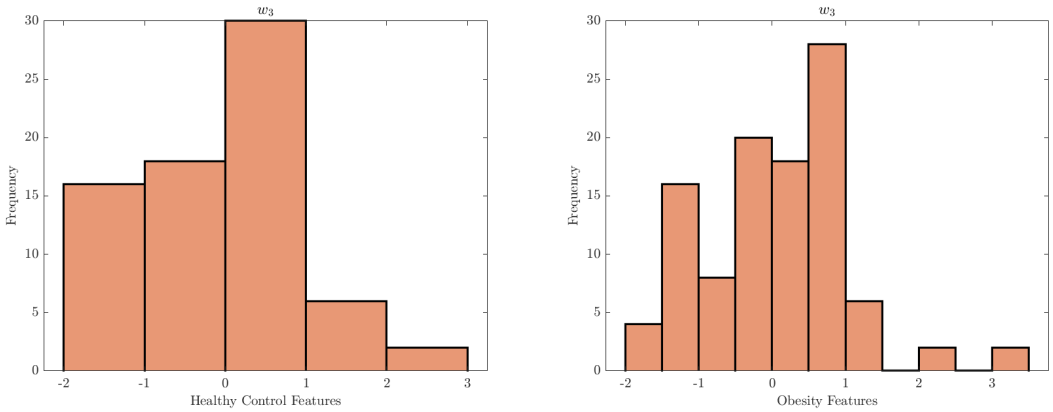


Figure 5.21: This figure shows the histogram plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_3. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

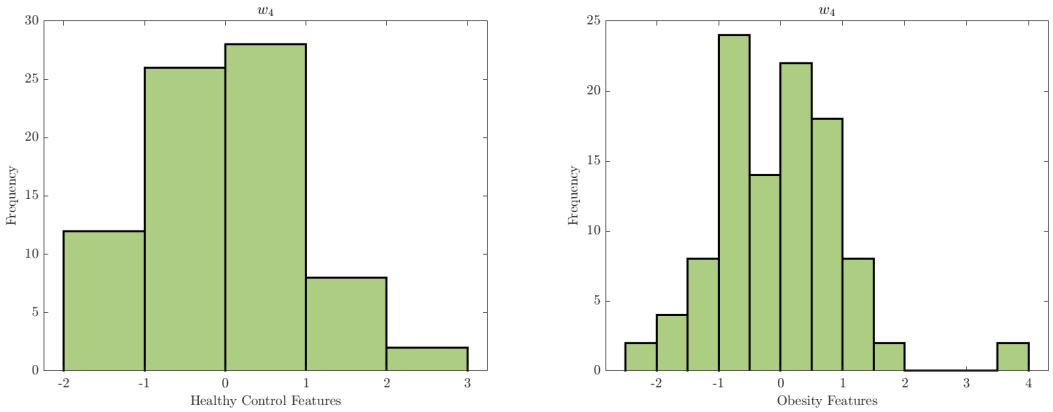


Figure 5.22: This figure shows the histogram plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_4. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

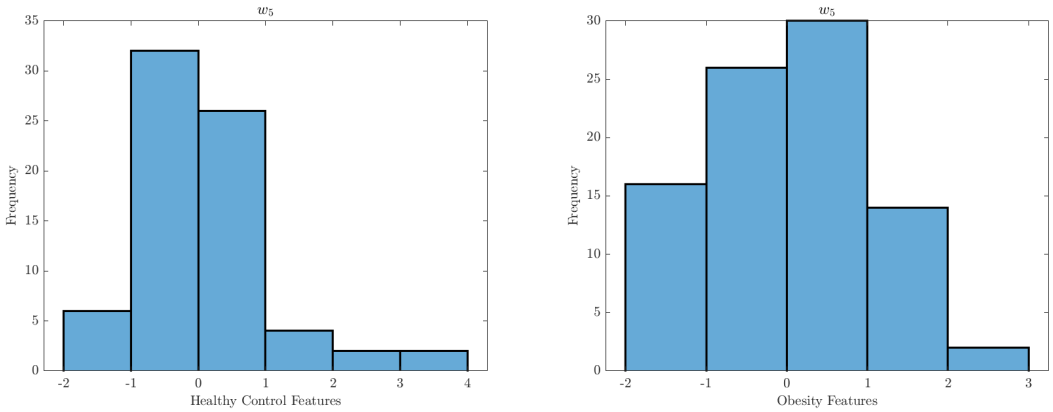


Figure 5.23: This figure shows the histogram plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_5. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

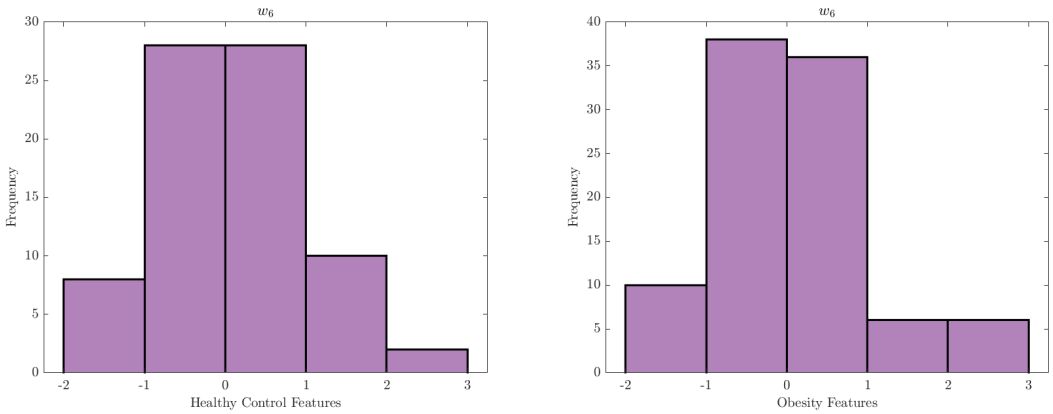


Figure 5.24: This figure shows the histogram plots for healthy control and obesity walk patterns that correspond to the clinical gait experiment conducted in w_6. On the left is the healthy control walk patterns portrait and on the right are the obesity walk patterns.

Table 5.4 presents the means of six variables (w_1, w_2, w_3, w_4, w_5, and w_6) for

two groups of participants: Healthy Control and Obesity. The means of w_2, w_3, and w_5 were higher in the Healthy Control group than in the Obesity group. On the other hand, the means of w_1, w_4, and w_6 were higher in the Obesity group than in the Healthy Control group. The results suggest that there might be significant differences between the two groups in terms of their weight-related factors. However, it is important to note that these differences might be due to various confounding factors, such as age, gender, physical activity, and dietary habits. Table 5.4 provides a summary of the mean values for each disorder category throughout the six-week period, as shown below:

Table 5.4: The mean of w_1, w_2, w_3, w_4, w_5 and w_6.

Category	**Mean**					
	$\boldsymbol{w_1}$	$\boldsymbol{w_2}$	$\boldsymbol{w_3}$	$\boldsymbol{w_4}$	$\boldsymbol{w_5}$	$\boldsymbol{w_6}$
Healthy Control	30.7	33.7	34.6	26.4	31.9	32.2
Obesity	37.8	33.5	29.2	30	27.3	36.7

5.5 The Proposed SVM Classifier

The proposed SVM model capitalises on the utilisation of the Kernel property outlined in Section 4.3.2, owing to its capability of converting this nonlinear dynamic gait data into a novel feature space, thereby streamlining the classification process. The DGAS software was utilised to extract features from the CARDIGAN dataset, which were then inputted into the CA analysis model. The controlled CA model extracts salient features, represented by f of equation 3.11 on the x-axis and m of equation 3.10 on the y-axis, that serve as inputs for the SVM algorithm, enabling the

assessment of classification performance under varying settings. The proposed SVM classifier underwent rigorous training and testing utilising these extracted features, specifically for both the healthy control and obese groups. The primary objective of this experiment was to evaluate the capacity of the SVM to discern gait patterns between the two groups (obesity and healthy control). Algorithm 1 outlines the implementation procedures of the proposed SVM algorithm. The SVM algorithm's performance was evaluated using several metrics, including the confusion matrix, receiver operating characteristics, and other measures, which were analysed in Appendix A.

5.6 The Proposed SVM Training

The CARDIGAN dataset for this study comprises 20 data points per gait subject, resulting in a total of 40 data samples for both the healthy control and obesity subjects. The training and testing of the Kernel SVM classifier employed a 50% splitting ratio. In preparation for the SVM algorithm training process, all CA gait features were normalised using the z-score method, resulting in a zero mean and unity standard deviation for the data samples. The Kernel SVM algorithm, as outlined in Algorithm 1, was implemented utilising MATLAB to examine the influence of Kernel characteristics, specifically the regularisation parameter C and control width σ, on the classification performance of gait data. The generalisation performance of the trained Kernel SVM was evaluated by measuring the prediction accuracy for each model, analysing the confusion matrix, plotting the Receiver Operating Characteristics (ROC) curve, and computing the area under the ROC curve. These evaluations are discussed in the subsequent sections of this thesis.

5.7 Simulation Results

In order to more accurately assess the efficacy of the suggested approach described earlier, the subsequent performance metrics will be employed:

5.7.1 Confusion Matrix

When dealing with a binary classification problem, the dataset labels are restricted to two possibilities: positive or negative. The decision of the SVM classifier in this scenario can be visualised using a structured contingency table called a confusion matrix. This matrix comprises actual and predicted binary classification data and is helpful in evaluating the classification algorithm's ability to differentiate between different categories. The confusion matrix is characterised by four primary metrics:

- **True Positive** (TP): indicates the number of positive samples (i.e., healthy control participants) that are correctly identified as positive (i.e., healthy).
- **False Positive** (FP): denotes the number of negative samples (i.e., obese participants) that are wrongly identified as positive (i.e., healthy).
- **True Negative** (TN): represents the number of negative samples (i.e., obese individuals) that are correctly identified as negative (i.e., obese).
- **False Negative** (FN): denotes the number of positive samples (i.e., healthy control participants) that are wrongly identified as negative (i.e., obese individuals).

A classifier's additional performance metrics include accuracy, precision, F_1-Score, Recall or Sensitivity or True Positive Rate (TPR), Specificity or True Negative Rate (TNR), and False Positive Rate (FPR). These metrics are computed based

on the values of TP, FP, TN, and FN as mentioned earlier. The following are the definitions of these metrics:

- **Accuracy**: measures the proportion of accurate predictions and reflects the general level of success. This ratio is calculated by dividing the number of correctly classified labels $(TP + TN)$ by the total number of data samples $(TP + TN + FP + TN)$, expressed mathematically as:

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

- **Precision**: is used to measure the ratio of accurately predicted positive labels indicating healthy control status to the total labels predicted to be in good health, expressed as:

$$Precision = \frac{TP}{TP + FP}.$$

- **Recall (TPR)**: is used to measure the proportion of healthy data samples with correctly predicted positive labels for healthy control status, defined as:

$$Recall = \frac{TP}{TP + FN}.$$

- **F_1−Score**: maintains a balance between precision and recall, calculated as follows:

$$F_1 - Score = \frac{2 \times Precision}{Precision + Recall}.$$

- **True Negative Rate (TNR)**: the negative predictive value for obesity is determined by calculating the ratio of correctly predicted negative labels among all samples with obesity. This value is expressed as:

$$TNR = \frac{TN}{TN + FP}.$$

- **False Positive Rate (FPR)**: the misclassification rate, which represents the ratio of healthy control data that are incorrectly labeled among all obesity data samples, is expressed as follows:

$$FPR = \frac{FP}{FP + TN}.$$

The objective of the aforementioned metrics is to assess the effectiveness of the proposed SVM algorithm in detecting abnormal gait patterns in various individuals. Tables A.1–A.12 in Appendix A present the performance metrics of the SVM classifier for each of the six-week trials, based on CA data features and different values of σ and C (0.1, 1, and 10). For instance, Tables A.3– A.4 of Appendix A report the performance of SVM classification for w_2 walk patterns with different values of σ and C. Table A.3 reports the results of classification at $\sigma = 10$ and for C values 0.1, 1, and 10. Also, Table A.4 reports the results of classification at $\sigma = 0.1$ and $\sigma = 1$ for the same C values. Comparing the two tables, we can observe that the performance of SVM classification is dependent on the values of σ and C. At $\sigma = 10$, the SVM model performs poorly as all TP, FP, FN, and TN values are low. However, when σ is reduced to 0.1 or 1, the model performs better with higher TP, FP, FN, and TN values. Also, the accuracy and AROC increase for $\sigma = 0.1$ and $\sigma = 1$ compared to $\sigma = 10$. The combinations of σ and C values demonstrated that relying solely on accuracy is inadequate for assessing the SVM classifier's overall effectiveness.

5.7.2 Receiver Operating Characteristics (ROC) Curve

The confusion matrix parameters include true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). These parameters are used

to calculate other measures, such as precision, recall, F_1-score, specificity, and the area under the ROC (AROC). The ROC curve is a graph that shows the trade-off between TPR and FPR for different classification thresholds. The ROC curve illustrates how sensitivity (TPR) and the complementary probability of specificity ($\approx 1 - TNR$) vary across a range of threshold levels from 0 to 1. Figures 5.25– 5.27 display the ROC curves for the optimal combination of σ and C values that resulted in the highest accuracy throughout the entire experiment.

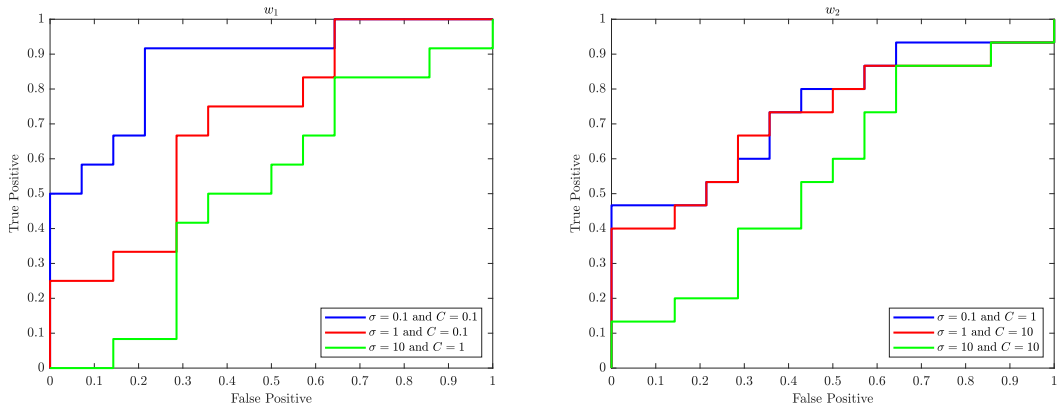


Figure 5.25: ROC (Receiver Operating Characteristic) curves of w_1 and w_2 show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

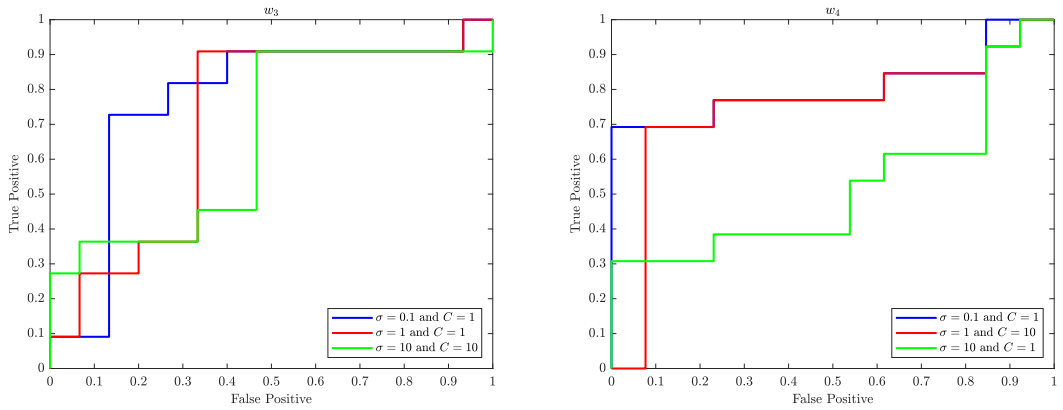


Figure 5.26: ROC (Receiver Operating Characteristic) curves of w_3 and w_4 show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

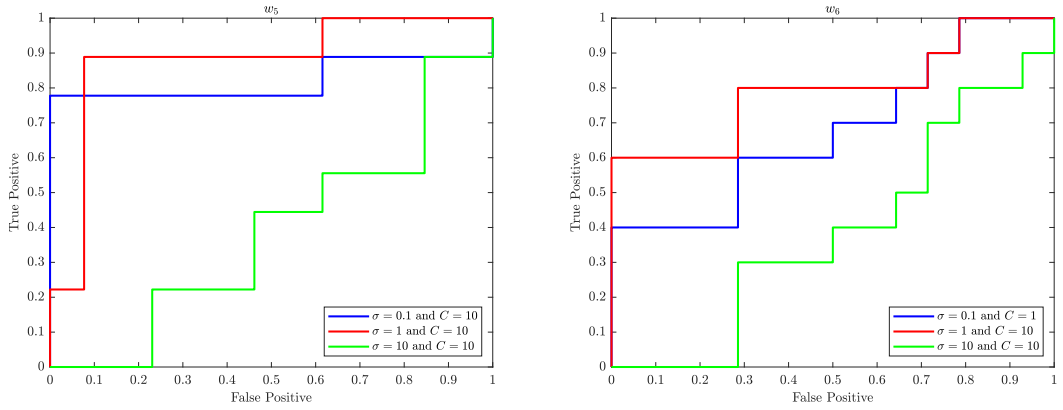


Figure 5.27: ROC (Receiver Operating Characteristic) curves of w_5 and w_6 show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

It is obvious from these ROC plots (Figures 5.25– 5.27) that, in the context of SVM, the parameter C controls the trade-off between maximising the margin and minimising the misclassification error. A smaller value of C results in a wider margin,

but with more misclassifications, while a larger value of C results in a narrower margin but with fewer misclassifications.

The parameter σ is used to control the width of the Kernel Gaussian function that is used to map the input data into a higher dimensional space, where a linear boundary can be found. A larger value of σ results in a wider Gaussian function, which leads to a softer decision boundary and a higher bias, while a smaller value of σ results in a narrower Gaussian function, which leads to a harder decision boundary and a higher variance.

When σ is small, the decision boundary is more sensitive to the input data, which can lead to overfitting. On the other hand, when σ is large, the decision boundary is less sensitive to the input data, which can lead to underfitting. Therefore, the value of σ has an impact on the generalisation performance of the SVM.

A good value for C and σ are the one that balance the trade-off of bias and variance, that is a good balance between overfitting and underfitting.

The ROC curves shown in Figures 5.25– 5.27 performed well with $\sigma = 0.1$ and 1 for various values of C of the SVM are likely because the classifier is able to find a good balance between overfitting and underfitting by adjusting the value of C and σ which in turn results in a good performance.

5.7.3 Area Under the ROC Curve (AROC)

The Area Under the Curve (AUC) is a measure of the overall performance of a binary classifier that considers all possible classification thresholds. A higher AROC value indicates better performance. The ROC curve and AUC are used to evaluate the performance of binary classifiers, especially when the classes are imbalanced. The AROC value ranges from 0 to 1, with 0.5 indicating random guessing and 1 indicating perfect classification.

In the given Tables A.1– A.12 of Appendix A , the AROC values are reported

for each combination of σ and C. We can observe that the AROC values increase with increasing values of C for $\sigma = 0.1$ and $\sigma = 1$. This indicates that increasing C improves the overall performance of the classifier. However, for $\sigma = 10$, the AROC values are low for all values of C, indicating that the classifier performs poorly. Overall, the results show that the performance of SVM classification is highly dependent on the values of hyperparameters σ and C. A small value of σ improves the performance of the model, while a large value of σ results in poor performance. Moreover, the results highlight the importance of selecting appropriate values of hyperparameters for machine learning models.

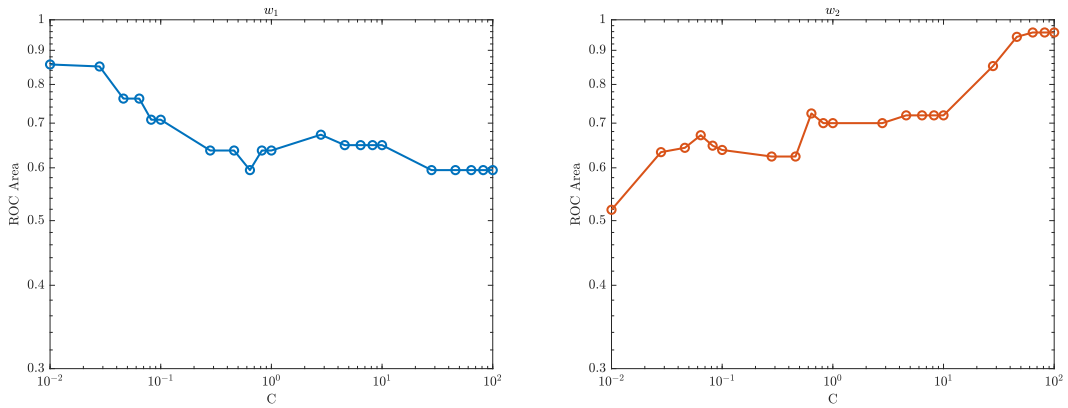


Figure 5.28: The AROC versus the regularisation parameter C of w_1 and w_2.

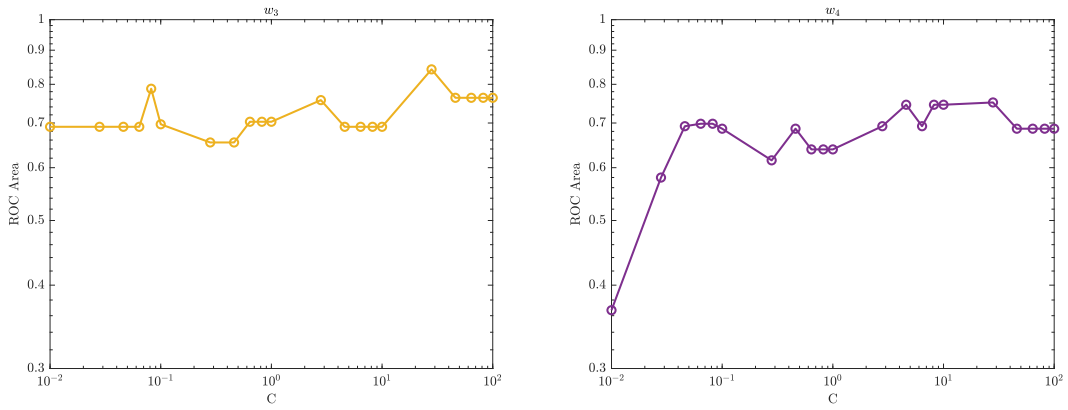


Figure 5.29: The AROC versus the regularisation parameter C of w_3 and w_4.

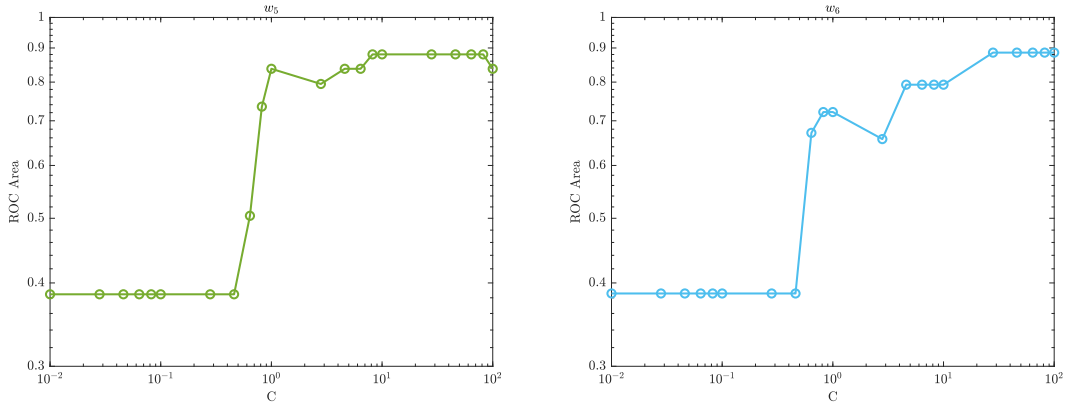


Figure 5.30: The AROC versus the regularisation parameter C of w_5 and w_6.

The area under the ROC curve, also known as the AUC (Area Under the Curve), is a measure of the performance of a binary classifier. In the context of SVM, the AUC represents the ability of the classifier to distinguish between the positive and negative classes. A higher AUC value indicates that the classifier is able to correctly classify more instances of the positive class as positive, while also correctly classifying more instances of the negative class as negative. An AUC of 1.0 represents a perfect

classifier, while an AUC of 0.5 represents a classifier that performs no better than random guessing.

Figures 5.28– 5.30 show how the performance of the SVM model changes as the regularisation parameter strength C is varied. The regularisation parameter C controls the trade-off between maximising the margin (the distance between the decision boundary and the closest training instances) and minimising the classification error. When C is small, the model will focus more on maximising the margin, which can lead to a simpler decision boundary but also a higher classification error. As C is increased, the model will focus more on minimising the classification error, which can lead to a more complex decision boundary but also lower classification error.

From Figures 5.28– 5.30, if the AUC increases as C increases, it means that the model's performance is improving as the regularisation strength C increases. This may suggest that the model was underfitting the data when C was small and that increasing the regularisation strength helped to improve the model's performance. On the other hand, if the AUC decreases as C increases, it means that the model's performance is worsening as the regularisation strength increases. This may suggest that the model was overfitting the data when C was small and that increasing the regularisation strength C caused the model to become too simplistic and lose important information from the data.

The optimal value of C is where the AUC is the highest, this is the sweet spot where the model is able to balance the trade-off between maximising the margin and minimising the classification error in a way that leads to the best classification performance.

5.7.4 Classification Decision Boundary

The decision boundary of an SVM classifier is determined by the support vectors, which are the data points closest to the boundary. The parameters C and σ, also

known as the regularisation and kernel parameters respectively, control the width of the margin and the shape of the decision boundary. For instance, When σ is set to 0.1 and C is set to 1, the decision boundary will be complex and more influenced by the individual data points. The width of the margin will be relatively small and the classifier will be more sensitive to the presence of outliers, as the algorithm tries to minimise misclassification errors. Moreover, When σ is set to 0.1 and C is set to 10, the decision boundary will be even more complex as C has a greater influence on the decision boundary. The width of the margin will be even smaller and the classifier will be even more sensitive to outliers. Furthermore, When σ is set to 0.1 and C is set to 0.1, the decision boundary will be relatively simple as C has a much smaller influence on the decision boundary. The width of the margin will be relatively large and the classifier will be less sensitive to outliers. The classification boundaries of the SVM model are depicted in Figures 5.31 to Figures 5.33 using the best classification parameters, enabling the model to accurately categorise participants into the appropriate group.

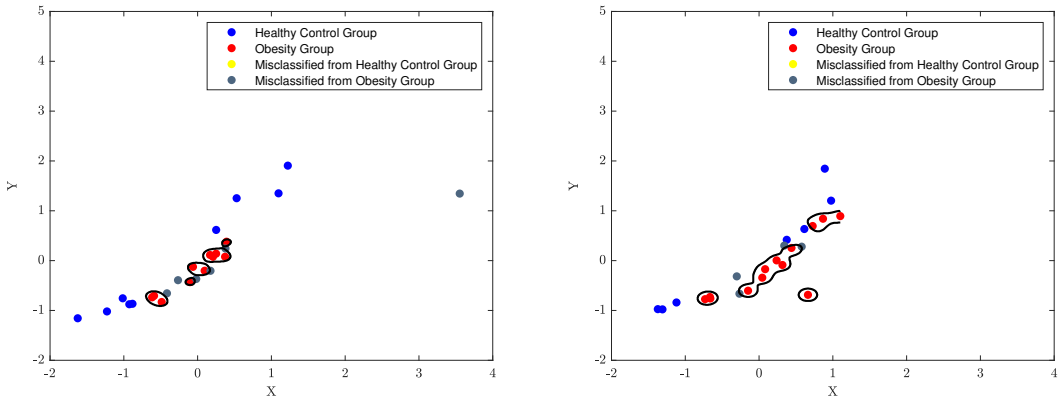


Figure 5.31: The boundary that separates the healthy control walk patterns from the obesity patterns in an SVM model, with $\sigma = 0.1$ and $C = 0.1$ for w_1 and with $\sigma = 0.1$ and $C = 1$ for w_2.

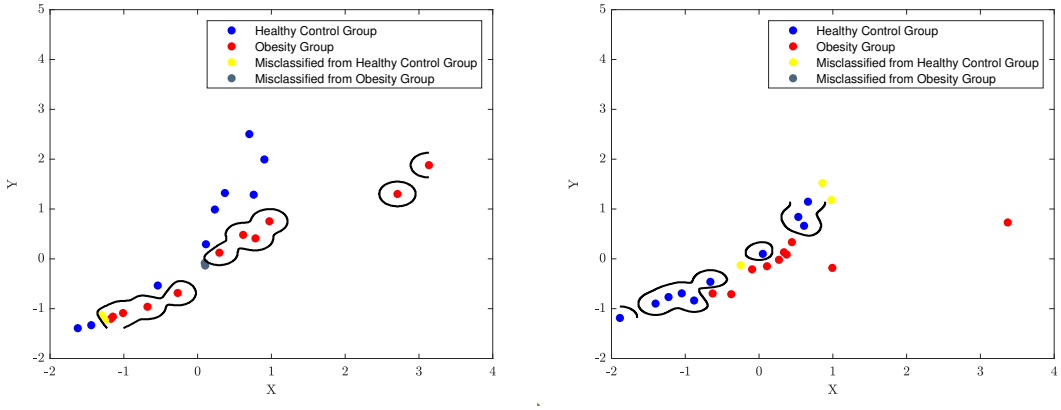


Figure 5.32: The boundary that separates the healthy control walk patterns from the obesity patterns in an SVM model, with $\sigma = 0.1$ and $C = 1$ for w_3 and w_4.

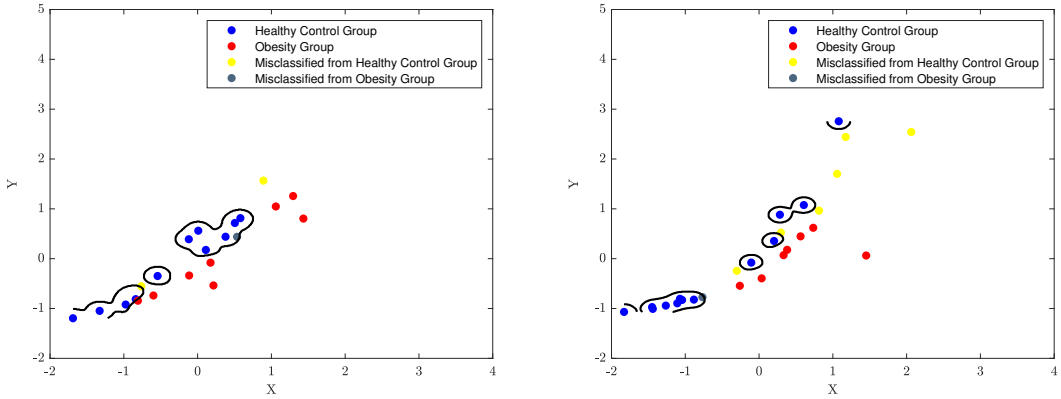


Figure 5.33: The SVM decision boundary that separates the healthy control walk patterns from the obesity patterns in an SVM model, with $\sigma = 0.1$ and $C = 10$ for w_5 and with $\sigma = 0.1$ and $C = 1$ for w_6.

5.7.5 Mean Square Error Rate and Standard Deviation Test

Figure 5.34 displays the average mean square error (MSE) and variations observed in clinical data trials conducted over a 6-week period. The average MSE for

each participant was obtained by computing the mean of the incorrect data patterns obtained for each curve depicted in Figures 5.28– 5.30. The average MSE for w_6 was found to be below 25% with minimum variation rates, indicating accurate classification of gait disorder patterns. However, the average MSE and standard deviation for individuals' data collected during w_2 and w_5 were considerably high in comparison.

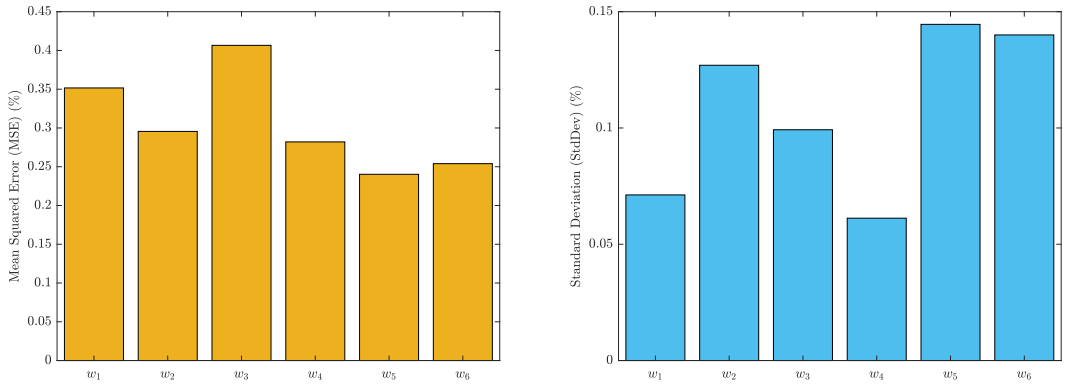


Figure 5.34: This figure shows the average mean square error (MSE)(%) (on the left) and the average standard deviation (%) (on the right) of the SVM classification performance for each individual piece of data.

The proposed SVM model's performance was evaluated as noticed in Appendix A, and it was found that the optimal generalisation performance was achieved with the best classification parameters ($\sigma = 0.1$ and $C = 0.1$). The 6-week evaluation of the SVM shows fluctuating accuracy in classifying participants into the healthy control and obesity groups. A high accuracy reflects consistent participant characteristics, facilitating accurate classification by the SVM, whereas a low accuracy indicates high variability in participant characteristics, making classification challenging. The results also revealed that the accuracy is highest in the first week, but drops dramatically in the third week, with slight improvement in the fourth week, followed by decreases in the fifth and sixth weeks. The fluctuation in the accuracy of the SVM

model during the 6 weeks period, despite the uniform diet and exercise regimen followed by the participants, could be attributed to various reasons such as variations in compliance levels, where some participants may have been more diligent in adhering to the regimen than others, leading to different classifications into healthy control or obesity groups. Other factors include individual differences such as genetics, medical history, and personal habits, measurement inaccuracies, and changes in the participants characteristics over time, even when following the prescribed diet and exercise regimen. For instance, participants stress levels or health status could impact their classification into healthy control or obesity groups.

5.8 Summary

The aim of this chapter was to investigate the utility of the CA technique in detecting gait abnormalities in children with metabolic disorders, specifically obesity. The study hypothesised that by incorporating gait features obtained from the CA biomechanical system into the SVM classifier, the detection rate of gait disturbances in children with obesity would be increased to 78.2% on average and their gait quality would be enhanced. Hence, the main objective of the research in this paper has been achieved, demonstrating that CA is an effective approach for multivariate data representation and dimension reduction.

The effectiveness of the proposed approach was assessed using the confusion matrix, receiver operating characteristics (ROC), and area under the ROC metrics. The combination of the CA method and the kernelised properties of the SVM model produced better results. These findings suggest that the proposed SVM model with CA support can be a reliable tool for identifying dynamic disruptions in biological data patterns, providing great potential for clinical diagnosis and rehabilitation.

The limited sample size of the CARDIGAN dataset, which consisted of only 40 samples, used in this study, could impact the performance of the SVM classification

process. However, our previous research [40] applied the CA methodology with a larger dataset of 6000 data samples and achieved an impressive 94% accuracy.

Despite this constraint, our results hold promise and offer various opportunities for further exploration, such as investigating how the CA method can be combined with other machine learning models to increase accuracy and performance. This could involve incorporating CA with other unsupervised methods such as dimensionality reduction, or with supervised methods such as neural networks or regression. The outcomes from CA can also serve as inputs or features for other models. The objective of this integration is to harness the strengths of different techniques and overcome their limitations.

Overall, the criticality analysis approach presented in this thesis was demonstrated as a valuable tool for representing gait data and improving medical outcomes. The detection of gait disorders is a complex and interdisciplinary process that can benefit from the integration of advanced technologies. By utilising a combination of criticality analysis and supervised machine learning methods, such as the SVM classifier, this research has the potential to significantly advance the field of gait analysis and lead to more precise diagnoses, improved patient outcomes, and empowered individuals in managing their health.

Chapter 6

Detecting Dynamic Disturbances in Human Gait

6.1 Introduction

Disturbances related to human gait are a prevalent concern, particularly among elderly individuals, which can be caused by reduced mobility resulting from sensory deficits, medical procedures, and falls caused by medication. These factors significantly affect quality of life and can increase morbidity and mortality rates [209]. This chapter presents a support vector machine learning-based approach for classifying gait disturbances. By analysing clinical data collected from multiple individuals, this method can identify abnormalities in human gait. Additionally, the KNN algorithm was utilised to distinguish normal and abnormal gait patterns, with further details provided in this chapter.

6.2 Methodology

This section delineates the methodology framework proposed, which comprises of four key stages: data collection, data processing, feature extraction technique, and the deployment of the SVM classifier. The graphical representation of this methodology is presented in Figure 6.1.

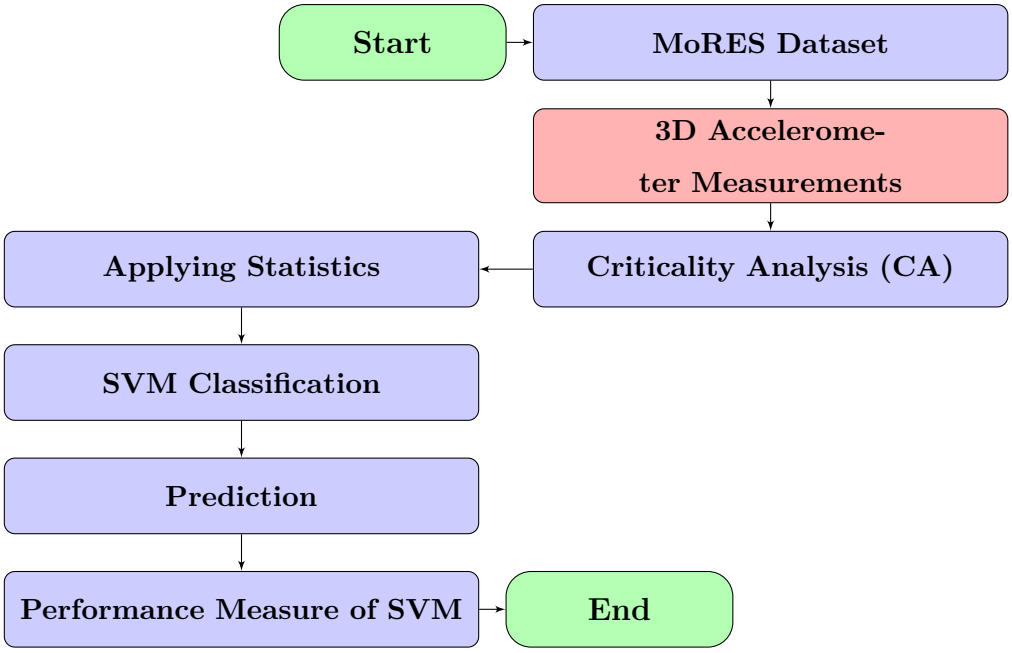


Figure 6.1: The flowchart of MoRES proposed methodology.

6.3 MoRES Dataset

Adult participants with both healthy normal walk and disturbed walk, known as strapped walk, were invited to participate in a provisional clinical study to support the Criticality Analysis of Gait in Adults (CAGA) project at the Movement, Occupational and Rehabilitation Sciences (MoRES) centre at Oxford Brookes University in the UK. All participants provided informed consent prior to participation. The

clinical study aimed to evaluate the ability of participants to walk independently as a primary means of mobility, identify the underlying causes of their health conditions, and assess their level of motivation to follow the program during the study. The psychological measurements were carried out by a medical professional prior to and during the walking exercise. Additionally, a group of healthy individuals, without peripheral injury or other conditions that may affect mobility without assistive devices, was recruited for the trial.

To obtain a feasible outcome, the five participants were instructed to walk back and forth on a flat surface for six seconds while their walking control, stride frequency, and step length were monitored. Critical changes were observed as participants changed their speed, effectively testing their mobility.

The procedures governing the operation of data collection are outlined below:

(1) The IMU movement sensor unit is employed to gather gait data from each participant, while the Axivity AX3 accelerometer, worn on the non-dominant wrist, is utilised to capture the participant's physical activity.

(2) The AX3 accelerometer is linked to a USB hub on a laptop using a Micro USB cable for the purpose of transferring data. The acquired data is then saved internally onto a memory device with a storage capacity of one Terabyte (1TB) in the form of a raw binary file.

(3) The AX3 OMGUI software, installed on a laptop, enables access to the collected data. Instructions for software installation and data recording configuration can be found in the documentation provided in [49].

(4) The Faculty of Health and Life Sciences at Oxford Brookes University's Movement, Occupational and Rehabilitation Sciences center oversaw the complete data collection process. The data collected has been designated as the MoRES dataset, acknowledging the center's role as a provider of data.

(5) The MoRES dataset was made compliant with the General Data Protection Regulation (GDPR) by anonymising it and securely sharing it with Oxford Brookes University's School of Engineering, Computing, and Mathematics (ECM) as part of the research work of this thesis. The University Research Ethics Committee (UREC) officially approved the dataset through their procedures.

(6) The MoRES dataset comprises the unprocessed data collected by the 3D IMU sensory unit. It includes two sets of measurements for each participant: normal walk patterns, which are observable gait subjects, and strapped walk patterns, which are non-observable gait subjects. Both sets of data are utilised as inputs for the CA methodology explained by equations (3.10)–(3.13) to attain a multivariate representation of the 3D IMU feature measurements, reducing them to only two features. More information on this process can be found in Section 6.3.1.

6.3.1 Features Extraction

The MoRES dataset contains raw data gathered by the 3D accelerometer, gyroscope and magnetometer IMU sensor unit, comprising two sets of measurements for each participant: observable gait patterns during normal walking and non-observable gait patterns while wearing straps. Both sets of data are utilised as inputs for the CA methodology described by equations (3.10)–(3.13) to derive a multivariate representation of the 3D IMU feature measurements, reducing them to only two features: f (x-axis) and m (y-axis). By analysing the phase space portraits of both features, it is possible to identify unique changes in the structural patterns of the dynamics of human gait.

6.3.2 Criticality Analysis as a Data Representation Method

The features extracted from the 3D IMU sensor data in the MoRES dataset are used to perturb a mathematical model represented by equations (3.10)–(3.13) through the Criticality Analysis method. This method converts multivariate dynamic data into a nonlinear representation space, where each sample is characterised by a unique orbit resulting from perturbing a critical system composed of a network of nonlinear controlled oscillators. This critical network remains stable despite small input changes leading to state changes in the orbit of the entire network. The resulting trajectory is an independent depiction of the initial dataset, devoid of any scaling, and can serve as evidence of the likeness of a particular subset of features representing either an abnormal or normal walking pattern to its corresponding group in the sample set when compared to other members of the dataset [208] [50]. This representation reduces the dimensionality of the MoRES raw data into two predominant dimensionless features, namely f and m, which are plotted on the x and y axes, respectively.

The phase space plots of the CA data representation for each categorised group of participants are presented in Figures 6.2–6.6. Figure 6.7 depicts the phase plot for all individuals in each walk category, where normal walk behaviour exhibits greater variability in the phase space as compared to strapped walk patterns. This indicates that healthy individuals have an unrestricted gait without any physical restrictions, resulting in more structured patterns. The CA approach enables the detection and classification of healthy and affected gait patterns in the time domain, without requiring machine learning training. Furthermore, the dynamic representation of individual data samples can be compared and monitored over time. In this experiment, CA is used to distinguish normal walk patterns from their strapped counterparts, demonstrating the varying behaviour of participants. The normal phase plots displayed in the left side of Figures 6.2–6.7 show that external perturbations to the network of coupled oscillators can affect gait behaviour patterns due to the scaled additive con-

nectivity strength to the external perturbations, while ensuring the system remains within a controlled domain.

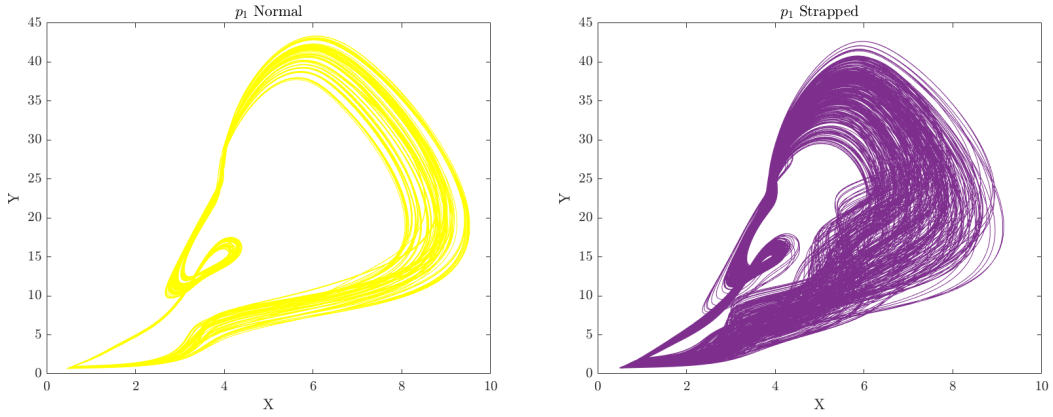


Figure 6.2: This figure shows the phase space plots for each walk pattern that correspond to individual of p_1. On the left is the normal walk patterns portrait, while the strapped patterns are on the right side of Figure 6.2.

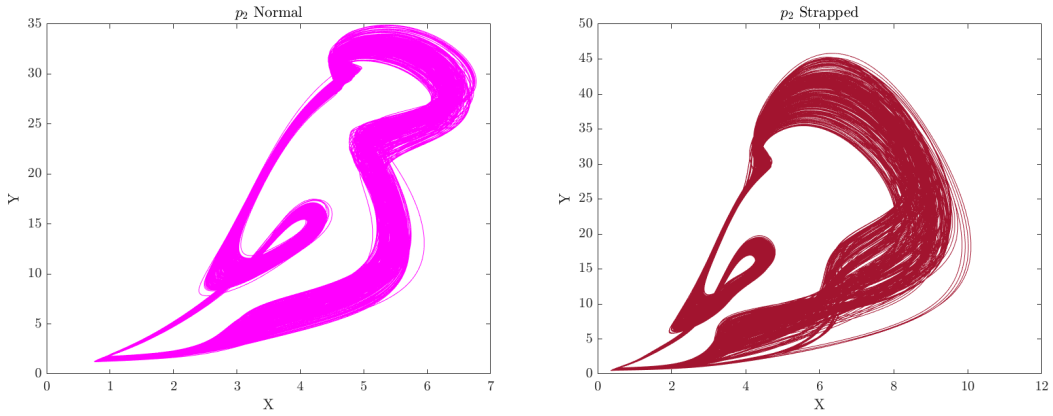


Figure 6.3: This figure shows the phase space plots for each walk pattern that correspond to individual of p_2. On the left is the normal walk patterns portrait, while the strapped patterns are on the right side of Figure 6.3.

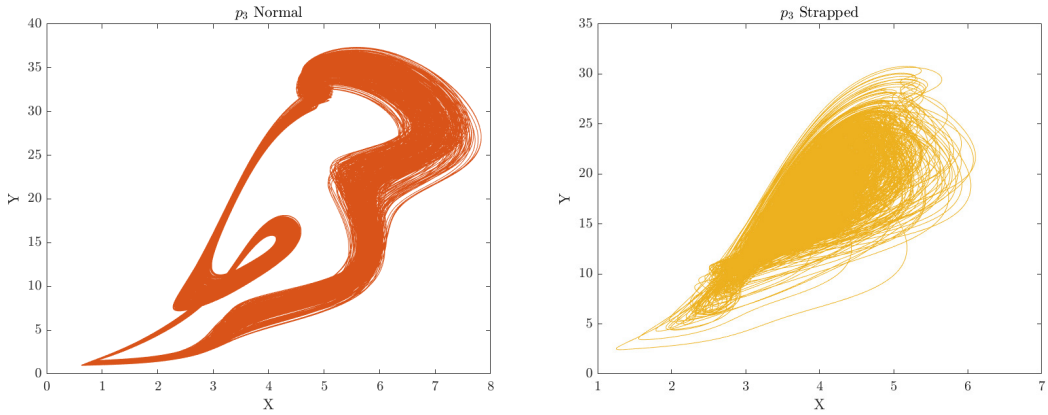


Figure 6.4: This figure shows the phase space plots for each walk pattern that correspond to individual of p_3. On the left is the normal walk patterns portrait, while the strapped patterns are on the right side of Figure 6.4.

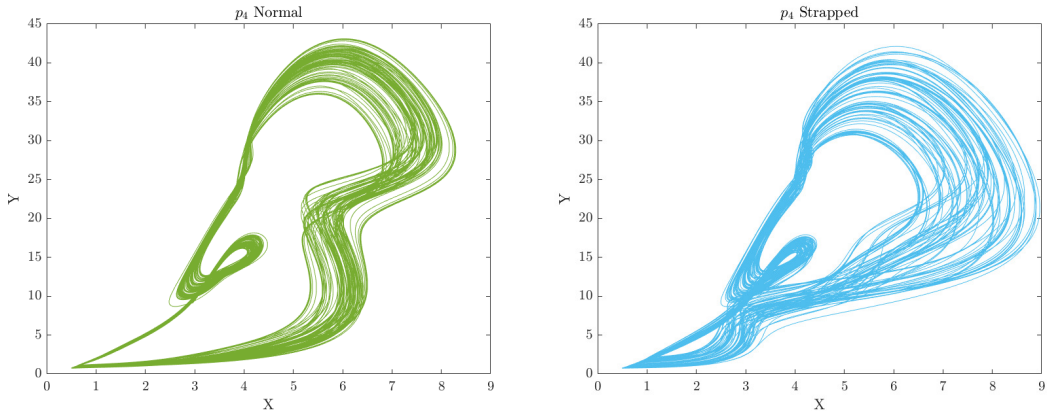


Figure 6.5: This figure shows the phase space plots for each walk pattern that correspond to individual of p_4. On the left is the normal walk patterns portrait, while the strapped patterns are on the right side of Figure 6.5.

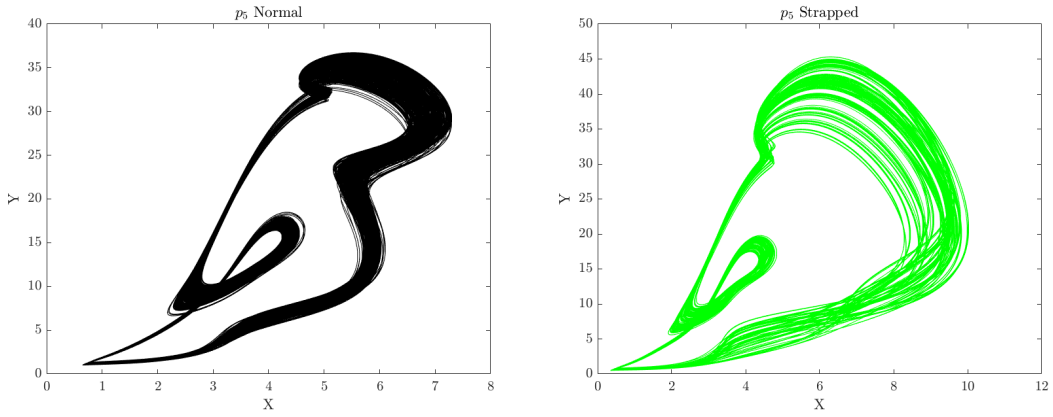


Figure 6.6: This figure shows the phase space plots for each walk pattern that correspond to individual of p_5. On the left is the normal walk patterns portrait, while the strapped patterns are on the right side of Figure 6.6.

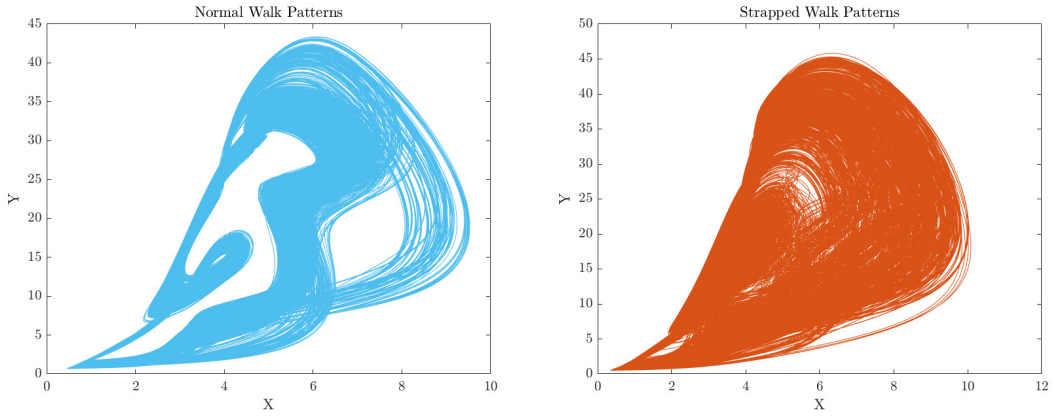


Figure 6.7: This figure shows the phase space plots for each walk pattern for all individuals of the same category. On the left is the normal walk patterns portrait, while the strapped patterns are on the right side of Figure 6.7.

The dynamic representations of CA analysis revealed that individuals with strapped walk experience more limitations in varying their gait, which is challenging to discern due to the high nonlinearity of the data. The CA phase plots presented in

Figures 6.2–6.6 as well as Figures 6.7 demonstrated that strapped walk individuals exhibit a gait pattern that is slower and more labored than that of individuals with normal healthy gait. Additionally, strapped walk individuals have a wider base of support and shorter steps, likely due to the increased stress on their joints and muscles from carrying extra weight. Strapped walk individuals are also at an increased risk for joint pain and arthritis, which can contribute to stiffness in their gait.

In contrast, individuals with normal healthy gait exhibit a smoother and more fluid gait pattern, with a narrower base of support and longer steps. The gait features depicted in the figures have an impact on the network of coupled CA oscillators, with the scaling of additive connectivity strength responding to external perturbations of gait. This is done to ensure that the input provided to each oscillator remains within the controlled domain and does not exceed a certain threshold.

6.4 Statistical Analysis of Criticality Analysis Data

6.4.1 Scatter Analysis

A scatter plot analysis for gait data can be a useful tool for discriminating between normal and strapped gait patterns. Scatter plot analysis involves plotting the CA gait data on a graph with the x-axis representing one variable and the y-axis representing another variable. By analysing the scatter plot, we can visually identify any patterns or trends in the data that may differentiate normal gait from strapped gait patterns. For example, we may observe that individuals with strapped gait patterns tend to have shorter stride lengths and wider step widths compared to those with normal gait patterns.

Figures 6.8–6.13 illustrate the distribution of CA data representation for each individual's walk pattern, corresponding to the figures depicted in Figures 6.2–6.7. Each plot shows the variability of continuous or successive gait interaction events

over time that an individual takes. Since there is a degree of association between the two dominant extracted features in the 2D space, the Pearson's correlation coefficient ρ, which measures the direction and strength of the dependency of extracted features on each other, is bounded by $0 \leq \rho \leq 1$. The average of ρ for normal walk patterns across all individuals is significantly higher and positive than that for strapped patterns. In other words, the effect of ρ for strapped features decreases steadily by an average of 0.28% compared to normal patterns.

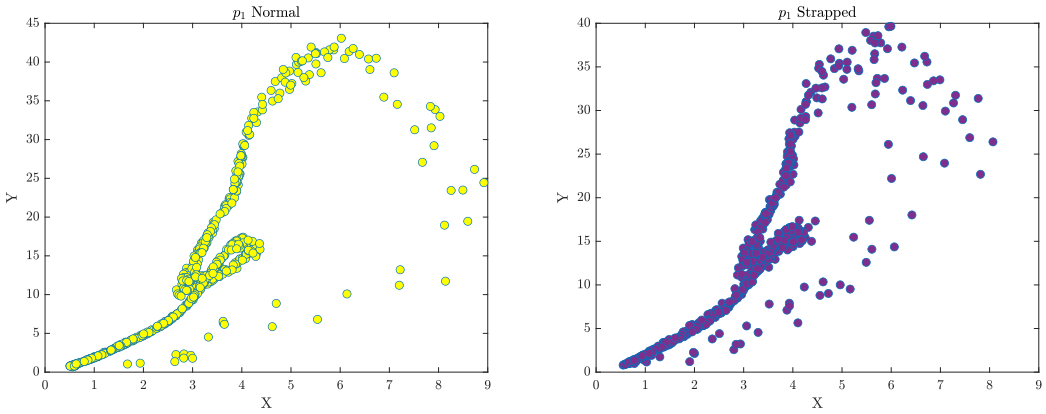


Figure 6.8: This figure shows the scatter plots for each walk pattern that correspond to the individual p_1. On the left is the normal walk patterns distribution, while the strapped patterns distribution is on the right side of Figure 6.8.

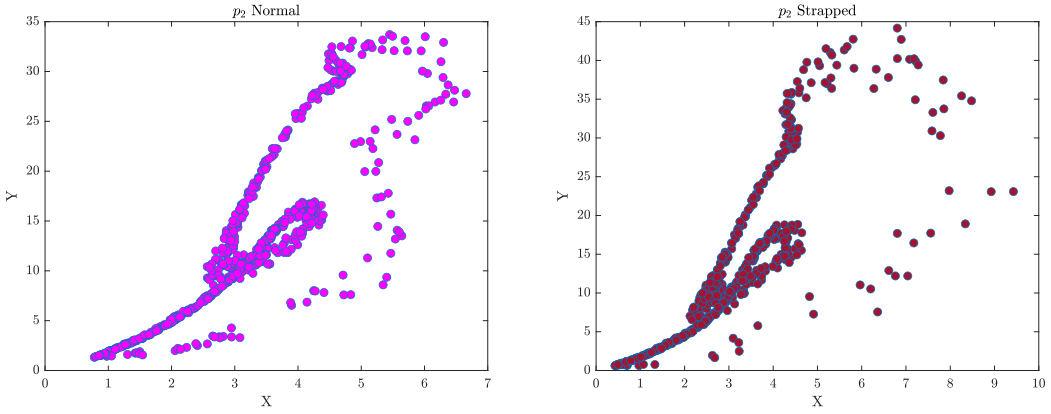


Figure 6.9: This figure shows the scatter plots for each walk pattern that correspond to the individual p_2. On the left is the normal walk patterns distribution, while the strapped patterns distribution is on the right side of Figure 6.9.

The figures depicted in Figures 6.8–6.13 exhibit the strapped gait on their right-hand side. These figures illustrate that there is a discernible concentration of data points over the range of data space or non-uniformity in the distribution of data samples, thereby distinguishing them from a typical walking pattern, as exemplified in Figure 6.10.

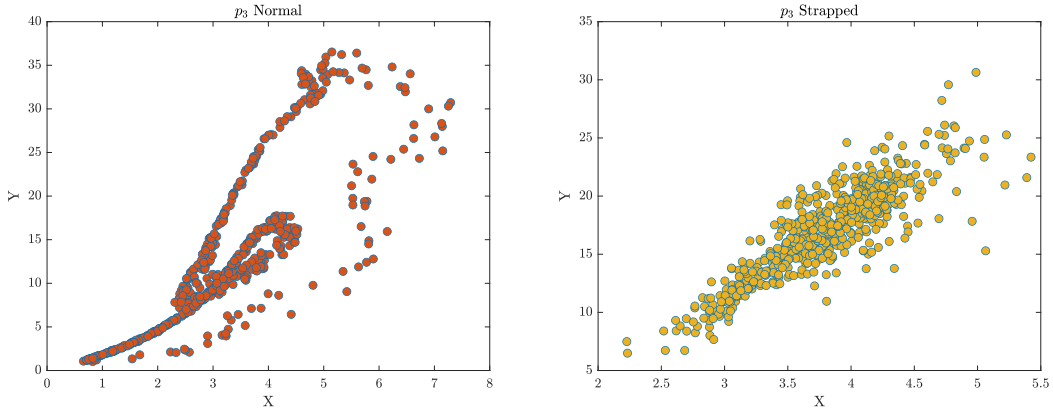


Figure 6.10: This figure shows the scatter plots for each walk pattern that correspond to the individual p_3. On the left is the normal walk patterns distribution, while the strapped patterns distribution is on the right side of Figure 6.10.

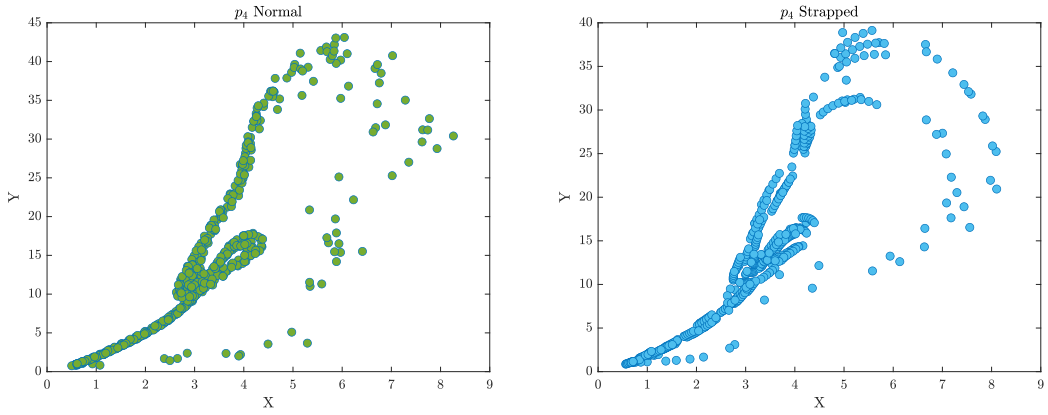


Figure 6.11: This figure shows the scatter plots for each walk pattern that correspond to the individual p_4. On the left is the normal walk patterns distribution, while the strapped patterns distribution is on the right side of Figure 6.11.

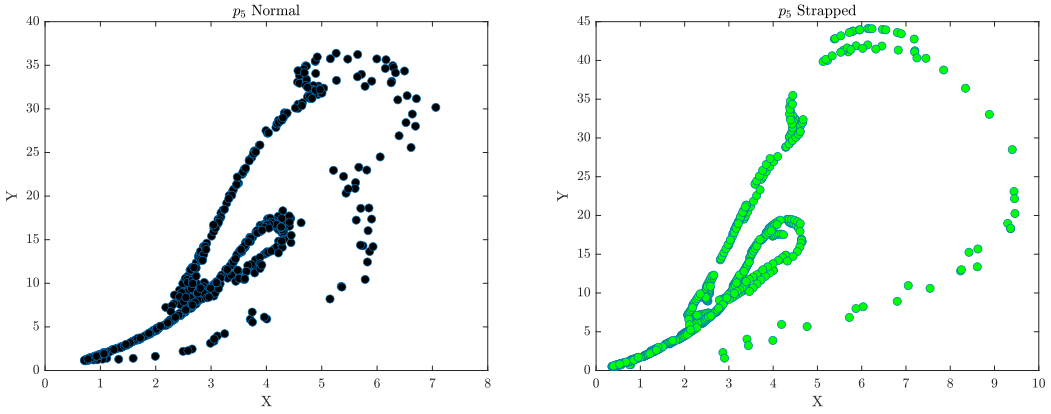


Figure 6.12: This figure shows the scatter plots for each walk pattern that correspond to the individual p_5. On the left is the normal walk patterns distribution, while the strapped patterns distribution is on the right side of Figure 6.12.

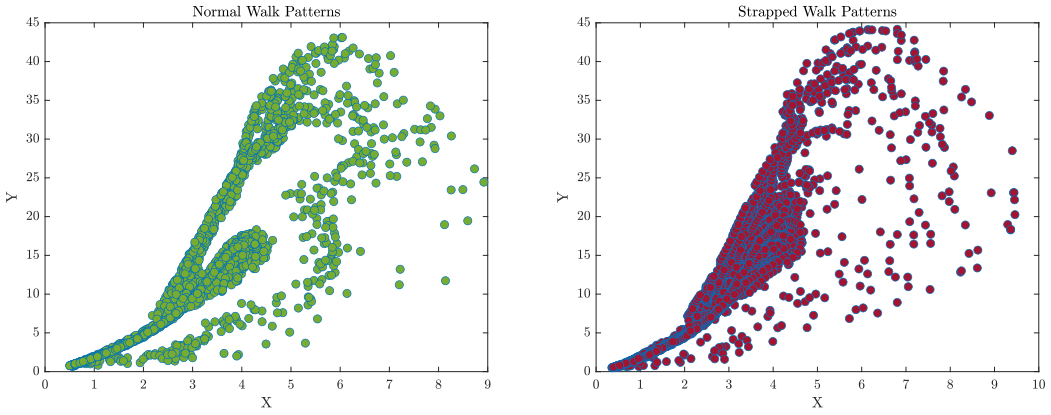


Figure 6.13: This figure shows the scatter plots for each walk pattern for all individuals of the same category. On the left is the normal walk patterns distribution, while the strapped patterns distribution is on the right side of Figure 6.13.

Table 6.1 shows the correlation coefficient values for five different variables, p_1 through p_5, between normal and strapped walking patterns. The correlation coefficients are approximately higher for the strapped walk pattern, indicating a stronger

relationship between these variables for this type of gait, while Table 6.2 summarises the correlation coefficient values for all individuals in both normal and strapped walking patterns. The correlation coefficients for both types of gaits are relatively high, with a slightly higher value for normal walking patterns. This suggests that there is a significant correlation between the variables across all individuals, regardless of the gait pattern.

Tables 6.1 and 6.2 provide a summary of the correlation coefficient values (ρ) for different walking patterns, including values for individual participants as well as for all individuals in the same category. Specifically, Table 6.1 reports the ρ values for different variables in normal and strapped walking patterns, while Table 6.2 summarises the ρ values for all individuals in each gait pattern category.

Table 6.1: The correlation coefficient ρ of p_1, p_2, p_3, p_4 and p_5.

Category	**Correlation Coefficient (ρ)**				
	$\boldsymbol{p_1}$	$\boldsymbol{p_2}$	$\boldsymbol{p_3}$	$\boldsymbol{p_4}$	$\boldsymbol{p_5}$
Normal Walk	0.815	0.846	0.832	0.84	0.847
Strapped Walk	0.86	0.819	0.871	0.838	0.778

Table 6.2: The correlation coefficient ρ of all individuals.

Category	**Correlation Coefficient (ρ)**
	All Individuals
Normal Walk Patterns	0.8337
Strapped Walk Patterns	0.8255

6.4.2 Spatiotemporal Analysis

Spatiotemporal analysis of gait data refers to the study of the space and time aspects of human walking patterns. It involves the measurement and analysis of various parameters related to gait, such as step length, stride length, cadence, velocity, and variability. When comparing normal and strapped walk patterns, spatiotemporal analysis can provide valuable insights into the differences between the two. Strapped walk patterns refer to walking while wearing some form of restriction or constraint, such as ankle weights or braces, which can affect gait parameters. Compared to normal walking, strapped walking patterns may exhibit changes in spatiotemporal parameters such as shorter stride length, reduced velocity, increased variability, and altered gait symmetry. The specific changes observed will depend on the type and severity of the restriction or constraint.

Figures 6.14 to 6.16 present a graphical comparison between normal gait dynamic patterns and strapped patterns, revealing significant effects of gait dynamic disturbances on gait progression. This analysis particularly highlights the variations and locations of gait patterns in individuals with certain gait abnormalities. The median is represented by the middle line of each categorical box, with the degree of dispersion and skewness of the CA data measured by the difference in variability between each category for each individual. Table 6.3 presents the median values of p_1 to p_5 for normal and strapped walking patterns. The data shows that the median values of p_1 to p_5 for the strapped walking pattern are lower compared to the normal walking pattern. The median value of p_1 for the strapped walking pattern is the lowest among all the categories, indicating that individuals with certain gait abnormalities experience significant disruptions in their gait dynamics during strapped walking. Overall, these findings suggest that spatiotemporal analysis of gait data can reveal important insights into the differences between normal and strapped walking patterns and can aid in clinical or rehabilitation interventions to improve gait function.

Table 6.3: The median of p_1, p_2, p_3, p_4 and p_5.

Category	**Median**				
	$\boldsymbol{p_1}$	$\boldsymbol{p_2}$	$\boldsymbol{p_3}$	$\boldsymbol{p_4}$	$\boldsymbol{p_5}$
Normal Walk	21.804	22.252	22.709	21.869	23.029
Strapped Walk	19.99	24.645	21.527	22.085	20.637

Table 6.4 provides the median values for two categories of walking, normal walk and strapped walk, for all individuals. The median value for normal walk is 22.253 while the median value for strapped walk is 20.130. This indicates that individuals in the normal walk category tend to have a higher median value compared to individuals in the strapped walk category.

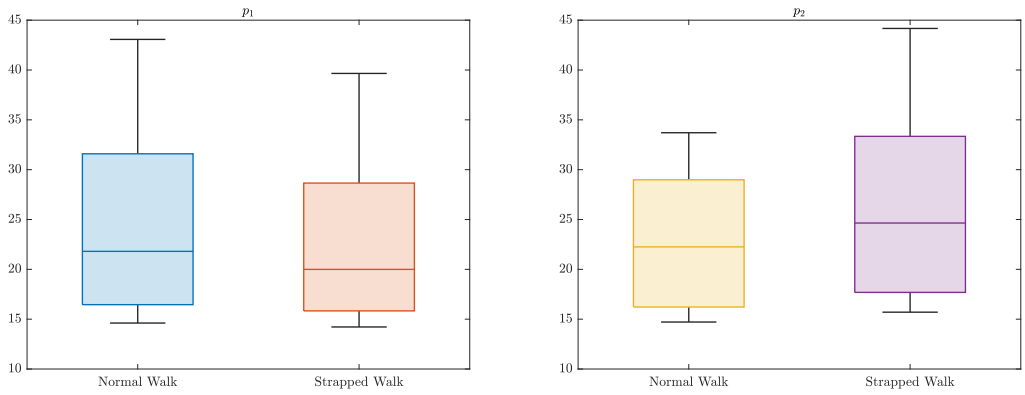


Figure 6.14: This figure shows the spatiotemporal plots for each walk pattern that correspond to each individual p_1 and p_2.

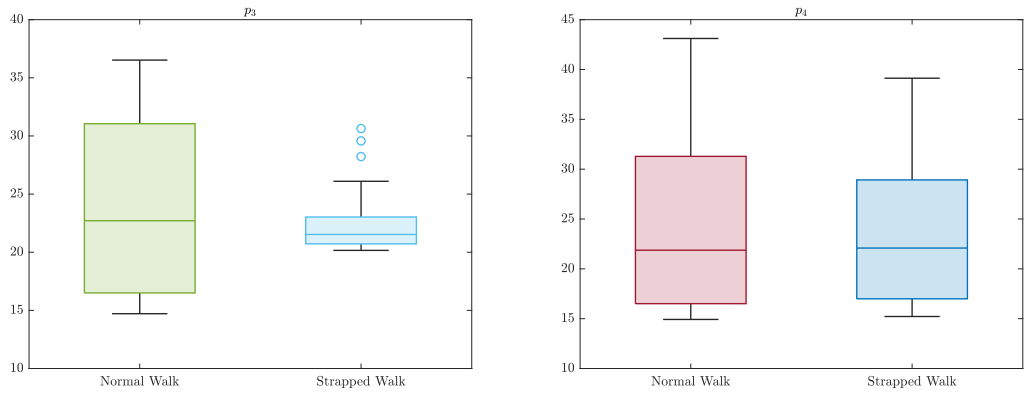


Figure 6.15: This figure shows the spatiotemporal plots for each walk pattern that correspond to each individual p_3 and p_4.

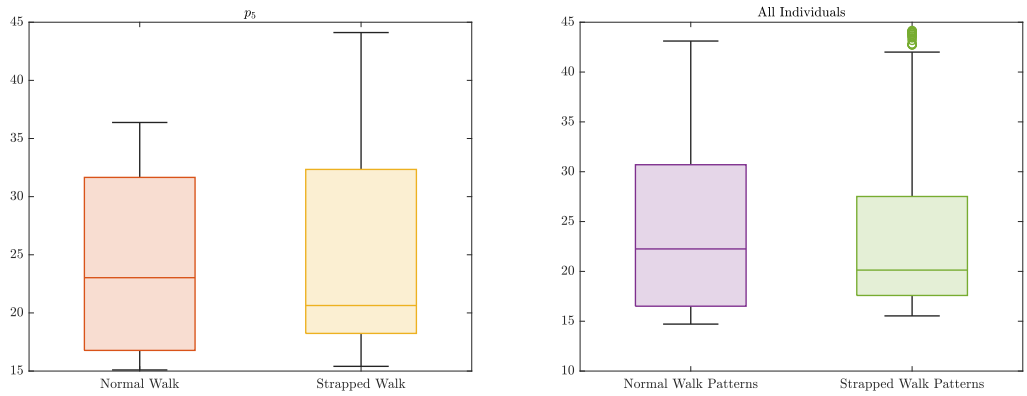


Figure 6.16: This figure shows the spatiotemporal plots for each walk pattern that correspond to each individual p_5 and to all individuals from the same category.

Table 6.4: The median of all individuals.

Category	**Median**
	All Individuals
Normal Walk	22.253
Strapped Walk	20.130

6.4.3 Histogram Analysis

Histogram analysis of gait data can provide insight into the frequency and distribution of certain gait parameters, such as step length, stride length, and cadence. By comparing the histograms of gait data between normal and strapped walk patterns, one can identify differences in the gait patterns. For example, if the histogram for a normal walk pattern shows a single peak centered around a certain value, and the histogram for a strapped walk pattern shows a broader distribution with multiple peaks, this could suggest that the strapped walk pattern is less consistent and less efficient than the normal walk pattern. Similarly, if the histogram for a normal walk pattern shows a roughly symmetric distribution, and the histogram for a strapped walk pattern shows a skewed distribution with a longer tail, this could suggest that the strapped walk pattern is characterised by longer and possibly more variable strides, which may be indicative of compensatory mechanisms due to the use of straps.

Visual representations in the form of histograms are depicted in Figures 6.17 through 6.21, demonstrating the unique statistical properties of the distribution of CA data by illustrating the frequencies of individual features. Table 6.5 displays a significantly higher average mean for the strapped walk distribution compared to the normal walk distribution. Furthermore, it is evident that the strapped walk data distribution for CA individuals is skewed to the right, which is expected for

individuals p_1, p_2, p_4 and p_5, while the distribution of p_3 is skewed to the left. As the degree of skewness increases, so does the level of variability. This one-sided distribution is caused by the lower bound of data features and results in outliers due to an unintended change in experiment settings that affects the distribution's direction. Additionally, Table 6.6 presents the mean values for all individuals in two categories of walking: normal walk and strapped walk. The mean value for normal walk is 23.838, while the mean value for strapped walk is 23.198. This suggests that the overall mean of normal walk is higher than that of strapped walk.

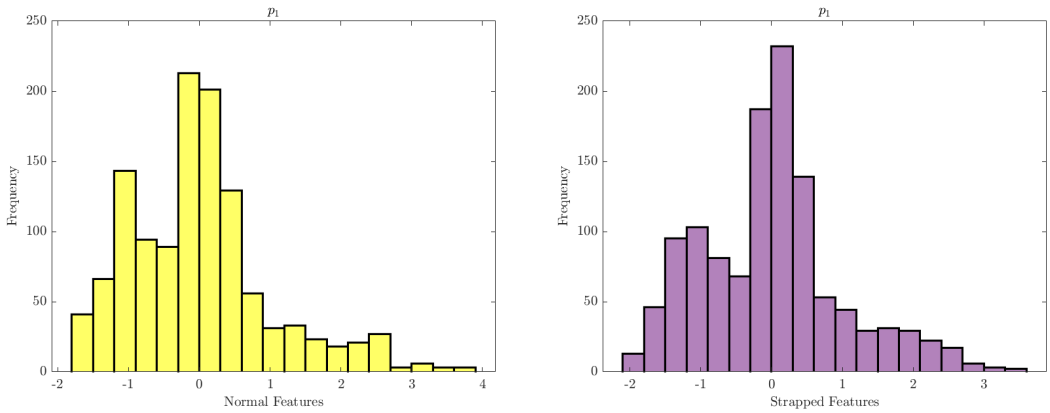


Figure 6.17: This figure shows the histogram plots for each walk pattern that correspond to the individual p_1.

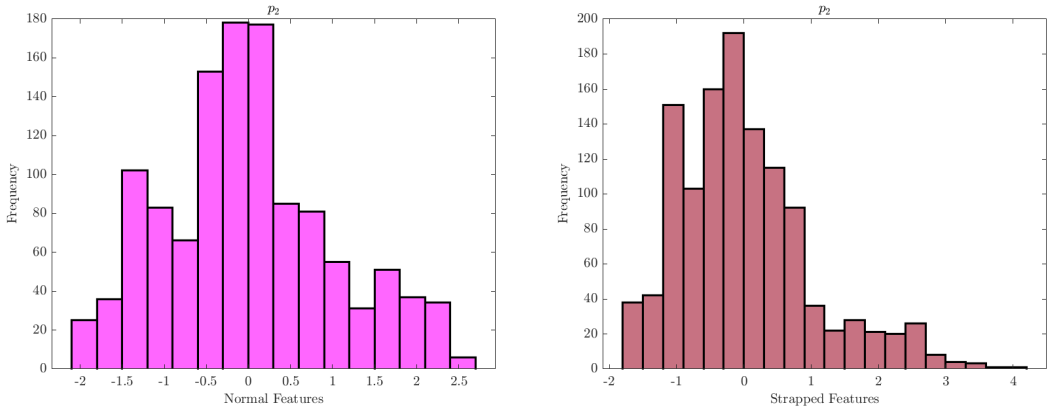


Figure 6.18: This figure shows the histogram plots for each walk pattern that correspond to the individual p_2.

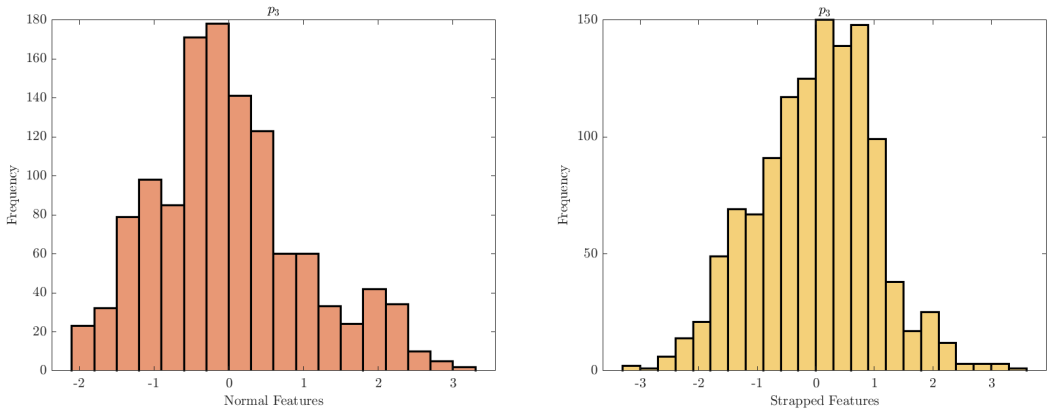


Figure 6.19: This figure shows the histogram plots for each walk pattern that correspond to the individual p_3.

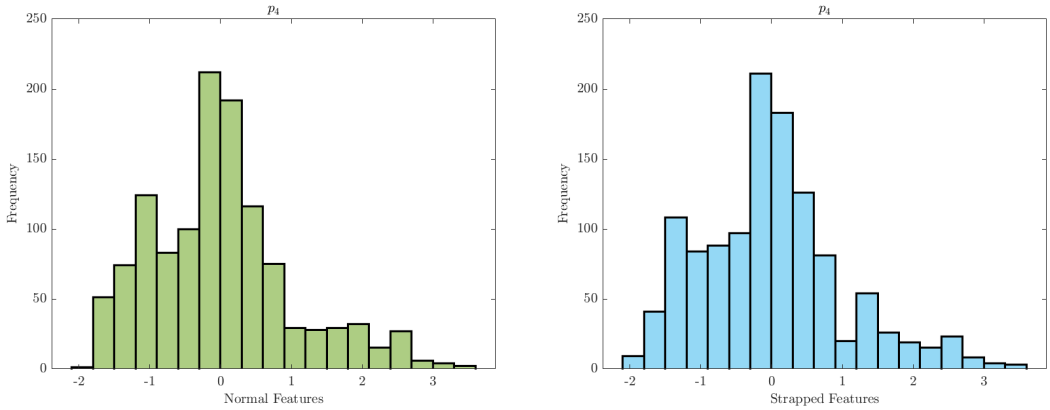


Figure 6.20: This figure shows the histogram plots for each walk pattern that correspond to the individual p_4.

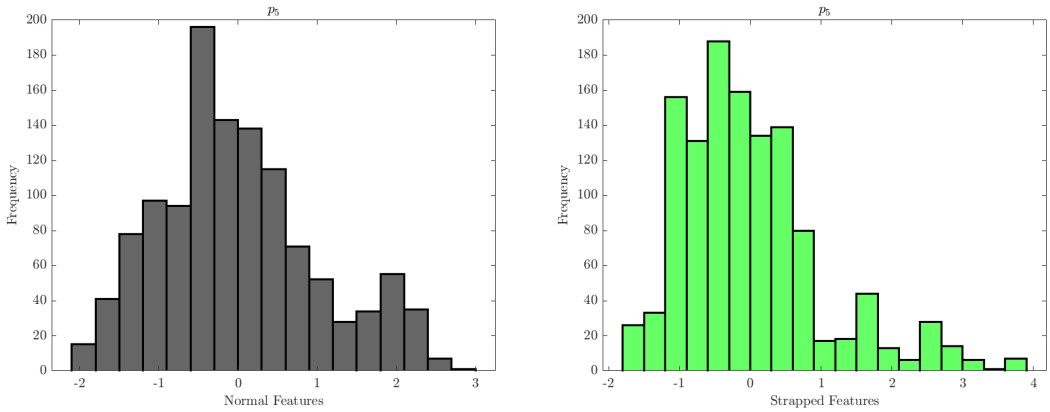


Figure 6.21: This figure shows the histogram plots for each walk pattern that correspond to the individual p_5.

Table 6.5: The mean of p_1, p_2, p_3, p_4 and p_5.

Category	**Mean**				
	$\boldsymbol{p_1}$	$\boldsymbol{p_2}$	$\boldsymbol{p_3}$	$\boldsymbol{p_4}$	$\boldsymbol{p_5}$
Normal Walk	24.483	22.803	23.522	24.257	24.287
Strapped Walk	22.497	25.99	22.124	23.914	25.835

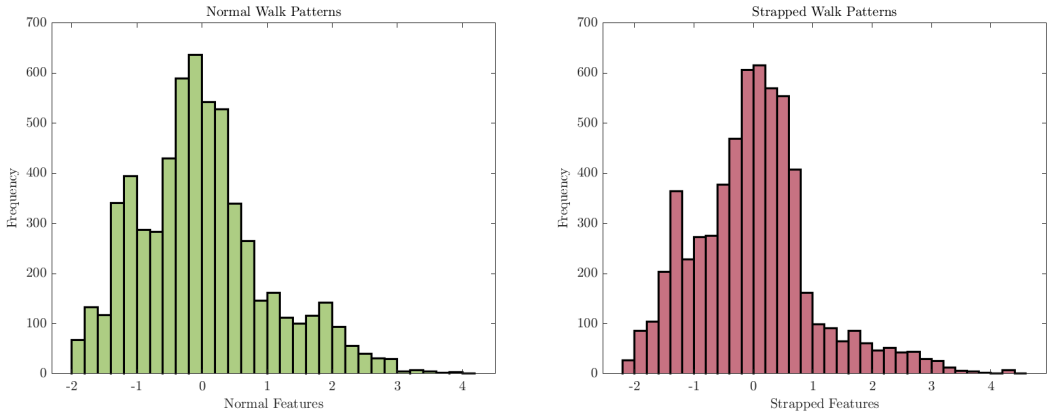


Figure 6.22: This figure shows the histogram plots for each walk pattern of individuals from the same gait category.

Table 6.6: The mean of all individuals.

Category	**Mean**
	All Individuals
Normal Walk	23.838
Strapped Walk	23.198

6.5 The Proposed SVM Classifier

The proposed Support Vector Machine (SVM) model is based on the kernel property discussed in Section 4.3.2. The most dominant gait features extracted by the RCC Berry model [29], represented by f (equation (3.11)) on the x-axis and m (equation (3.10)) on the y-axis, exhibit high correlation and nonlinearity. The Kernel SVM is well-suited to handle and transform such dynamic and nonlinear gait data into a new feature space to facilitate the classification process for various data subjects. To train and test the classification performance of various models, the extracted features from the controlled CA model were used as input to the SVM algorithm. The SVM classifier was trained and tested on the nonlinear CA extracted features for each individual of various gait subjects. Specifically, the two-dimensional normal and strapped walk subjects were used for training and testing, respectively. This scenario was repeated to train the Kernel SVM on all individual data, where the normal patterns were used for training and the strapped ones were used for testing to reduce the number of parameters that could lead to unnecessary overfitting in training the gait data. The purpose of conducting these two experiments is to evaluate the performance of the SVM and its ability to classify various dynamic gait subjects, as well as detect disturbances associated with individual walk behavior. The implementation procedures of the proposed SVM algorithm are described in

Algorithm 1.

6.6 The Proposed SVM Training

A total of 3000 data points are available for each gait subject, namely normal or strapped, with an equal distribution among five participants. This results in a total of 6000 data samples for both subjects, corresponding to the extracted features input to the classification algorithm. Each participant in a given gait category has 600 data samples. The training and testing of the Kernel SVM algorithm were carried out using a 50% splitting ratio. Prior to training the SVM, all CA gait features were normalised using the z-score to center the data samples at zero mean and unity standard deviation. The implementation of the Kernel SVM algorithm was conducted using MATLAB for statistical analysis of gait data, including the development of various Kernel SVM models to examine the impact of Kernel characteristics (regularisation parameter C and control width σ) on the classification performance of gait data. The generalisation performance of the trained Kernel SVM was evaluated by measuring the prediction accuracy for each model, analysing the confusion matrix (refer to Appendix B), plotting the Receiver Operating Characteristics (ROC) curve, and calculating the area under the ROC curve. Further details on this evaluation process are discussed in subsequent sections.

6.7 Simulation Results

In order to ascertain the efficacy of the methodology proposed earlier, the subsequent performance metrics are employed:

6.7.1 Confusion Matrix

In a binary classification problem, the dataset labels are classified as either positive or negative. The output of the SVM classifier in such a scenario can be represented by a structured contingency table, commonly known as a confusion matrix. This matrix contains information about the actual and predicted binary classification and is used to evaluate the performance of the classification algorithm in differentiating between different categories, such as normal or strapped. The confusion matrix has four key components, as outlined by the literature [52]:

- **True Positive** (TP): the term pertains to the accurate identification and classification of positive data samples, which in this context refers to participants engaging in normal walking behavior.
- **False Positive** (FP): the term "false positives" pertains to the count of negative data instances, specifically those from participants who engaged in a strapped walk, that have been inaccurately identified as positive or normal.
- **True Negative** (TN): this pertains to the count of negative instances within the data set (namely individuals participating in a strapped walk) that have been accurately identified as negative (i.e., subjected to strapping).
- **False Negative** (FN): the term "false negatives" pertains to the count of positive data points, which correspond to individuals exhibiting normal walking behavior, that are erroneously identified as negative, indicating that they are wearing a brace or other similar assistive device.

The evaluation of a classifier's performance includes several metrics, including accuracy, precision, F_1-score, recall or sensitivity (also known as True Positive Rate or TPR), specificity (also known as True Negative Rate or TNR), and false positive rate (FPR). These metrics are calculated using the True Positive (TP), False Positive

(FP), True Negative (TN), and False Negative (FN) values. The definitions of these metrics are as follows:

- **Accuracy**: denotes the comprehensive rate of successful predictions, which is calculated as the proportion of accurately classified labels (represented by $TP+TN$) to the total number of data samples ($TP+TN+FP+TN$). This metric is mathematically defined as:

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

- **Precision**: used for detecting normal walking status. It is computed as the ratio of accurately predicted positive labels for normal walking to the total number of predicted labels indicating a normal walking condition. It is calculated using the following formula:

$$Precision = \frac{TP}{TP+FP}.$$

- **Recall (TPR)**: pertains to the accuracy of correctly predicting positive labels with normal walk status, in relation to all normal walk data samples. Its calculation is determined by the following formula:

$$Recall = \frac{TP}{TP+FN}.$$

- **F_1-Score**: is a metric that maintains a balance between precision and recall. Its calculation is as follows:

$$F_1 - Score = \frac{2 \times Precision}{Precision + Recall}.$$

- **True Negative Rate (TNR)**: refers to the ratio of accurately predicted negative labels under strapped walk conditions, as a proportion of all the samples in the strapped walk dataset. Mathematically, it is expressed as:

$$TNR = \frac{TN}{TN+FP}.$$

- **False Positive Rate (FPR)**: quantifies the proportion of strapped walk samples that are inaccurately labeled as normal walk data, in relation to the total number of strapped data samples. Mathematically, FPR is represented as:

$$FPR = \frac{FP}{FP + TN}.$$

The purpose of the aforementioned metrics is to assess the effectiveness of the proposed SVM algorithm in identifying the atypical gait patterns associated with diverse individuals. The performance metrics of the SVM classifier for each individual at specific levels of regularisation parameters $C = (0.1, 1, 10)$ are summarised in Appendix B. Among these metrics, the accuracy is the primary validation measure that provides a comprehensive evaluation of the SVM classifier's overall performance. Notably, the accuracy of the classifier is contingent on the values of σ and C. In general, the optimal outcomes were obtained when σ was set to 0.1 or 1. In light of one of the primary discoveries concerning the confusion matrix, it is essential to reflect upon its implications. For instance, Table B.9 presents the results of SVM classification for p_5 walk patterns at $\sigma = 10$. The table demonstrates the classification performance of SVM with varying values of C (i.e., 0.1, 1, and 10). The performance measures such as TP (true positives), FP (false positives), FN (false negatives), TN (true negatives), FPR (false positive rate), precision, recall, F_1-score, specificity, AROC (area under the receiver operating characteristic curve), and accuracy are provided. The table indicates that for all values of C, the SVM correctly classified 298 instances as true positives and misclassified 302 instances as false positives. No instances were misclassified as false negatives or true negatives. The F_1-score for all values of C was 0.663, indicating that the precision and recall were balanced.

Table B.10 shows the SVM classification results for p_5 walk patterns at $\sigma = 0.1$ and $\sigma = 1$. The table provides performance measures for SVM with different values of C (i.e., 0.1, 1, and 10) at both values of sigma. For $\sigma = 0.1$, the SVM correctly

classified 268, 264, and 266 instances as true positives for C values of 0.1, 1, and 10, respectively, while misclassifying 9, 1, and 0 instances as false positives. Additionally, 30, 34, and 32 instances were misclassified as false negatives for C values of 0.1, 1, and 10, respectively. The F_1-score for C values of 0.1, 1, and 10 was 0.932, 0.937, and 0.943, respectively. The accuracy was 93.5%, 94.17%, and 94.67% for C values of 0.1, 1, and 10, respectively.

For $\sigma = 1$, the SVM correctly classified 298 instances as true positives for all C values, while misclassifying 134, 66, and 45 instances as false positives for C values of 0.1, 1, and 10, respectively. No instances were misclassified as false negatives or true negatives. The F_1-score for C values of 0.1, 1, and 10 was 0.816, 0.9, and 0.929, respectively. The accuracy was 77.67%, 89%, and 92.33% for C values of 0.1, 1, and 10, respectively.

6.7.2 Receiver Operating Characteristics (ROC) Curve

The Receiver Operating Characteristic (ROC) plot is an additional tool for assessing the performance of the SVM classifier. This plot displays a representation of True Positive Rate (TPR) or sensitivity versus the complement of True Negative Rate ($\approx 1 - TNR$) across a range of threshold levels, ranging from 0 to 1. In Figures 6.23–6.28, ROC curves were generated for the optimal σ and C values that produced the highest accuracy during the overall trial period for each individual, as well as for all individuals as depicted in Figure 6.28. The results show that the Kernel SVM performed well in detecting normal and strapped features for each individual, as well as for the different gait subjects of all individuals. The area under the ROC curve (AROC) also increased with increasing values of C, indicating a better ability to distinguish between the positive and negative classes. However, at $\sigma = 1$, the accuracy remained at its maximum value, but the AROC value decreased as the value of C increased. This suggests that the model is overfitting to the training data

and not performing well on the test data. The false positive rate (FPR) increased as the value of C increased, indicating that the model is classifying more negative instances as positive. Overall, the AROC and ROC curve suggest that the model's performance is highly dependent on the values of σ and C. The ROC curve in the context of the SVM classifier is influenced by both the choice of the regularisation parameter C and the kernel parameter σ. The regularisation parameter C controls the balance between the training error and the complexity of the decision boundary, while the kernel parameter σ affects the shape and flexibility of the decision boundary. The choice of C and σ can have a significant impact on the performance of the SVM classifier, as well as the shape and location of the ROC curve. Therefore, it is important to carefully tune both parameters to achieve the best possible performance. In the given context, the ROC curve was plotted for the best pair of σ and C values that satisfy the highest accuracy during the overall trial period for each individual and for all individuals. It was shown that the kernel SVM performed well in most cases for detecting normal and strapped features per individual, as well as between the different gait subjects of all individuals.

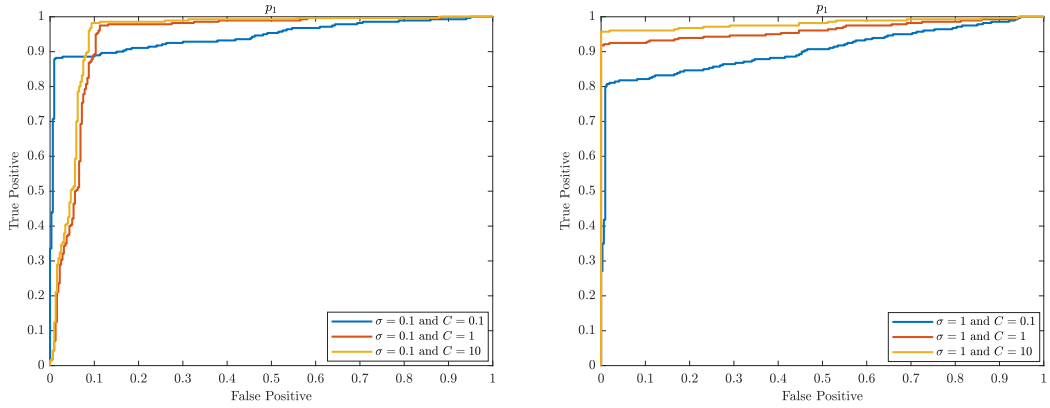


Figure 6.23: ROC (Receiver Operating Characteristic) curves of p_1 show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

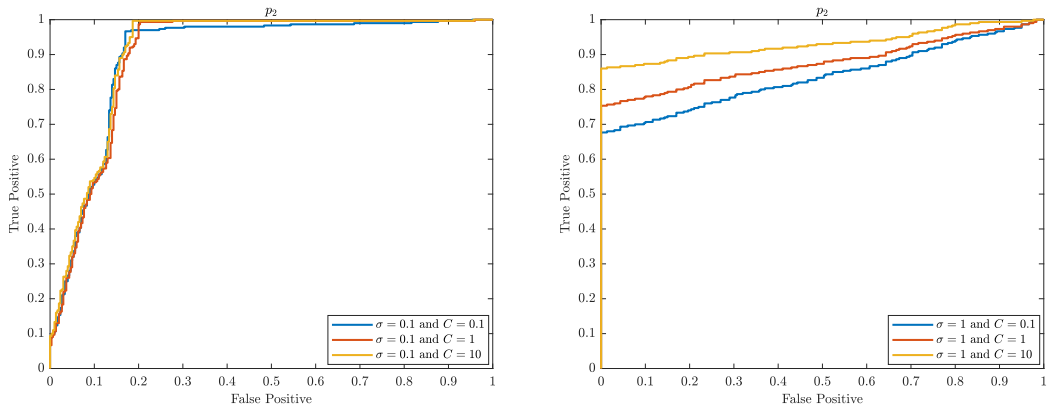


Figure 6.24: ROC (Receiver Operating Characteristic) curves of p_2 show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

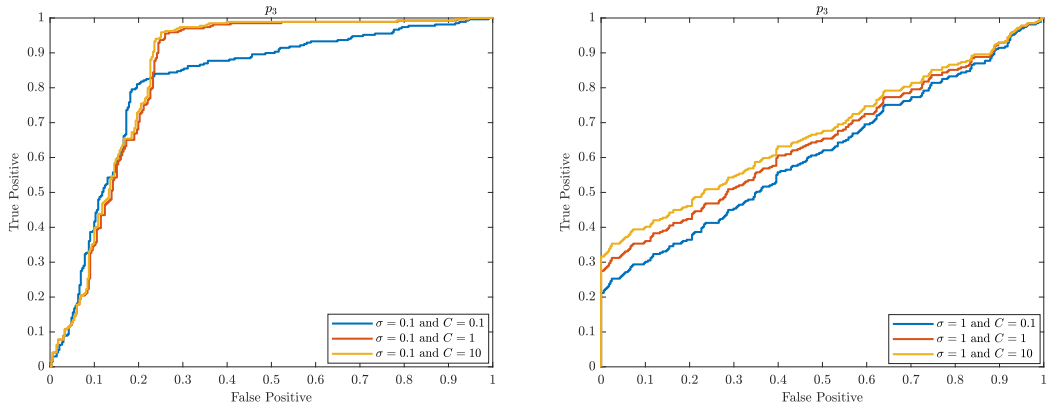


Figure 6.25: ROC (Receiver Operating Characteristic) curves of p_3 show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

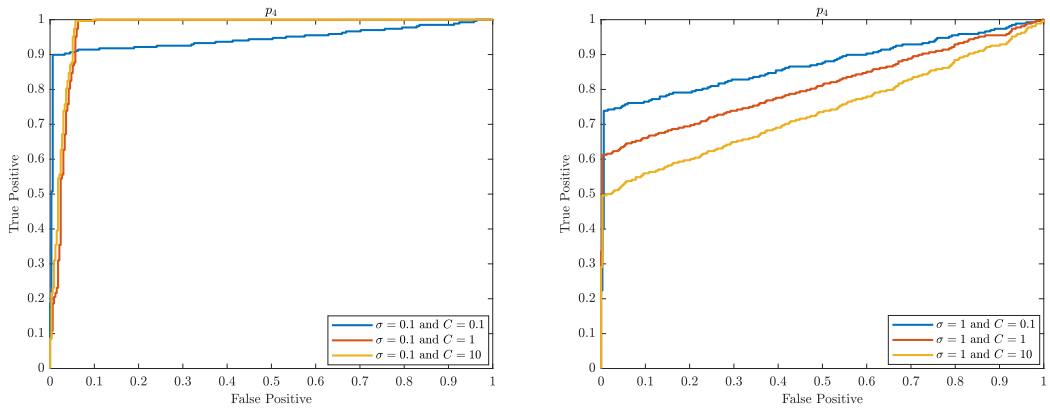


Figure 6.26: ROC (Receiver Operating Characteristic) curves of p_4 show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

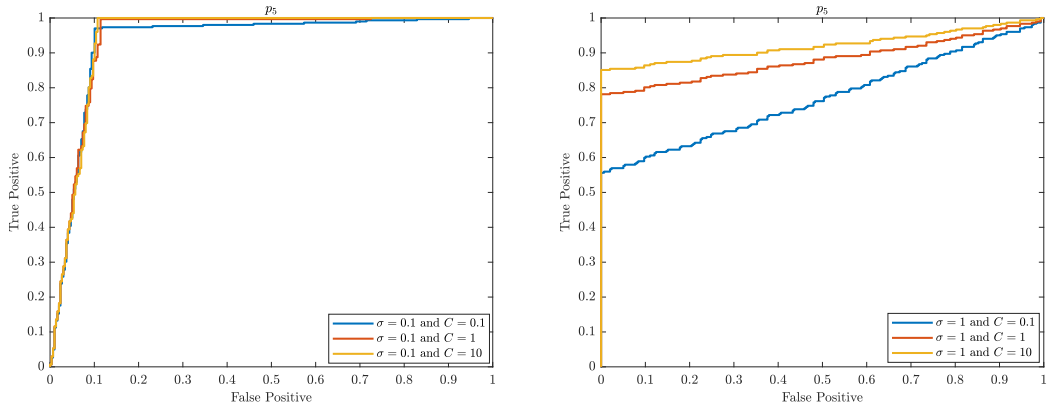


Figure 6.27: ROC (Receiver Operating Characteristic) curves of p_5 show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

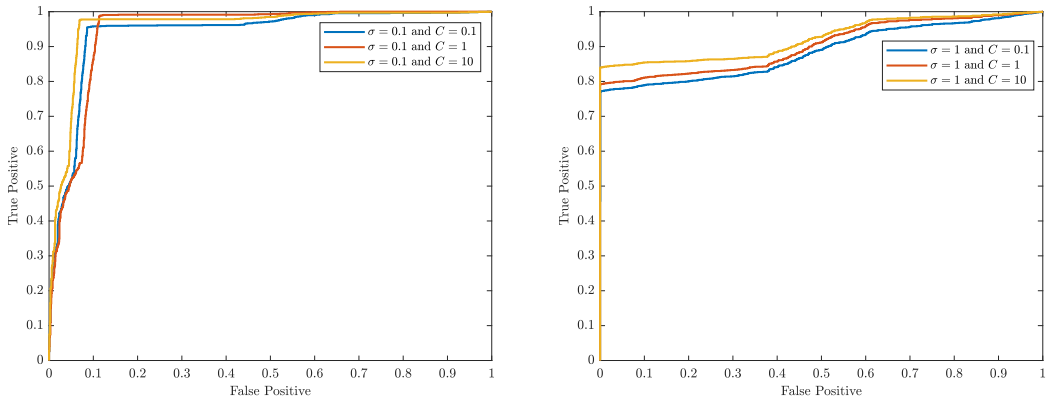


Figure 6.28: ROC (Receiver Operating Characteristic) curves of all individuals' gait subjects show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

6.7.3 Area Under the ROC Curve (AROC)

The Area Under the Curve (AUC), which takes into account all possible classification thresholds, provides an overall assessment of binary classifier performance. A higher AUC value signifies superior performance, with the ROC curve and AUC serving as useful tools for evaluating binary classifiers, particularly in cases of imbalanced classes. AROC values range from 0 to 1, with 0.5 representing random guessing and 1 signifying flawless classification. The ROC area for each individual walk's Gaussian kernel as a function of the regularisation parameter C is depicted in Figures 6.29–6.31. The SVM classifier's performance can be influenced by the C parameter, and there is a particular value or range that can enhance efficiency relative to others. Tables B.2–B.12 in Appendix B demonstrate that the SVM classifier achieves optimal results when $\sigma = 0.1$ and, in some cases, when $\sigma = 1$, compared to other σ values. Based on this, Figures 6.29–6.31 were generated, illustrating that for certain individuals, a range of C values can produce optimal performance. Addi-

tionally, it is demonstrated in Figures 6.29– 6.31 that for some individuals, a higher sensitivity level results in a larger Area under the Curve and better performance.

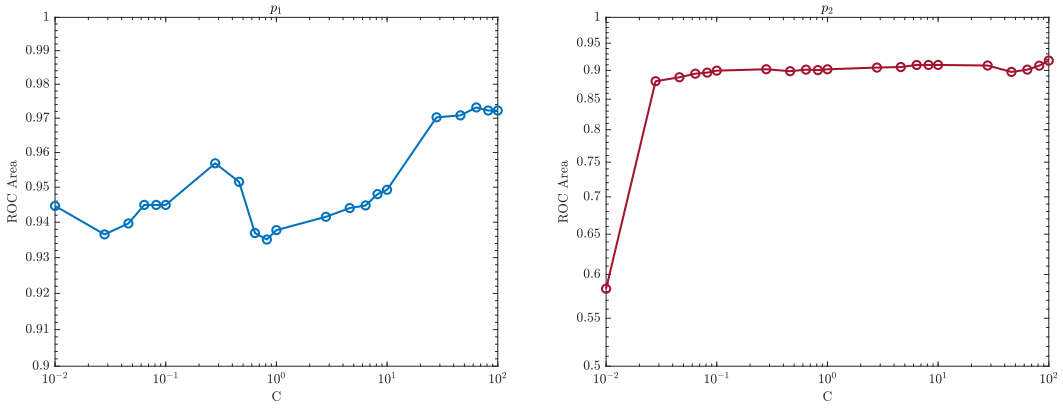


Figure 6.29: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when $\sigma = 0.1$ for each individual p_1 and p_2, respectively.

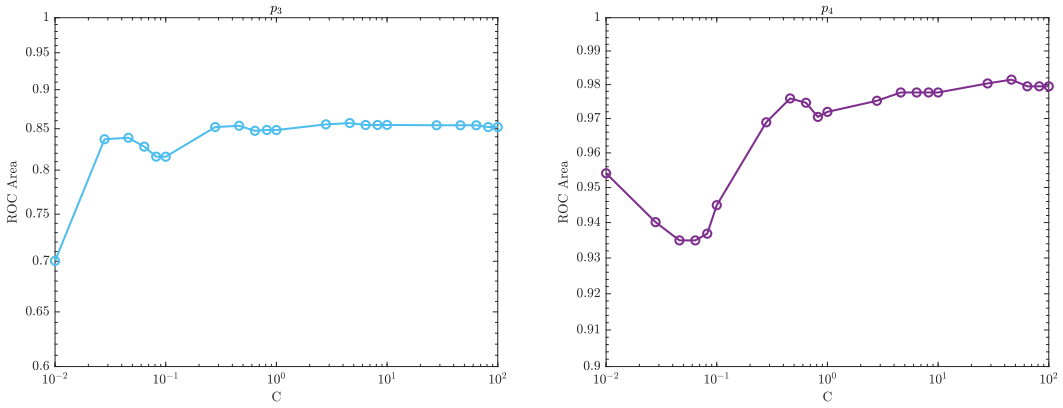


Figure 6.30: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when $\sigma = 0.1$ for each individual p_3 and p_4, respectively.

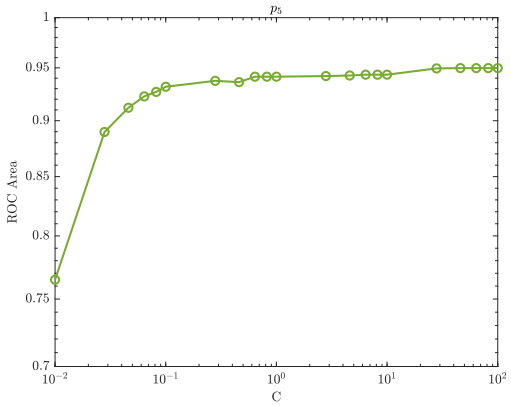


Figure 6.31: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when $\sigma = 0.1$ for the individual p_5.

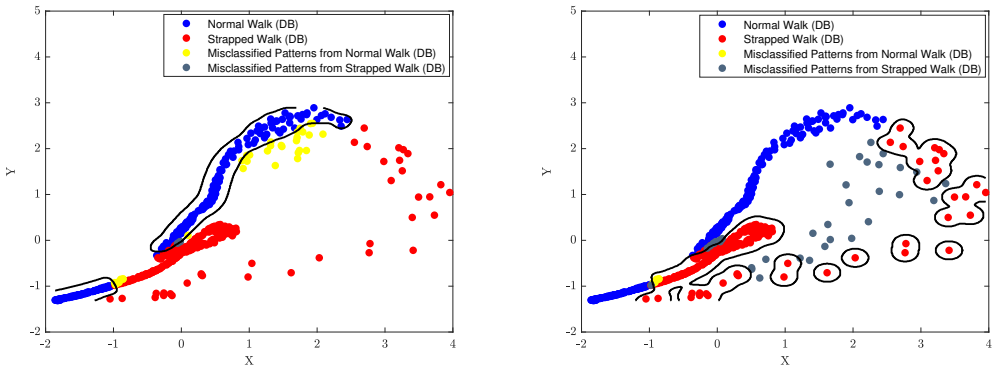
Tables B.2–B.12 in Appendix B demonstrate that the SVM model measurements indicate a high level of efficiency in distinguishing between normal and strapped gait disturbances, as the AUC and accuracy values are closely aligned. The AUC was computed through the custom algorithm software routines developed in MATLAB.

6.7.4 Classification Decision Boundary

The decision boundary of a SVM is a hyperplane that separates the different classes of data. The position and orientation of the decision boundary are determined by the SVM algorithm during the training process, based on the training data. The SVM algorithm has two key parameters, C and σ, that can influence the position and orientation of the decision boundary. The C parameter is a regularisation parameter that controls the trade-off between maximising the margin (i.e., the distance between the decision boundary and the nearest data points from each class) and minimising the classification error. A higher value of C will result in a narrower margin and a more complex decision boundary, which can potentially lead to overfitting to the training data. Conversely, a lower value of C will result in a wider margin and

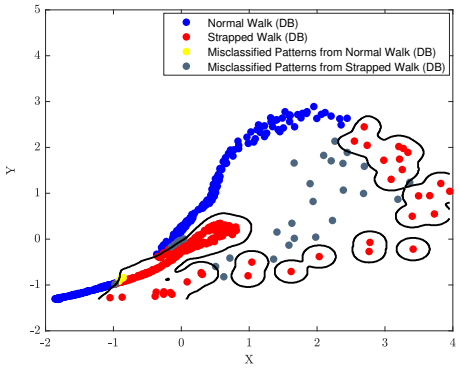
a simpler decision boundary, which can potentially lead to underfitting. The σ parameter is the width of the Gaussian kernel function used by the SVM algorithm to transform the input data into a higher-dimensional space. A smaller value of σ will result in a more peaked kernel function, which can lead to overfitting, while a larger value of σ will result in a smoother kernel function, which can lead to underfitting. Therefore, the values of C and σ can affect the position, orientation, and complexity of the decision boundary, and can impact the performance of the SVM classifier on unseen data. It is important to choose appropriate values of C and σ through cross-validation or other techniques to achieve the best classification accuracy.

The decision boundary visualisations in Figures 6.32– 6.42, which are in support of the results presented in the left graph of Figure 6.23– 6.28, demonstrate the effectiveness of the Kernel SVM model in identifying gait abnormalities or dynamics that deviate from normal patterns. These figures illustrate the decision boundary for the SVM trained on the walk patterns of all individuals as opposed to just the strapped individual. The robustness of the model in this regard is noteworthy. Notably, the SVM model achieves optimal performance at $\sigma = 0.1$ and 1 for a range of C values.



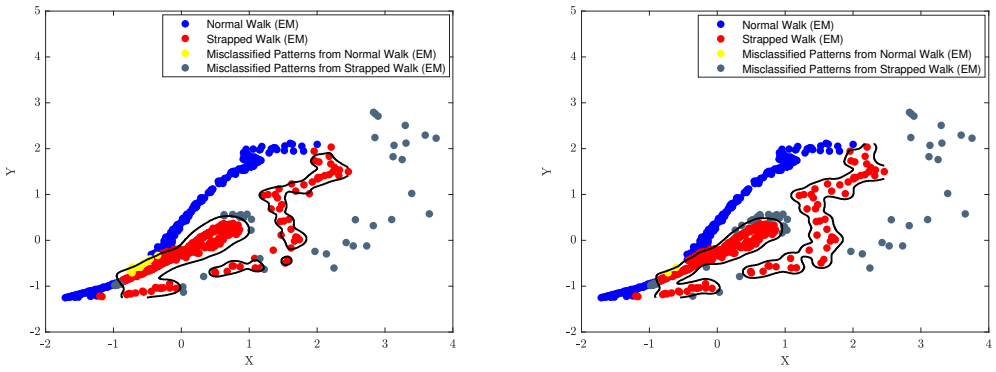
(a) $\sigma = 0.1$ and $C = 0.1$ (b) $\sigma = 0.1$ and $C = 1$

Figure 6.32: The decision boundary when the SVM model trained on p_1 normal walk patterns versus that of the strapped one, when: (**a**) $\sigma = 0.1$ and $C = 0.1$ and (**b**) $\sigma = 0.1$ and $C = 1$.



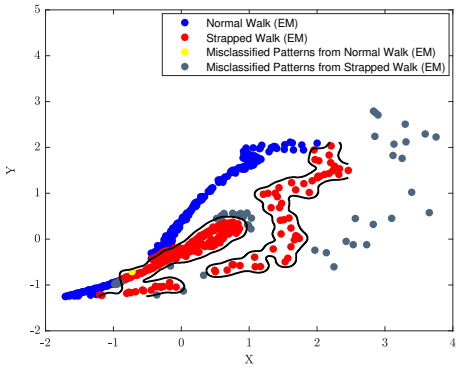
(a) $\sigma = 0.1$ and $C = 10$

Figure 6.33: The decision boundary when the SVM model trained on p_1 normal walk patterns versus that of the strapped one, when: $\sigma = 0.1$ and $C = 10$.



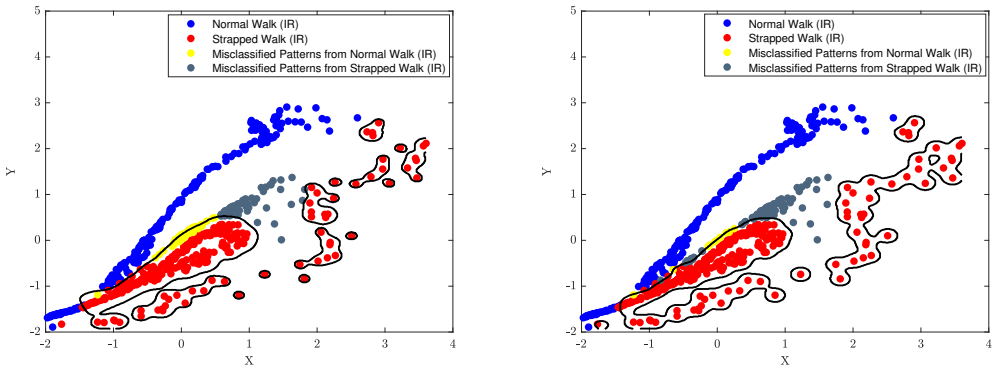
(a) $\sigma = 0.1$ and $C = 0.1$ (b) $\sigma = 0.1$ and $C = 1$

Figure 6.34: The decision boundary when the SVM model trained on p_2 normal walk patterns versus that of the strapped one, when: (**a**) $\sigma = 0.1$ and $C = 0.1$ and (**b**) $\sigma = 0.1$ and $C = 1$.



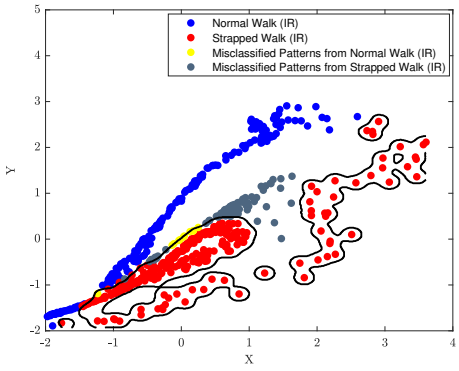
(a) $\sigma = 0.1$ and $C = 10$

Figure 6.35: The decision boundary when the SVM model trained on p_2 normal walk patterns versus that of the strapped one, when: $\sigma = 0.1$ and $C = 10$.



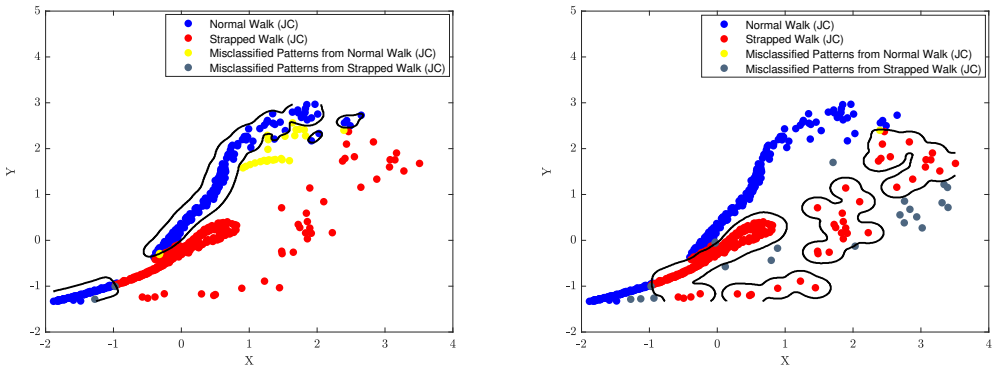
(a) $\sigma = 0.1$ and $C = 0.1$ (b) $\sigma = 0.1$ and $C = 1$

Figure 6.36: The decision boundary when the SVM model trained on p_3 normal walk patterns versus that of the strapped one, when: (**a**) $\sigma = 0.1$ and $C = 0.1$ and (**b**) $\sigma = 0.1$ and $C = 1$.



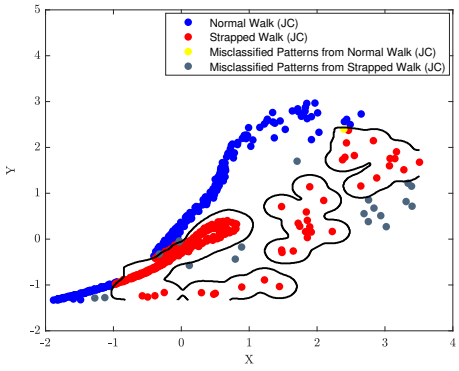
(a) $\sigma = 0.1$ and $C = 10$

Figure 6.37: The decision boundary when the SVM model trained on p_3 normal walk patterns versus that of the strapped one, when: $\sigma = 0.1$ and $C = 10$.



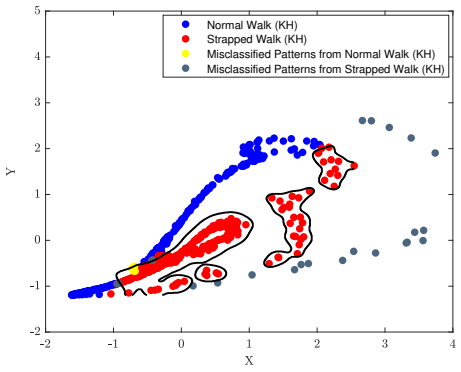
(a) $\sigma = 0.1$ and $C = 0.1$ (b) $\sigma = 0.1$ and $C = 1$

Figure 6.38: The decision boundary when the SVM model trained on p_4 normal walk patterns versus that of the strapped one, when: (**a**) $\sigma = 0.1$ and $C = 0.1$ and (**b**) $\sigma = 0.1$ and $C = 1$.

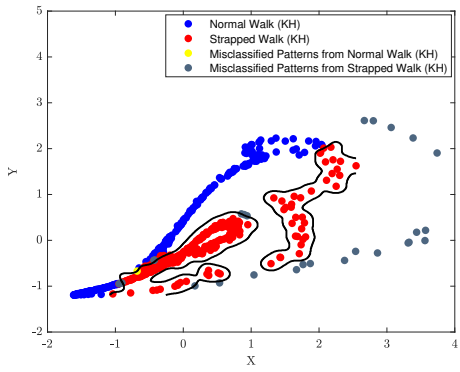


(a) $\sigma = 0.1$ and $C = 10$

Figure 6.39: The decision boundary when the SVM model trained on p_4 normal walk patterns versus that of the strapped one, when: $\sigma = 0.1$ and $C = 10$.

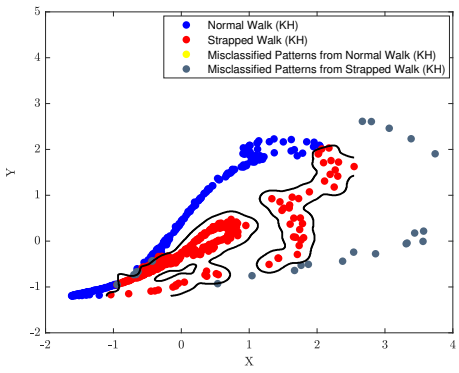


(a) $\sigma = 0.1$ and $C = 0.1$



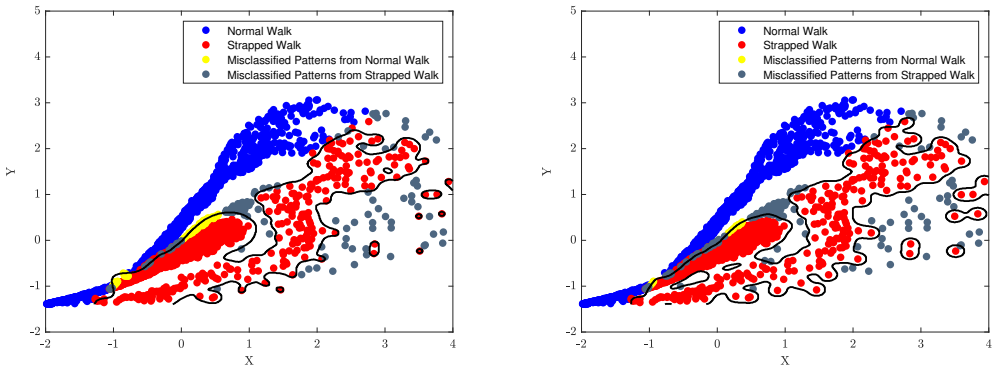
(b) $\sigma = 0.1$ and $C = 1$

Figure 6.40: The decision boundary when the SVM model trained on p_5 normal walk patterns versus that of the strapped one, when: (**a**) $\sigma = 0.1$ and $C = 0.1$ and (**b**) $\sigma = 0.1$ and $C = 1$.



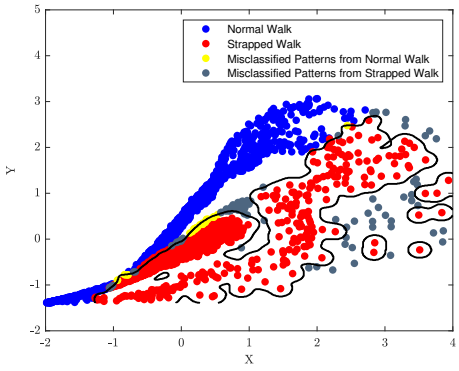
(a) $\sigma = 0.1$ and $C = 10$

Figure 6.41: The decision boundary when the SVM model trained on p_5 normal walk patterns versus that of the strapped one, when: $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$ (b) $\sigma = 0.1$ and $C = 1$

Figure 6.42: The decision boundary when the SVM model trained on all individuals p_{All} normal walk patterns versus that of the strapped one, when: (**a**) $\sigma = 0.1$ and $C = 0.1$ and (**b**) $\sigma = 0.1$ and $C = 1$.



(a) $\sigma = 0.1$ and $C = 10$

Figure 6.43: The decision boundary when the SVM model trained on p_{All} normal walk patterns versus that of the strapped one, when: $\sigma = 0.1$ and $C = 10$.

6.7.5 Mean Square Error Rate and Standard Deviation Test

The figure presented in Figure 6.44 displays the SVM model's average mean square error (MSE) rate for the data of five individuals. To calculate the average MSE for each person, the accuracy obtained for each curve shown in Figures 6.29–6.31 was averaged. The individuals p_1 and p_4 had an average MSE of less than 6% with minimal variation rates, respectively, in their ability to classify patterns of gait abnormalities. On the other hand, when compared to individuals p_2, p_3, and p_5, their average MSE and average standard deviation were found to be extremely high.

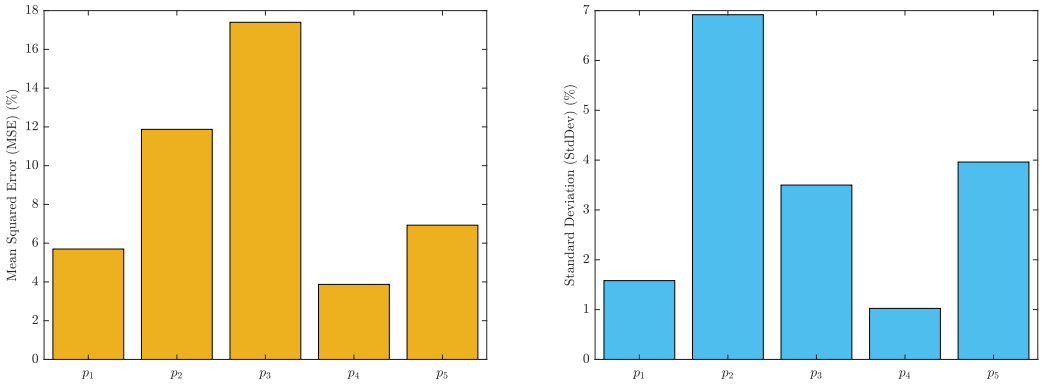


Figure 6.44: This figure shows the average mean square error (MSE)(%) (on the left) and the average standard deviation (%) (on the right) of the SVM classification performance for each individual piece of data.

6.8 Additional Experiment: K-Nearest Neighbour (KNN) Algorithm for Classifying Human Gait Disturbances

K-Nearest Neighbour (KNN) is a supervised machine learning technique that is highly effective in solving classification problems by learning from data. It relies on a binary criterion to model data and assigns categorical labels to input samples. KNN's binary classification basis is particularly useful for detecting and diagnosing gait pattern abnormalities, which can aid in early clinical interventions to prevent the disruption of the biological system and facilitate motor recovery, especially in patients with chronic illnesses who have lost their independence. This type of classification has significant potential for future KNN applications, especially in gait diagnosis. KNN is a reliable gait classifier because it can construct robust predictive models by considering the optimal k nearest Neighbour s in a Euclidean space based on majority votes from Neighbour ing labels.

6.9 Methodology of KNN Experiment

The KNN algorithm was implemented using the entire set of the CA gait features as inputs. A software routine was created in Matlab R2018b to conduct various tests, such as tuning the k parameter through cross-validation to improve classification performance. The KNN model's generalisation performance was determined by calculating the average of the mean square error from both training and cross-validation tests, using different values of k. To ensure fair and independent distributions, the dataset was randomly split into training and testing sets, which is a critical aspect of KNN. This allows unbiased estimation of the generalisation mean error rate, where one portion of data is used for fitting and the other for evaluation. This approach

provides promising results for actual performance.

6.10 Cross Validation

Cross-validation is a widely used method for testing classification models that aims to avoid bias in the performance results. It involves systematically excluding a small portion of the gait data during the training process to maintain the performance of each gait class prediction, while the excluded data points are used to test the trained model. This process is repeated until all data points are included in the testing dataset. Given that the gait dataset comprises only 6000 trials from five individuals, it is crucial to validate the data using different scales and observe the error rates associated with each data split. To draw accurate conclusions about the effectiveness of the KNN classification model, it is essential to test the MoRES dataset across various data splitting scenarios.

6.11 Measuring Mean Square Error Rate

In order to assess the effectiveness of the KNN classification model algorithm, it is essential to evaluate the mean square error (MSE) between the actual data samples and the predicted data points. This measurement allows for a determination of the predictive model's accuracy. Ideally, the model performs well when the average error rate $\mathcal{E}$ is minimised. The MSE can be calculated as follows:

$$\mathcal{E} = \frac{1}{N} \sum_{i=1}^{N} \left(y_i - \widehat{y}_i \right)^2 \tag{6.1}$$

where y_i is the original data samples (N) and $\widehat{y}_i$ is the predicted data points.

6.12 Optimal Value of K in KNN Model

The KNN model, named for its reliance on Neighbour ing data points, requires careful consideration when selecting the optimal value for the parameter, k. This parameter specifies the number of Neighbour s to consider during classification and can significantly impact model complexity and efficacy. The ideal value of k depends on several factors, including dataset size and data point distribution within their coordinate space. To achieve minimal classification error and accurately reflect the KNN model's accuracy, k must be explicitly selected using cross-validation techniques.

6.13 Simulation Results of KNN Experiment

In this experiment, a 70-fold cross-validation test was conducted on 6000 gait data points, divided into approximately 86 subsets. The majority of the data (90%) was allocated for training, while the remaining 10% was used for testing. This experimental setup resulted in minimal test and cross-validation error rates. The K-nearest Neighbour (KNN) model performed optimally with a value of $k = 4$. The model's performance was evaluated using the mean square error quantifiable approach. The percentage of unclassified data points was averaged over 100 runs of the training and testing data sample sets to ensure fair and reliable results.

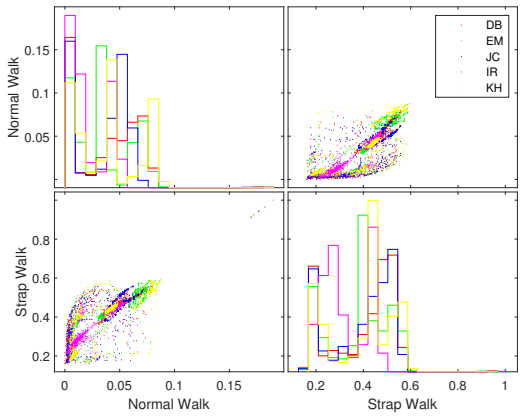


Figure 6.45: Pre-trained class labels associated with their medical conditions

The histogram distribution of MoRES data for normal and strapped gait patterns is depicted in Figure 6.45. The figure clearly illustrates qualitative differences between the two groups, such as increased variability, significant decrease in central tendency, and high skewness in the normal well-apparent plots. The experimentally derived features from Figure 6.45 were utilised to train a KNN model to classify gait patterns corresponding to the five different data labels: EM, IR, KH, DB, and JC. These labels are defined to comply with data protection rules and regulations. The prediction results of the trained KNN model are shown in Figure 6.46. The figure indicates that all label patterns were accurately classified, but the predicted data distribution for strapped gait patterns is clustered more closely and becomes more dense due to the limited mobility caused by stiffness.

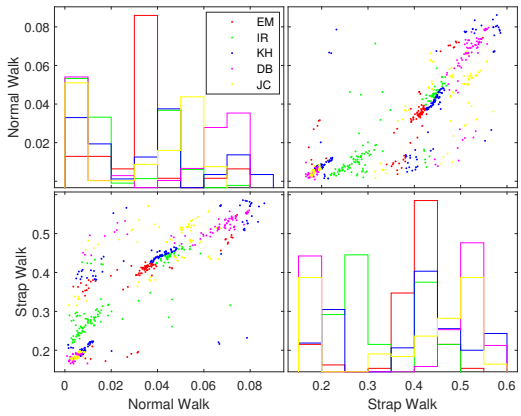


Figure 6.46: Predicted class labels associated with their medical conditions

The confusion matrix shown in Figure 6.47 illustrates the KNN model's capacity to categorise data labels based on their corresponding features. The diagonal elements of the matrix represent the number of instances where the predicted and true class labels agree, indicating how closely the predicted labels align with their actual mean or average values. It also indicates the degree of joint variability of each data label with itself, highlighting well-classified (positively correlated) test data points. Conversely, the off-diagonal elements signify incorrectly classified (negatively correlated) test data samples by the classifier. Notably, there is a pronounced correlation between the IR and KH data labels along the diagonal line, as well as a comparable correlation in medical conditions patterns between the other three labels.

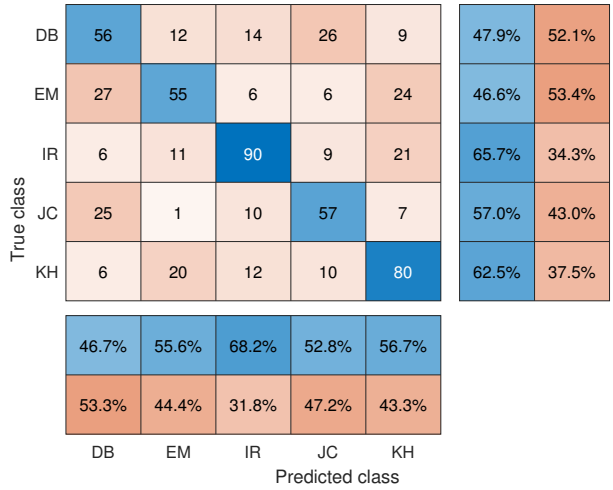


Figure 6.47: Confusion Matrix

The normalised row of the primary matrix presents the true positive and false positive rates as percentages for each genuine category, while the normalised column displays the positive predictive values and false predictive rates as percentages for each predicted category. The accuracy of accurately classified data patterns is estimated to be approximately 67.7%.

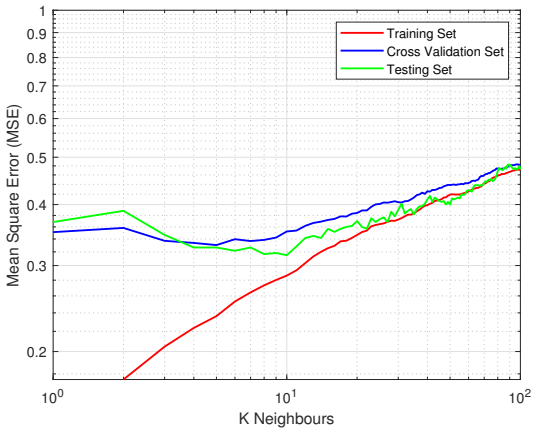


Figure 6.48: MSE rates with respect to various k values

In Figure 6.48, the KNN model underwent 100 iterations for various values of k, with classification mean square error evaluated against cross-validation, training, and testing. Overfitting occurred when k exceeded 4, leading to a significant increase in the error rate. The optimal value of k for minimising cross-validation error was found to be 4, while the lowest testing error was achieved at $k = 10$. In this case, the KNN model displayed reasonably high accuracy and consistency, with significantly low classification error rates. Figure 6.49 displays the observed cross-validation and testing error rates across multiple data splitting ratios.

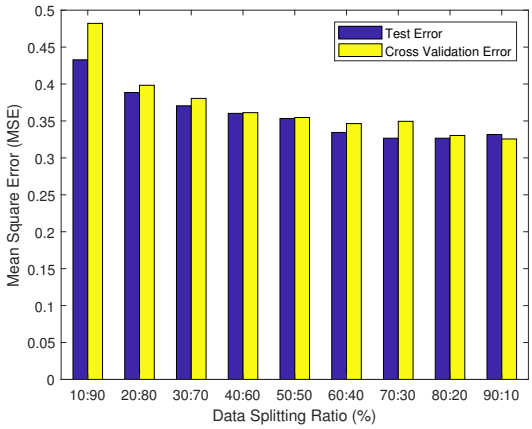


Figure 6.49: MSE rates against different data splitting ratios

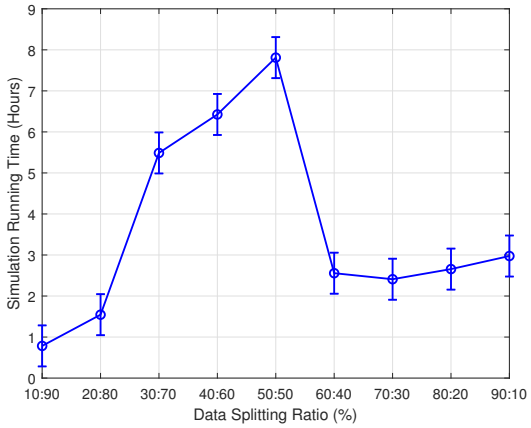


Figure 6.50: Simulation time versus various splitting ratios

The optimal split ratio for MoRES data appears to be a quarterly division of the testing and training sets, resulting in the lowest error rates for both cross-validation and testing. However, it is important to consider the simulation time of the KNN model. As demonstrated in Figure 6.50, the training time gradually increases with larger split ratios. Notably, at the halfway point of the MoRES data splitting, the training time reaches a maximum before decreasing. This figure presents a vertical error bar plot for each data point, where the range of values indicates the length of the error bar above and below the data point, with a total length twice the size of the error value.

6.14 Summary

This chapter presented the implementation of a proposed SVM algorithm, which utilises a criticality analysis technique to classify and detect dynamic abnormalities and disturbance patterns in human gait. The study demonstrates that gait features extracted through the criticality analysis methodology provide valuable information regarding the dynamics of human walking behavior, which can be used to train

a machine learning model to recognise walking deviations from the normal gait of individuals. Early detection of gait disturbances through this supervised pattern recognition technique can enhance the chances of identifying gait impairments and preventing gait injuries. Previous research has utilised Neural Networks (NNs) and fuzzy clustering techniques for automated gait classification to diagnose pathological human gait [53] [54] [55]. However, the SVM algorithm with the support of the CA method demonstrated superior performance in gait classification, making it a robust and reliable tool for detecting dynamic disturbances of biological data patterns. This study creates a tremendous opportunity for clinical diagnosis and rehabilitation.

As explained in this chapter, the CA method has been utilised to extract data that represents various gait patterns and feed them into SVM. This approach serves a dual purpose. Firstly, the CA-extracted data offer a more realistic measure of the kinematic motor system and provide a comprehensive representation of the dynamics that impact individual gait. Secondly, the SVM's kernelised nonlinearity can effectively classify the CA data patterns, which are inherently nonlinear. Additionally, the CA methodology generated sufficient data samples that can facilitate a proper analysis of human gait. The distribution of data can deteriorate, and classification performance can suffer when there is an insufficient number of gait samples due to short data measurements. To obtain more stable and controlled CA data patterns of human gait, the pre-processed raw data is recorded for a longer walk cycle period of 6 seconds, producing a reliable and useful statistical distribution of gait patterns.

The effective performance of the proposed Kernel SVM in classification is primarily dependent on the appropriate selection of the regularisation parameter C, as demonstrated in Figures 6.29– 6.31 and Tables B.2–B.12 of Appendix B. The penalty parameter C compensates for misclassification accuracy, and careful selection is necessary to achieve optimal performance. Comparing the performance of SVM at different C values corresponding to different σ control width values, it was found that SVM performed best when $\sigma = 0.1$, followed by $\sigma = 1$, as smaller values

of σ result in faster decline of the Kernel Gaussian function. The Kernel Gaussian function proved to be the most effective in detecting abnormal gait data patterns, as evident from the recorded accuracy at these specific values. Furthermore, Figures 6.29– 6.31 indicate that the optimal value of C varies according to individual data patterns. C can be selected using a trial and error method, and one way to accomplish this is by visualising the dependence of classification performance on C.

The results of this chapter indicated that the proposed SVM models, supported by the CA method, can serve as a reliable and robust tool for detecting dynamic disturbances in human gait patterns. To achieve good generalisation performance, it is essential to target the best features that can contribute to improved performance. Notably, the CA method is suggested in this thesis for feature selection, which can represent the entire data set without excluding any features. The findings presented in Figures 6.2– 6.7 demonstrated that changes in gait patterns can be detected and distinguished even without machine learning training. Also, the scatter and histogram plots of CA gait features extracted from dynamic gait data were helpful in distinguishing between normal and abnormal gait patterns. This indicates that changes in gait patterns due to an individual's condition or behavior are reflected in their visual patterns. The histogram plots presented in Figures 6.17– 6.22 provide a statistical interpretation of the CA data distribution, while the scatter plots illustrated in Figures 6.8– 6.13 describe the variability of the CA gait data in space. These plots are useful for detecting gait abnormalities and monitoring the progress of individual walks in clinical settings. The CA gait features, represented by the distribution and variability measures, serve as the basis for the classification process between the two distinct gait patterns.

Chapter 7

Prediction of Human Age Based on Human Gait

7.1 Introduction

Human gait analysis is essential in many areas of healthcare, including rehabilitation, geriatrics, and sports medicine. As we age, our gait patterns undergo changes, which can be indicative of various age-related health conditions [211]. Detecting these changes is crucial for the diagnosis and management of these conditions. However, identifying gait patterns based on age can be challenging, and traditional methods of analysis may not be effective in distinguishing age-related changes. This is where criticality analysis, a multivariate data representation tool, comes into play.

The criticality analysis approach is based on the concept of criticality, which refers to the state of a system when it is at the brink of a transition from one state to another. In the case of human gait analysis, criticality analysis considers gait patterns as a complex system, which undergoes changes with age. The approach analyses the system properties in a multivariate space, enabling the detection of subtle differences that may not be evident through traditional analysis methods.

The rationale behind using criticality analysis in gait analysis is its ability to identify associated factors and distinguish between gait patterns based on age. For instance, criticality analysis can help identify age-related changes in gait that are not directly associated with age but may be caused by other factors such as obesity, diabetes, or musculoskeletal disorders. By analysing gait patterns in a multivariate space, criticality analysis can distinguish between gait patterns that are directly related to age and those that are associated with other factors.

An example of how criticality analysis could be used is in the diagnosis of Parkinson's disease, a neurodegenerative disorder that affects the nervous system, resulting in tremors, stiffness, and balance problems. Parkinson's disease is often diagnosed by observing changes in gait patterns. However, these changes may not be apparent in the early stages of the disease, making diagnosis difficult. By using criticality analysis, gait patterns can be analysed in a multivariate space, enabling the detection of subtle changes that may indicate the early stages of Parkinson's disease.

This chapter uses the CA system's representationof gait data to train an SVM classifier, which can subsequently be employed to classify new representations. Similar to binary comparisons, the aim was to develop an age-category classifier based on gait data. The chapter simulation results demonstrated that the SVM model, supported by the CA approach, significantly improves age classification accuracy and efficiency. Hence, it serves as a reliable classification tool for detecting human ages based on gait data patterns, presenting an immense opportunity for future clinical diagnosis and rehabilitation.

7.2 Methodology

The proposed methodology framework is detailed in this section, which includes data collection, data processing, feature extraction techniques, and implementation of the SVM classifier. Figure 7.1 illustrates the flowchart of this methodology.

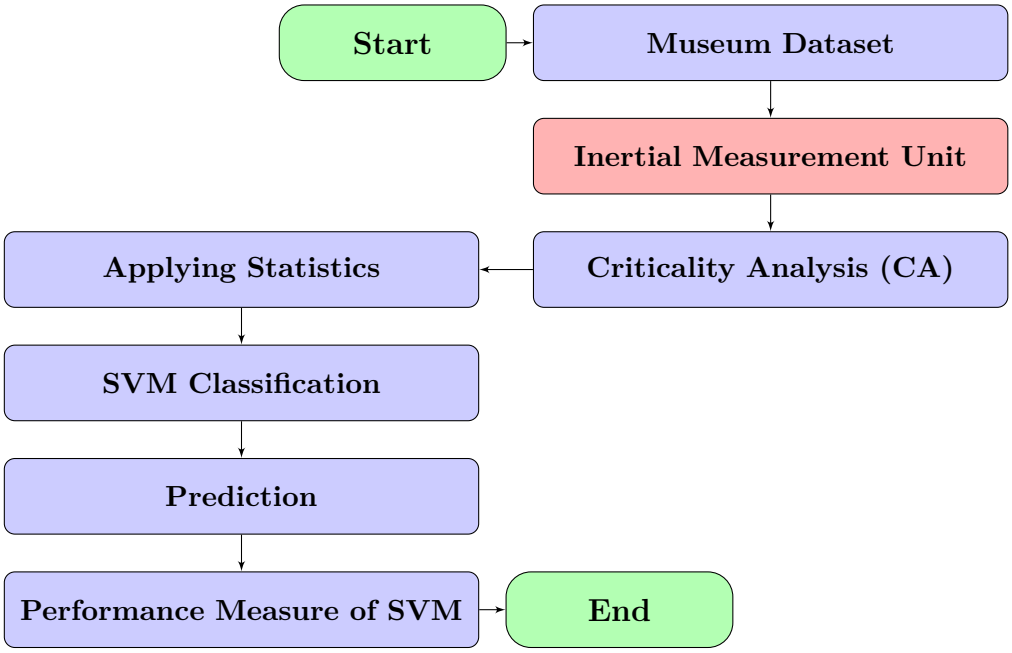


Figure 7.1: The flowchart of the proposed methodology of Museum Dataset.

7.3 Museum Dataset

The Science Museum in London, United Kingdom sought the participation of individuals between the ages of 5 and 80 in its Live Sciences section. Prior to participation, informed consent was obtained with approval from the Oxford Brookes University Research Ethics Committee (UREC). As part of the research work of this thesis, Oxford Brookes University received access to the Museum dataset for research purposes, which has been anonymised in accordance with the regulations outlined in the General Data Protection Regulation (GDPR)

7.3.1 Data Collection

The study collected information about the participants, which included age in years, gender, country of birth, height in meters, weight in kilograms, foot length based on shoe size, and leg length in meters. The participants were asked to walk a distance of 10 meters at their preferred speed while wearing an inertial measurement unit on their lower lumbar spinal region [206], [212]. Gait models that rely on leg length [213] and foot size [214], [215], [216] were used to measure the spatiotemporal features of gait, such as step time in milliseconds, cadence in steps per minute, step length and stride length in meters, as well as walking speed in meters per second. These gait characteristics were established and validated by means DGAS software.

7.3.2 Data Analysis and Features Extraction

The Museum dataset comprises unprocessed data collected from a 3D accelerometer, gyroscope, and magnetometer IMU sensor unit. To derive spatiotemporal features, the 3D IMU data underwent a transformation using the DGAS software discussed in Section 5.3.1 and described in Section 7.3.1. The DGAS feature extraction tool produced an output dataset consisting of two gender categories and five age groups: Children, Adolescents, Young Adults, Middle Aged, and Old Aged. The CA methodology, as described by equations (3.10)–(3.13), was applied to the output dataset. This transformed the spatiotemporal measurements into a two-feature representation, namely f (x-axis) and m (y-axis). The resulting phase space portraits for both features were analyzed, revealing distinctive changes in the structural patterns of human gait dynamics.

7.3.3 Criticality Analysis as a Data Representation Method

Using the Criticality Analysis method, it is possible to create a controlled self-organised critical system from RCC-controlled networks of oscillators. This system

can be used to represent arbitrary data in a nonlinear space, allowing for easy classification without the need for training. Each data sample has a unique orbit that is a result of perturbing the critical system underlying it. This system is considered critical because it is composed of a network of nonlinear controlled oscillators, which results in emergent critical properties. Even small changes in inputs can cause state changes in the total network's orbit, but it remains stable. The resulting orbit represents the original data in a scale-free manner, which allows for comparing the gait of an individual patient or control to the matching category of the sample versus other data members [50]. This representation also helps to reduce the dimensionality of the dynamic data.

The phase space plots in Figures 7.2–7.3 illustrate the walking differences between gender and age groups in the CA data representation. The plots at the bottom right of the Figures 7.2–7.3 show the walking patterns associated with each gender and age group. The male age groups exhibit greater variability in their walking patterns compared to the female age groups [217]. The visualizations also reveal that the walking rate for the elderly group is generally lower than the other groups, and has more variations. This rate is particularly lower for males due to their less frequent movement or normal walking routine. These findings suggest that Neighbour hood environments can influence the walking behavior of men and women across different age groups similarly. The CA dynamic representation of human gait data can be used to compare and monitor changes in walking behavior over time based on age and gender. The phase plots of Figures 7.2–7.3 demonstrate that even small or large impacts to the motor system resulting from psychological or environmental factors can affect the gait behavior patterns across age and gender groups, as reflected in the CA response of each walk pattern.

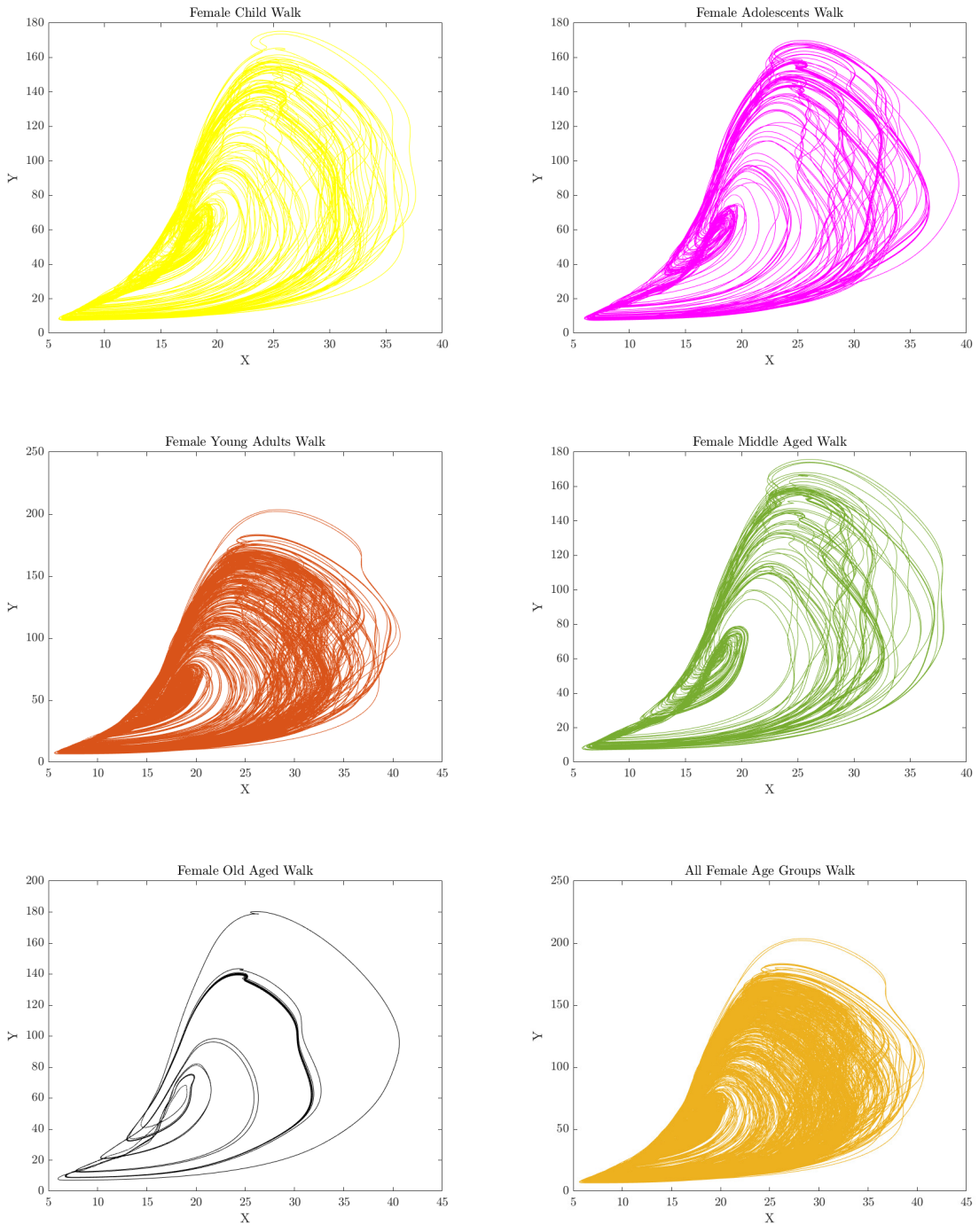


Figure 7.2: This figure shows the phase space plots for each walk pattern that corresponds to each age category of the female group.

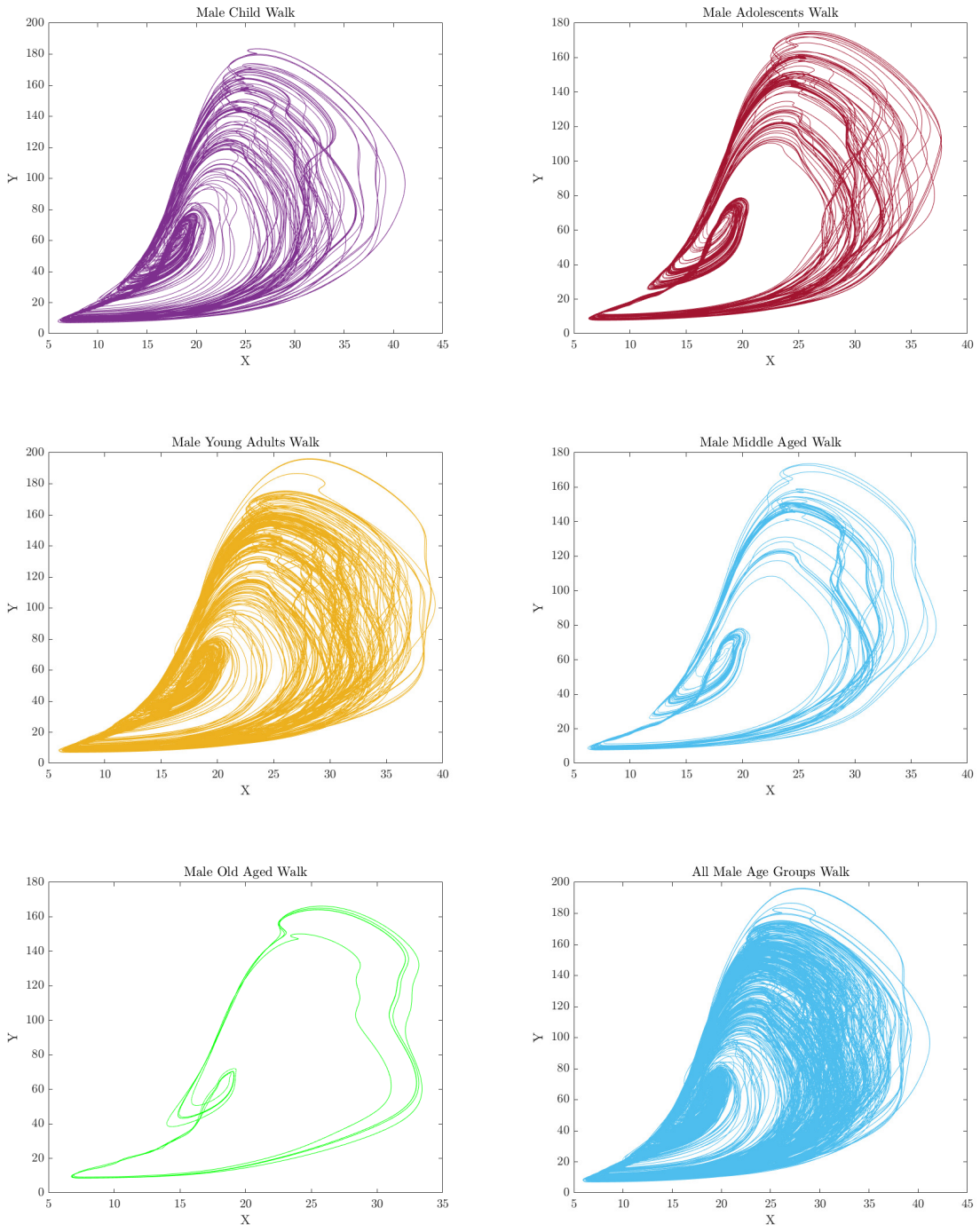


Figure 7.3: This figure shows the phase space plots for each walk pattern that corresponds to each age category of the male group.

7.4 Statistical Analysis of Criticality Analysis Data

7.4.1 Scatter Analysis

Scatter analysis can be used to compare the gait patterns of different age and gender groups, and to identify any differences or similarities in gait behavior. It can also be used to monitor changes in gait patterns over time, such as changes in response to injury or rehabilitation. Additionally, scatter analysis can provide insights into the underlying mechanisms that influence gait, such as changes in muscle strength, balance, and coordination.

Figures 7.4–7.9 illustrate the distribution of CA data representation for different age/gender walk patterns, corresponding to the visualisations shown in Figures 7.2–7.3. Each graph represents the changes in successive gait interaction patterns over time. The CA system demonstrates a strong correlation between the two dominant features extracted from the CA model, which is measured using the Pearson's correlation coefficient ρ, ranging between 0 and 1. Tables 7.1 and 7.2 provide the values of ρ for different walk patterns and age/gender categories. Table 7.1 provides the correlation coefficient ρ for different age groups and gender categories. The values of ρ range from 0.713 to 0.896 and indicate a strong positive correlation between the two dominant features extracted from the CA model. Also, Table 7.2 summarises the correlation coefficient ρ for all age groups and gender categories, which is 0.807 for female walk patterns and 0.811 for male walk patterns. These values suggest a strong positive correlation between the two dominant features for all age groups and gender categories.

Table 7.1: The correlation coefficient ρ of female and male age groups.

Category	**Correlation Coefficient (ρ)**				
	Child	Adolescents	Young Adults	Middle Aged	Old Aged
Female Walk	0.873	0.713	0.822	0.821	0.875
Male Walk	0.769	0.844	0.807	0.829	0.896

Table 7.2: The correlation coefficient ρ of walk patterns for each age group in total.

Category	**Correlation Coefficient (ρ)**
	All Age Groups
Female Walk	0.807
Male Walk	0.811

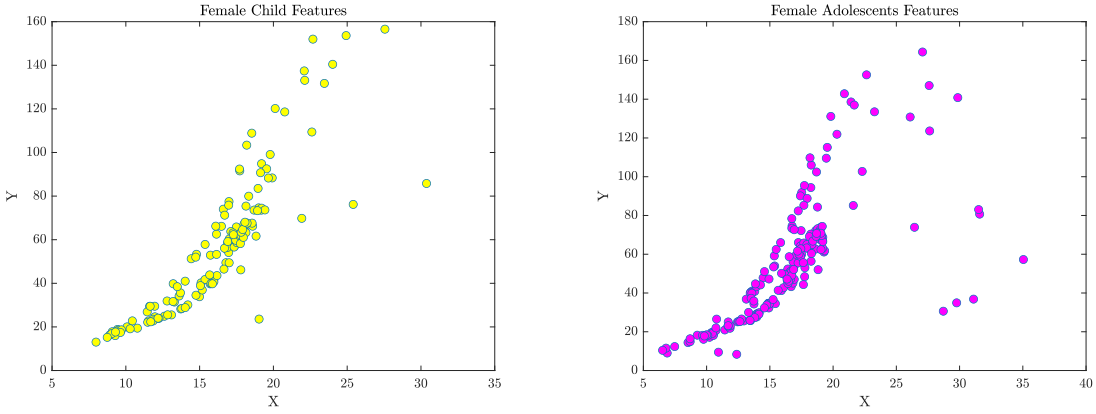


Figure 7.4: This figure shows the scatter plots for each walk pattern that corresponds to each age category (Children and Adolescents) of the female group.

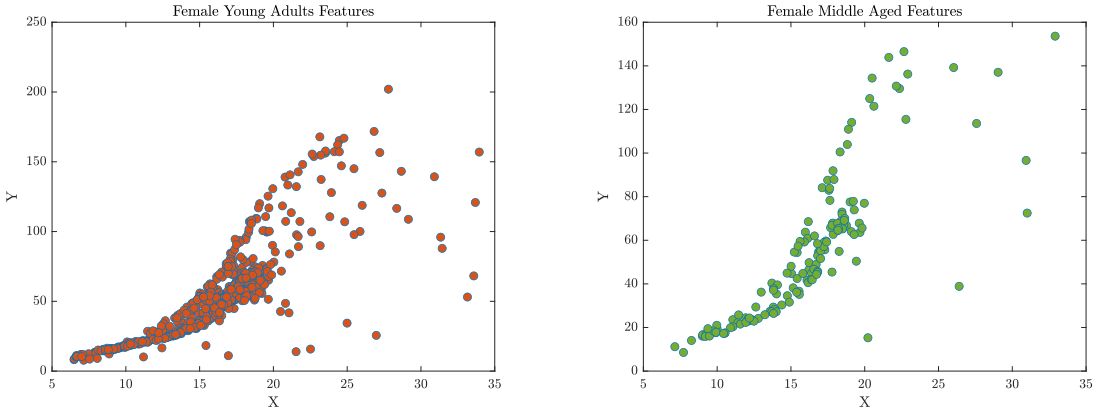


Figure 7.5: This figure shows the scatter plots for each walk pattern that corresponds to each age category (Young Adults and Middle Age) of the female group.

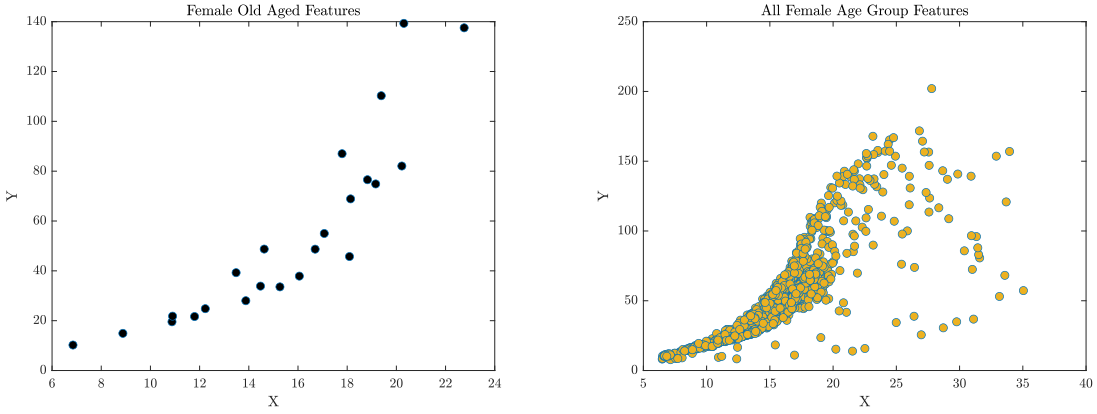


Figure 7.6: This figure shows the scatter plots for each walk pattern that corresponds to each age category (Old Aged and All Ages Group) of the female group.

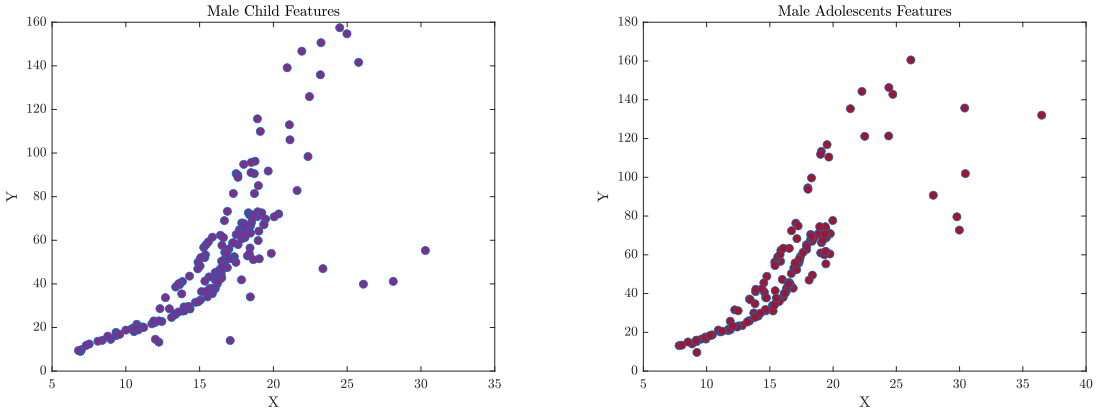


Figure 7.7: This figure shows the scatter plots for each walk pattern that corresponds to each age category (Children and Adolescents) of the male group.

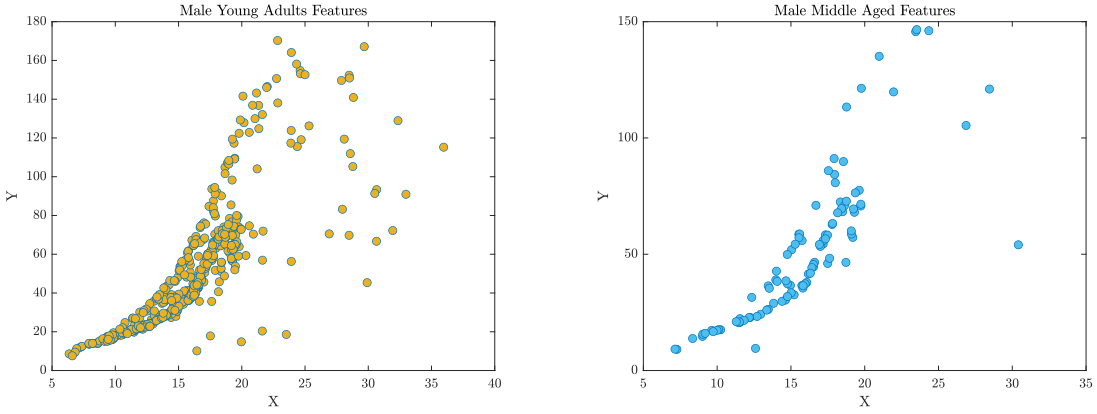


Figure 7.8: This figure shows the scatter plots for each walk pattern that corresponds to each age category (Young Adults and Middle Age) of the male group.

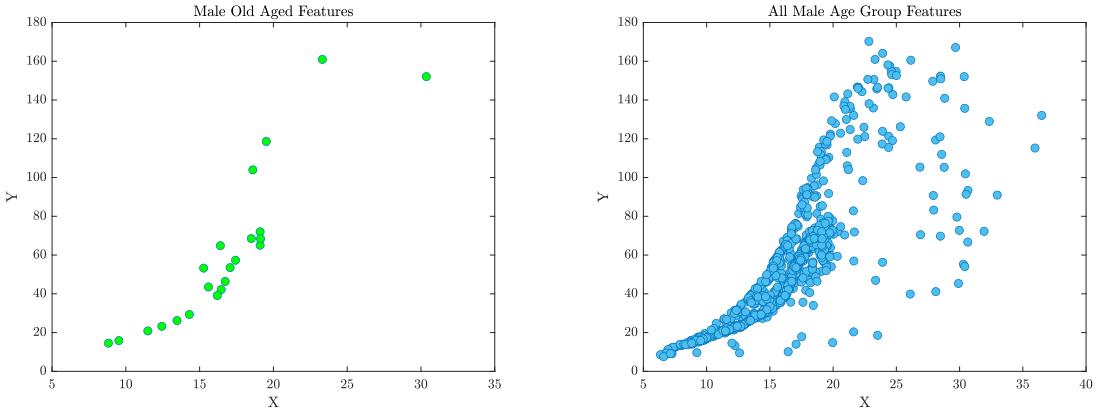


Figure 7.9: This figure shows the scatter plots for each walk pattern that corresponds to each age category (Old Age and All Ages Group) of the male group.

7.4.2 Spatiotemporal Analysis

Spatiotemporal analysis refers to the study of how spatial and temporal aspects of human movement change over time. In the context of analysing human gait based on age-gender groups, spatiotemporal analysis involves measuring and quantifying the

changes in the patterns of movement that are specific to each age-gender group. This may include analysing the speed, stride length, step width, and other parameters of gait that vary with age and gender. By using spatiotemporal analysis, useful insights can be gained into how gait patterns change as people age, and how these changes differ between males and females. These insights can be useful for developing interventions to improve gait and mobility in older adults and other populations.

Figure 7.10 displays the location and fluctuations of gait patterns for each age/-gender group. The middle line in each categorical box represents the median. The extent of variation in gait data between categories in each age/gender group determines the degree of dispersion and skewness of gait in various individuals. Table 7.3 presents the median values of CA gait data for each age/gender category. The measurements reveal that the median value for male gait is slightly higher than that of female walk patterns. Moreover, Table 7.4 shows that the median for the combined age group of male walk patterns is higher than that of female walk patterns. The parallel whisker lines illustrate the range of gait data for each age/gender category beyond the quartile interval. The CA gait data for females in each age group are devoid of outliers, whereas some undesired data points exist between certain age groups of males' walk patterns.

Figure 7.10 shows the location and variations of gait patterns across age/gender category. The middle line of each categorical box represents the median. The amount of difference in the variability of gait data between each category for each age/gender group measures the degree of dispersion and skewness of various individuals' gait. Table 7.3 describes the median values of CA gait data for each age/gender category. The measurements indicated that the average median of the male gait is slightly higher than that one of the female walk patterns. Also, the median of the combined age group of male walk patterns shows higher records when compared with that of the female walk patterns, as shown in Table 7.4. Furthermore, the parallel lines of whiskers depict the variation of gait data for each age/gender category outside the

limits of the quartiles interval. Also, the CA gait data of females per age category are free of outliers, while some unwanted data points exist between some age groups of males walk patterns.

Table 7.3: The median of female and male age groups.

Category	**Median**				
	Child	Adolescents	Young Adults	Middle Aged	Old Aged
Female Walk	19.72	19.56	19.56	19.85	19.89
Male Walk	19.54	19.71	19.77	19.49	20.18

Table 7.4: The median of walk patterns for each age group in total.

Category	**Median**
	All Age Groups
Female Walk	19.63
Male Walk	19.69

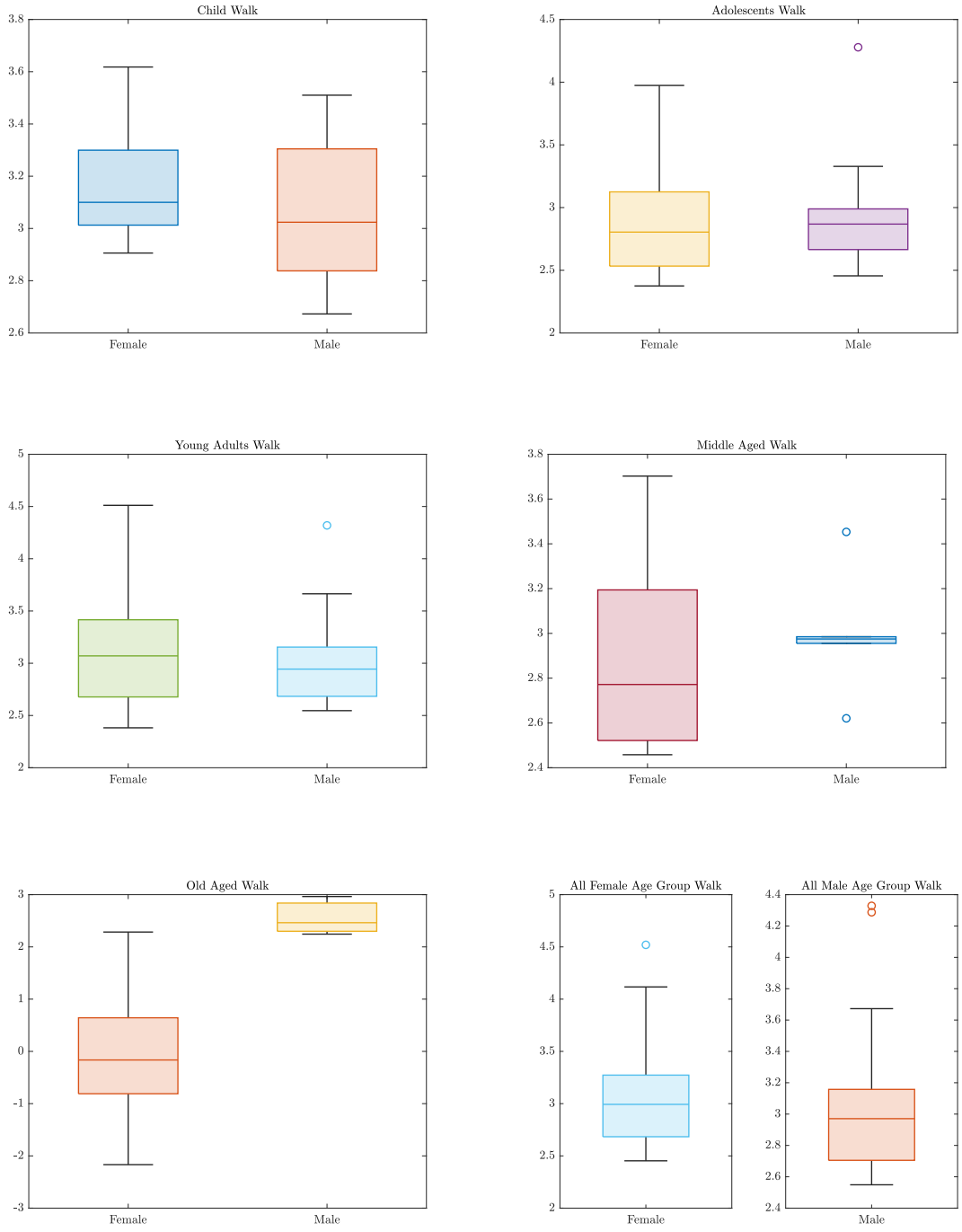


Figure 7.10: This figure shows the box plots for each walk pattern that corresponds to each age category of the female group.

7.4.3 Histogram Analysis

In the context of human gait analysis based on age-gender groups, histogram analysis involves examining the distribution of data values for each age-gender group to identify patterns and differences in gait characteristics. The histogram plots of both female and male gait features and their corresponding frequencies, which describe the distribution of CA gait data in space, are graphically represented in Figures 7.11–7.12. Table 7.5 presents the mean values of CA gait data for each age/gender category. The measurements show that, on average, females have a slightly longer stride length than males, except for older age groups where males have a longer stride. However, the mean of the combined age group of male walk patterns is lower than that of the female walk patterns, as presented in Table 7.6.

Table 7.5: The mean of female and male age groups.

Category	**Mean**				
	Child	Adolescents	Young Adults	Middle Aged	Old Aged
Female Walk	36.04	35.85	35.09	35.93	35.17
Male Walk	34.78	35.11	34.76	34.01	38.83

Table 7.6: The mean of walk patterns for each age group in total.

Category	**Mean**
	All Age Groups
Female Walk	35.46
Male Walk	34.83

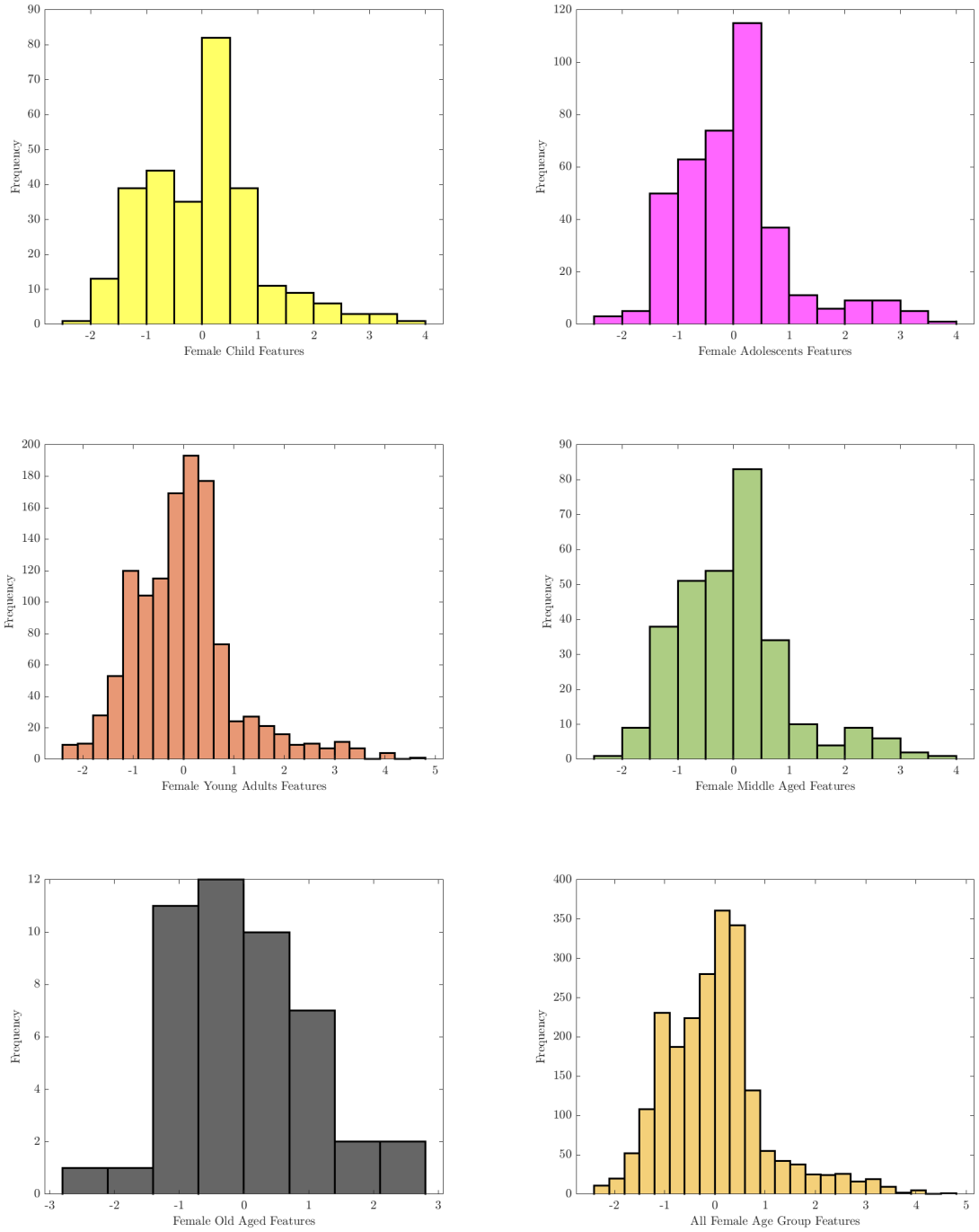


Figure 7.11: This figure shows the histogram plots for each walk pattern that corresponds to each age category of the female group.

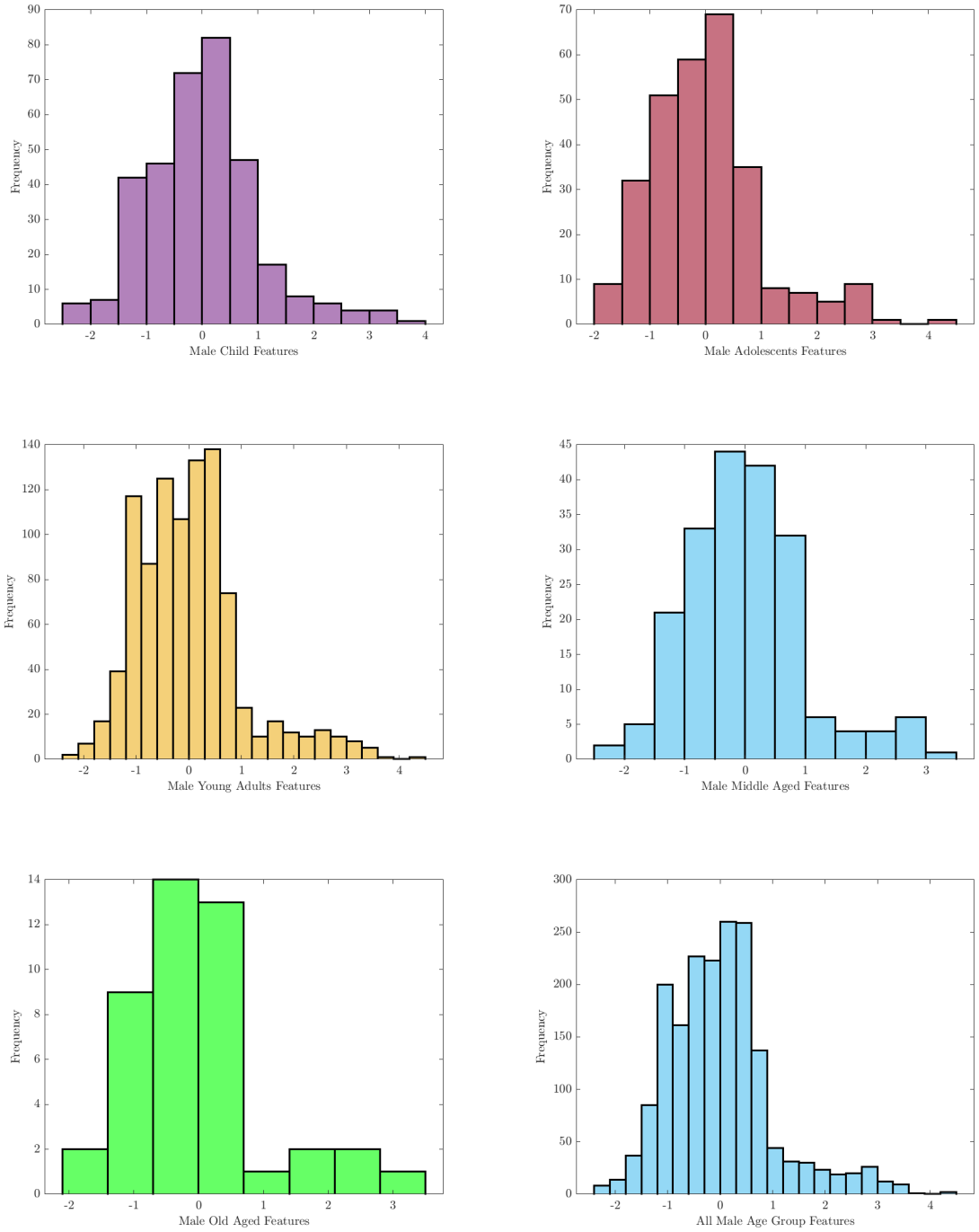


Figure 7.12: This figure shows the histogram plots for each walk pattern that corresponds to each age category of the male group.

7.5 The Proposed SVM Classifier

In order to fully validate the efficacy of our CA method in terms of detecting and classifying age based on human gait, we conducted a series of experiments using various models based on the SVM's Kernelised property, as explained in Section 4.3.2. These experiments included classifying a single age group against other age groups within the same gender category, classifying an age group from one gender against the corresponding age group from the opposite gender, and classifying the combined age groups from one gender against the combined age groups from the other gender.

7.6 The Proposed SVM Training

7.6.1 Experimental Settings

Based on the details oulined in section 7.3.1, Museum dataset is obtained. It is made up of 2019 data samples, distributed as follows: 1105 data points for female group and 914 for male group. There are five age categories (Children [2–12 Yrs], Adolescents [13–18 Yrs], Young Adults [19–44 Yrs], Middle Ages [45–64 Yrs], and Old Aged [65+ Yrs]) that correspond to each gender group. Four experimental scenarios have been conducted, aiming at:

- **Scenario 1:** Classify age category per female gender group, for example: (x_{f1}) Female Children as a training set versus Female Adolescent as a testing set, (x_{f2}) Female Children as a training set versus Female Young Adults as a testing set, (x_{f3}) Female Children as a training set versus Female Middle Aged as a testing set, (x_{f4}) Female Children as a training set versus Female Old Aged as a testing set, (x_{f5}) Female Adolescent as a training set versus Female Young Adults as a testing set, (x_{f6}) Female Adolescent as a training set versus Female Middle Aged as a testing set, (x_{f7}) Female Adolescent as a training set versus

Female Old Aged as a testing set, (x_{f8}) Female Young Adults as a training set versus Female Middle Aged as a testing set, (x_{f9}) Female Young Adults as a training set versus Female Old Aged as a testing set, and (x_{f10}) Female Middle Aged as a training set versus Female Old Aged as a testing set. So in total, 10 SVM models have been implemented for each scenario.

- **Scenario 2:** Classify age category per male gender group, for example: (x_{m1}) Male Children as a training set versus Male Adolescent as a testing set, (x_{m2}) Male Children as a training set versus Male Young Adults as a testing set, (x_{m3}) Male Children as a training set versus Male Middle Aged as a testing set, (x_{m4}) Male Children as a training set versus Male Old Aged as a testing set, (x_{m5}) Male Adolescent as a training set versus Male Young Adults as a testing set, (x_{m6}) Male Adolescent as a training set versus Male Middle Aged as a testing set, (x_{m7}) Male Adolescent as a training set versus Male Old Aged as a testing set, (x_{m8}) Male Young Adults as a training set versus Male Middle Aged as a testing set, (x_{m9}) Male Young Adults as a training set versus Male Old Aged as a testing set, and (x_{m10}) Male Middle Aged as a training set versus Male Old Aged as a testing set. So in total, 10 SVM models have been implemented for each scenario.

- **Scenario 3:** One-to-one mapped classification per age category between female and male gender groups, for example: (x_{fm1}) Female Children as a training set versus Male Children as a testing set, (x_{fm2}) Female Adolescents as a training set versus Male Adolescents as a testing set, (x_{fm3}) Female Young Adults as a training set versus Male Young Adults as a testing set, (x_{fm4}) Female Middle Aged as a training set versus Male Middle Aged as a testing set, and (x_{fm5}) Female Old Aged as a training set versus Male Old Aged as a testing set. So in total, 5 SVM models have been implemented for each scenario.

- **Scenario 4:** (x_{All}) All Female age groups are trained against all Male age

groups, which used as a testing set.

The RCC model, specified by equation (3.10)–equation (3.13), utilises the 24 gait parameters of the Museum dataset as perturbation input to extract the most significant gait features (f from equation (3.11) on the x-axis and m from equation (3.10) on the y-axis). These extracted features are highly nonlinear and correlated, and thus a Kernel SVM is proposed to transform the dynamic gait data samples into a new feature space to aid in classifying various data subjects. The SVM algorithm is trained and tested using the extracted features as input, with the third scenario being implemented to reduce the number of parameters and avoid overfitting. The main objective of performing the three experiments is to assess the SVM's classification performance and ability to classify human age based on walk behavior. Algorithm 1 details the implementation steps of the proposed SVM algorithm.

The CA gait features underwent normalisation using the z-score to standardise the data samples to have a mean of zero and a standard deviation of one prior to training the SVM algorithm. The Kernel SVM algorithm was employed in MATLAB to perform statistical analysis of the gait data and create various Kernel SVM models to investigate the impact of Kernel characteristics (such as the regularisation parameter C and control width σ) on the classification performance of gait data. The trained Kernel SVM's ability to generalise was evaluated by determining the prediction accuracy for each model, analysing the confusion matrix in Appendix C, generating the Receiver Operating Characteristics (ROC) curve, and calculating the area under the ROC curve. Further details on these procedures are discussed in subsequent sections.

7.7 Simulation Results

In order to thoroughly assess the efficacy of our proposed methodology, the experimental results for each individual task in the subsequent subsections will be

presented in-depth.

7.7.1 Confusion Matrix

When dealing with a binary classification task, the labels within the dataset are limited to either positive or negative. The output of the SVM classifier can be analysed through a well-defined contingency table called a confusion matrix, which contains information regarding the actual and predicted classifications. This matrix can provide insights into the effectiveness of the classification algorithm in distinguishing between different categories, such as children or adolescents walk. According to sources cited as reference [52], the confusion matrix can be characterised by four primary properties:

- **True Positive** (TP): refers to the number of positive data samples (participants of targeted age walk) that are correctly classified and detected as positive (targeted age).
- **False Positive** (FP): refers to the number of negative data samples (participants of non-targeted age walk) that are incorrectly classified as positive (targeted age).
- **True Negative** (TN): refers to the number of negative data samples (participants of non-targeted age walk) that are correctly classified as negative (non-targeted age).
- **False Negative** (FN): refers to the number of positive data samples (participants of targeted age walk) that are incorrectly classified as negative (non-targeted age).

The other performance metrics of a classifier are accuracy, precision, F_1-Score, Recall or Sensitivity or True Positive Rate (TPR), Specificity or True Negative Rate

(TNR), and False Positive Rate (FPR), which are evaluated on the basis of the above-stated TP, FP, TN, and FN numbers. Their definitions are as follows:

- **Accuracy**: represents the overall success rate of correct predictions. It is mathematically expressed as the ratio between the correctly classified labels ($TP + TN$) and the total number of data samples ($TP + TN + FP + TN$). It is given by:

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

- **Precision**: is the ratio of correctly predicted positive labels with targeted walk status to the total labels predicted to have a targeted age walk condition. It is measured by the following expression:

$$Precision = \frac{TP}{TP + FP}.$$

- **Recall (TPR)**: is defined as the proportion of correctly predicted positive labels with targeted age walk status among all targeted age walk data samples. It is measured by the following expression:

$$Recall = \frac{TP}{TP + FN}.$$

- **F_1-Score**: is known as the F measure, which preserves the equilibrium between the precision and the recall. It is measured as:

$$F_1 - Score = \frac{2 \times Precision}{Precision + Recall}.$$

- **True Negative Rate (TNR)**: is defined as the proportion of correctly predicted negative labels with non-targeted age walk conditions among all non-targeted age walk data samples. It is formulated as:

$$TNR = \frac{TN}{TN + FP}.$$

- **False Positive Rate (FPR)**: is measured as the proportion of the non-targeted age walk samples mislabeled as targeted age walk data among all the non-targeted age data samples. It is given as:

$$FPR = \frac{FP}{FP + TN}.$$

The metrics outlined above serve the purpose of evaluating the ability of the proposed SVM algorithm to differentiate targeted age gait patterns from non-targeted walk patterns within each gender/age group. Appendix C summarises the SVM classifier's performance metrics for each individual at different values of the regularisation parameter $C = (0.1, 1, 10)$ in Tables C.1–C.20. The primary validation metric that provides insight into the overall performance of the SVM classifier is accuracy. The accuracy of the classifier depends on the values of σ and C, and the best performance was achieved when $\sigma = 0.1$ and 1. For instance, Tables C.51– C.52 demonstrate how to assess the overall performance of the confusion matrix parameters through reflection. Tables C.51– C.52 present the SVM classification results of x_{All} walk patterns at different values of C and σ. The values of true positive (TP), false positive (FP), false negative (FN), and true negative (TN) are presented along with the performance measures, such as precision, recall, F1-score, specificity, area under the ROC curve (AROC), and accuracy. At $\sigma = 10$, the values of TP, FP, FN, and TN are the same for all values of C. However, the performance measures are not meaningful because the values of specificity and AROC are zero. This indicates that the model is not capable of distinguishing between positive and negative cases. Also, at $\sigma = 0.1$ and $\sigma = 1$, the performance measures improve significantly. At $\sigma = 0.1$, the model achieves high precision, recall, F1-score, specificity, AROC, and accuracy for all values of C. Similarly, at $\sigma = 1$, the model achieves high precision, recall, F1-score, AROC, and accuracy for all values of C. However, the specificity is relatively low for $C = 0.1$ and $C = 1$. This indicates that the model has a higher false positive rate for these values of C.

7.7.2 Receiver Operating Characteristics (ROC) Curve

The ROC curve is a useful tool for assessing the performance of the SVM classifier, as it illustrates the relationship between the true positive rate (TPR) (sensitivity) and the complement of the true negative rate ($\approx 1 - TNR$) at various threshold levels ranging from 0 to 1. Section 7.6.1 outlines the experimental settings, and the following is a summary of the results obtained.

Scenario 1: The ROC curves depicted in Figures C.1–C.10 illustrate the results for the optimal σ and C values that yield the highest accuracy for female age groups. These curves indicate that the Kernel SVM approach performs well in most cases when it comes to detecting the age categories of female individuals. The SVM model achieves its best performance when σ takes values of 0.1 and 1 for various values of C, as evidenced by the results presented in Tables C.2– C.20.

Scenario 2: Figures C.11–C.20 display ROC curves generated for the optimal σ and C values that maximise accuracy within each male age group. The Kernel SVM method proves to be highly effective in detecting male individuals across a range of age categories, as evidenced by the majority of cases. Specifically, the SVM model achieves optimal performance when $\sigma = 0.1$ and 1 across various C values, as indicated in Tables C.22– C.40.

Scenario 3: The figures, labeled as Figures C.21–C.25, exhibit ROC curves for the optimal pair of σ and C values that yield the highest accuracy for female age groups compared to their corresponding age groups of male individuals. The results demonstrate that the Kernel SVM performed well in detecting similar age categories between males and females. The SVM model's optimal performance is achieved when σ values of 0.1 and 1 are paired with various values of C, as illustrated in Table C.50.

Scenario 4: The ROC curve illustrated in Figure C.26 depicts the optimal pair of σ

and C values that achieve the highest accuracy across all male and female age groups. Notably, the Kernel SVM exhibited strong performance in detecting various age categories of male individuals. Specifically, the SVM model demonstrated optimal performance for a range of C values, particularly when σ was set to either 0.1 or 1, as indicated in Tables C.52.

7.7.3 Area Under the ROC Curve (AROC)

The SVM classifier's ability to distinguish between different age categories for males, females, and both genders can be measured using the Area Under the Curve (AUC) estimation or ROC area. A higher AUC value indicates a better classifier performance. In Section 7.6.1, we present four scenarios, and the ROC area of the Gaussian kernel is shown as a function of the regularisation parameter C for each age group in Figures C.27– C.39. It is important to note that the C parameter can affect the SVM classifier's performance, and there is a specific value or range that can maintain better efficiency than others. The analysis of Tables C.2–C.20 reveals that the SVM classifier demonstrates superior performance at $\sigma = 0.1$, and occasionally at $\sigma = 1$, in contrast to other values of σ. The corresponding Figures C.27–C.39 were generated based on these results, indicating that certain C values can produce optimal performance. Additionally, some age groups exhibit greater Area under the Curve and superior performance with higher sensitivity levels. The measurements of the SVM model in Tables C.2–C.52 confirm that the AUC and accuracy values are similar, implying that the proposed SVM algorithm is highly efficient at differentiating between normal and strapped gait disturbances. The optimal values utilised to generate Figures C.27– C.39 are indicated in the description of each graph. The AUC was calculated using our MATLAB software routines.

7.7.4 Classification Decision Boundary

The SVM finds a decision boundary that separates the data points of different classes. In SVM, the decision boundary is known as the hyperplane. The hyperplane is defined by a set of parameters known as the weights and bias. It also tries to find the hyperplane that maximises the margin between the two classes. The margin is the distance between the hyperplane and the closest data points of each class. The closer the data points are to the hyperplane, the more likely they are to be misclassified. Therefore, maximising the margin is a way to increase the robustness of the classifier. The σ and C values in SVM are hyperparameters that affect the decision boundary. The sigma parameter is used to control the width of the Gaussian kernel function, which is used in the radial basis function (RBF) kernel. The RBF kernel is a popular kernel function used in SVM for non-linear classification problems. The kernel function computes the similarity between two data points in the feature space. A smaller σ value leads to a wider kernel function, which makes the decision boundary smoother. On the other hand, a larger σ value leads to a narrower kernel function, which makes the decision boundary more complex. The C parameter is used to control the trade-off between the margin maximisation and the classification error. A larger C value leads to a narrower margin and a more complex decision boundary. This can result in overfitting, where the model fits the training data too closely and performs poorly on new, unseen data. On the other hand, a smaller C value leads to a wider margin and a simpler decision boundary. This can result in underfitting, where the model is too simple and cannot capture the complexity of the data. To support the results presented in Figures C.1–C.39, Figures C.65– C.49 are displayed, depicting the decision boundaries obtained when the SVM was trained on all experimental scenarios using the optimal values of σ and C. This demonstrates the capability of the Kernel SVM in accurately detecting age categories based on gait dynamics of each gender group. The decision boundaries in all the graphs are depicted for $\sigma = 0.1$ and C values of 0.1, 1, and 10.

7.7.5 Mean Square Error Rate and Standard Deviation Test

Figures 7.13–7.15 illustrate the mean square error (MSE) rate of the SVM model for female and male individuals across various age categories, as well as for combined age categories within each gender group. The average MSE for each group was calculated by averaging the accuracy obtained from each curve presented in Figures C.1–C.39. The results reveal that the average MSE for females is lowest for age categories s_{f8}, s_{f9}, and s_{f10}, while the SVM trained on male individuals' data produced the highest values for age categories s_{m5} and s_{m6}. Additionally, when the SVM classifier was tested on combined similar age categories between gender groups, high average MSE rates were recorded for s_{fm1} and s_{fm5}. The standard deviation plots presented on the right-hand side of Figures 7.13– 7.15 display the degree of variability among individual data points.

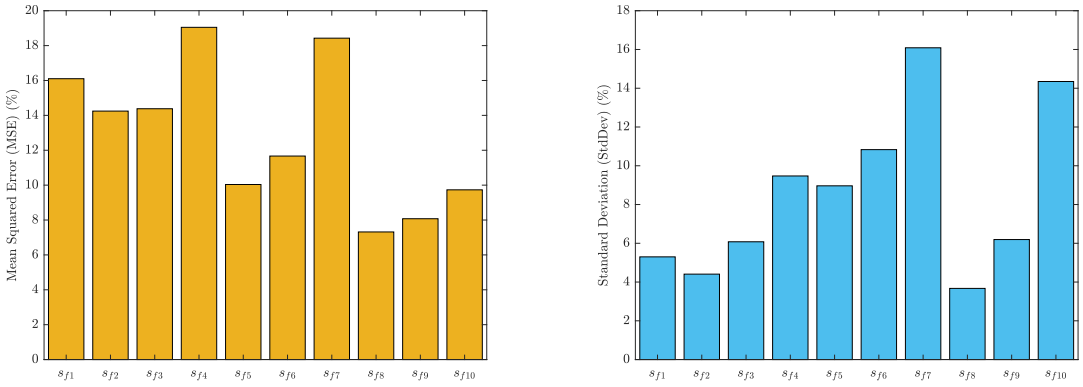


Figure 7.13: This figure shows the average mean square error (MSE)(%) (on the left) and the average standard deviation (%) (on the right) of the SVM classification performance for s_{f1}, s_{f2}, s_{f3}, s_{f4}, s_{f5}, s_{f6}, s_{f7}, s_{f8}, s_{f9} and s_{f10}.

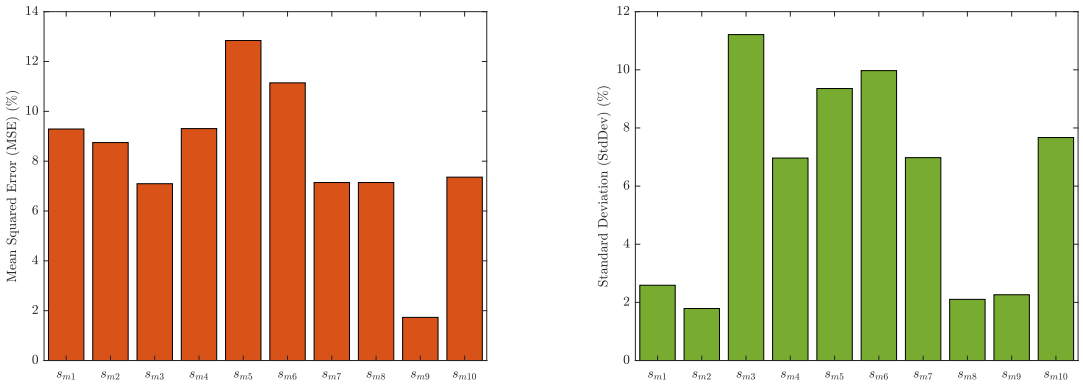


Figure 7.14: This figure shows the average mean square error (MSE)(%) (on the left) and the average standard deviation (%) (on the right) of the SVM classification performance for s_{m1}, s_{m2}, s_{m3}, s_{m4}, s_{m5}, s_{m6}, s_{m7}, s_{m8}, s_{m9} and s_{m10}.

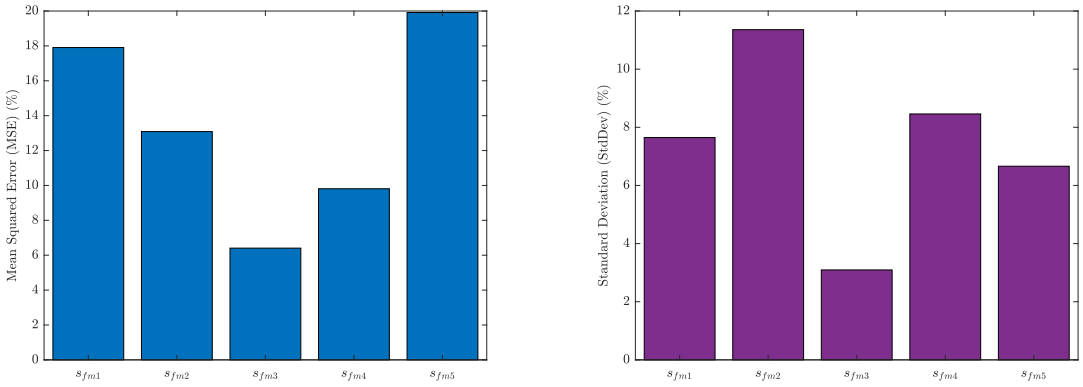


Figure 7.15: This figure shows the average mean square error (MSE)(%) (on the left) and the average standard deviation (%) (on the right) of the SVM classification performance for s_{fm1}, s_{fm2}, s_{fm3}, s_{fm4} and s_{fm5}.

7.8 Summary

In this chapter, the proposed SVM algorithm utilising a criticality analysis technique was applied to classify and distinguish between male and female individuals of different age groups based on their walking patterns. The results demonstrate that the SVM, when trained with the support of the CA method, can be an effective and reliable tool for accurately detecting gait patterns based on demographic factors such as age and gender.

The classification performance of the proposed kernel SVM in this study largely depends on the selection of the regularisation parameter C, which is discussed in Figures C.1–C.26 and Tables C.1– C.52. This parameter serves as a penalty for misclassification accuracy and must be carefully optimised to achieve the highest performance accuracy. Results indicate that the SVM performs best when σ is set to 0.1 or 1, as these values correspond to a faster decline in the Kernel Gaussian function. This function was found to effectively capture the non-linearities of different gait patterns, resulting in high SVM accuracy at these specific values. Moreover, Figures C.27– C.39 illustrate that the optimal value of C varies depending on the individual data pattern, and thus trial and error may be required to determine the appropriate value.

The criticality analysis (CA) approach was implemented in conjunction with the support vector machine (SVM) algorithm on the CARDIGAN dataset, as discussed in Chapter 5, resulting in an average accuracy of 78.2%. Similarly, when applied to the MoRES dataset presented in Chapter 6, the CA approach achieved an accuracy of 94%. In the current chapter, the CA approach was applied to the Museum dataset, resulting in an accuracy of 92%.

Overall, criticality analysis is a valuable tool in gait analysis, especially in detecting age-related changes in gait patterns. The approach's ability to identify associated factors and distinguish between gait patterns based on age makes it a powerful tool

in healthcare, rehabilitation, and sports medicine. Its ability to detect subtle changes that may not be apparent through traditional analysis methods makes it an essential tool in early disease diagnosis and management.

Chapter 8

Conclusion and Future Work

8.1 Conclusion

This section provides a synopsis of the primary contribution of the thesis, which addressed two key research inquiries initially introduced in Chapter 1.

In the initial stage of this research, an extensive examination was conducted to investigate the dynamical models that have been employed for the analysis of human gait. Out of the various models, only two dynamic models were identified, which have tremendous potential for providing profound insights into the analysis of human gait, particularly during the stance phase of the gait cycle. The first model is the mass-spring system, which is used to identify gait during running, while the second is the simple inverted pendulum, which is utilised to recognise gait during walking. The hybrid pendulum mass-spring system, which combines both models, was employed to analyse the gait cycle during the swing phase and account for the stiffness of the locomotor system's tissues. However, these physical harmonic systems solely focused on the structure of the musculoskeletal system, which comprises muscles, bones, and other connective tissues that facilitate body movement. Moreover, these models only considered the two stages of gait: walking and running, and neglected the abnormal

gait, which is associated with several unidentified disorders and pathological conditions. Therefore, it is imperative to comprehend the nonlinear complexity of human gait as a biological mathematical model, which not only reflects normal gait patterns but also provides insights into spatiotemporal patterns when they are disrupted or violated. To achieve this, the thesis has been developed to address the first research question.

Is it possible to identify the non-linear dynamics of human gait using the criticality analysis technique?

Chapter 3 presented an established bioenzymatic system that offers biological relevance in explaining the biological interactions of biological system functions through the concentrations of metabolic reactions. This system's strength lies in its ability to capture the dynamic interactions that affect biological mechanisms using a flexible set of Ordinary Differential Equations (ODEs). Additionally, the system's criticality analysis technique proved effective in detecting the biological dynamics of human gait by generating new data representations from the system's chaotic behavior. The improved detection technique resulted from the application of the rate control of chaos (RCC) principle, which regulates the rate of change of system variables that exhibit nonlinear and chaotic effects. A comprehensive evaluation of the RCC approach, as discussed in Chapter 3, demonstrated its effectiveness in regulating the time-varying chaotic response of the adopted Berry model [29]. By adjusting certain parameters, the model remained stable and in a state of harmony.

The utilisation of a novel criticality analysis technique for human gait analysis in conjunction with supervised machine learning approaches has effectively addressed a significant research gap in the medical field pertaining to gait therapy. This approach has yielded conclusive findings on the detection of the dynamic characteristics of human gait. Additionally, the study has effectively dealt with the issue of non-

linearity in gait patterns through the use of machine learning automation methods. Consequently, the second research question was comprehensively discussed.

How effective are the supervised machine learning models, specifically the Support Vector Machine (SVM) classifiers, that have been developed into practice for identifying and categorising patterns of human gait disturbance?

Chapter 4 of this thesis examined the effectiveness of the proposed support vector machine algorithm in classifying nonlinear patterns of data. The algorithm offers customisable design parameters that can adapt to various classification challenges. In Chapter 5, 6, and 7, the successful implementation of supervised machine learning models on three datasets, namely CARDIGAN, MoRES, and Museum, respectively, were extensively analysed using the SVM algorithm.

Additionally, we exclusively investigated the K-Nearest Neighbouring algorithm on the MoRES dataset, which showed adequate performance, but not as good as the SVM approach.

The Kernelised SVM technique was applied to analyse the CARDIGAN, MoRES, and Museum datasets. In the CARDIGAN dataset, the accuracy of the technique in detecting metabolic disorders associated with human gait was found to be 78.2% on average for distinguishing obese individuals from those in a healthy state. The classifier trained on the MoRES dataset showed a higher classification accuracy of approximately 94% for identifying gait abnormalities in individuals. The Museum dataset was used to classify different gait patterns based on age-gender categories and investigate the effect of aging on gait progression. The evaluation performance of the classification was relatively high at around 92.6% compared to the other datasets (CARDIGAN and MoRES) with an average performance of 86.1%. The regularisation parameter C and control width σ of the Gaussian kernel function were impor-

tant factors affecting the overall classification performance. The best results were obtained when sigma was set to 0.1 and 1 due to the minimal variance of data points around the average of neighbouring data samples.

The primary focus of the three dataset analyses was to establish the suitability of the confusion matrix as a measure of accuracy, particularly in evaluating the sensitivity and specificity of the model in categorising different categories. Sensitivity and specificity were calculated across a range of threshold values, and their relationship was inversely related. The Receiver Operating Characteristics (ROC) plot was utilised to graphically depict the results, including the area under the curve, which served as an effective evaluation of the overall accuracy of the model. The greater the area under the curve, the more accurate and robust the model is in detecting and classifying different categories.

8.2 Future Work

In biological processes, the human body expends significant amounts of energy to operate and select appropriate functions to accomplish various tasks. Therefore, there is a need to create a unique model that can comprehensively describe the integrated functions of the locomotion system. Future research should assess current dynamic systems and their potential for aiding individuals with gait disorders. As dynamic systems are increasingly utilised in evaluating proof-of-concepts for machine learning and artificial intelligence approaches, it is crucial to demonstrate the clinical validity and reliability of these systems for gait function. Moreover, the criticality analysis methodology could be an effective tool for analyzing the behavior of human gait function using physical nonlinear harmonic oscillation models, such as the van der Pol oscillator, Kuramoto oscillator, Rayleigh oscillator, Nos*é*-Hoover oscillator, Munmuangsaen oscillator, KBB oscillator, thermostated ergodic oscillator, MKT doubly-thermostated ergodic oscillator, Sinusoidally-forced parametric oscilla-

tor, Symmetric parametrically-coupled oscillators, and Asymmetric parametrically-coupled oscillators [218].

There are various research avenues to explore, including the possibility of further extending and optimizing the model presented in Chapter 3 to achieve controllable phase space trajectories. One potential approach involves modulating the bifurcation parameter r_{im}, which governs system stability, with external perturbent inputs using an exponential function, such that ${r_{im}}^* = r_{im}e^{-\frac{1}{2}{x_i}^2}$. This method eliminates the need for intricate adjustment of nonlinear system variables, resulting in stable, structured, and differentiable phase space trajectories.

Chapter 4 discusses machine learning classification algorithms that can be further explored to create an effective model with universal classification performance. Graph neural networks are a popular choice for analysing high-dimensional healthcare data and extracting important information from biological systems, as noted in [223]. The automation, identification, and classification tasks can benefit from this approach. Specifically, human gait can be represented as a time-varying signal that provides detailed information about the functional status of the biological system. Using deep learning techniques, the aim is to convert the spatiotemporal dynamics of gait into graph signals and establish a graph-based architecture of gait signal features that can be inputted into graph convolutional networks for identification and classification tasks, going beyond the use of graph neural networks.

Appendix A

The Performance Measure Metrics of CARDIGAN Dataset

This appendix displays the performance measures, as represented by the confusion matrix, of the proposed SVM on the CARDIGAN dataset, as discussed in Chapter 5.

Table A.1: SVM Classification Results of w_1 Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	0	14	14
FP	0	12	12
FN	14	0	0
TN	12	0	0
FPR	0	1	1
Precision	0	0.538	0.538
Recall	0	1	1
F1−Score	0	0.7	0.7
Specificity	1	0	0
AROC	0.511	0.511	0.511
Accuracy(%)	46.15	53.84	53.84

Table A.2: SVM Classification Results of w_1 Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	11	14	14
FP	1	5	5
FN	3	0	0
TN	11	7	7
FPR	0.083	0.416	0.416
Precision	0.916	0.736	0.736
Recall	0.785	1	1
F1−Score	0.846	0.848	0.848
Specificity	0.916	0.583	0.583
AROC	0.875	0.755	0.755
Accuracy(%)	84.61	80.76	80.76
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	14	14	14
FP	9	10	9
FN	0	0	0
TN	3	2	3
FPR	0.75	0.833	0.75
Precision	0.608	0.583	0.608
Recall	1	1	1
F1−Score	0.756	0.736	0.756
Specificity	0.25	0.166	0.25
AROC	0.708	0.636	0.648
Accuracy(%)	65.38	61.53	65.38

Table A.3: SVM Classification Results of w_2 Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	0	13	14
FP	0	13	13
FN	14	1	0
TN	15	2	2
FPR	0	0.866	0.866
Precision	0	0.5	0.518
Recall	0	0.928	1
F1−Score	0	0.65	0.682
Specificity	1	0.133	0.133
AROC	0.519	0.5	0.557
Accuracy(%)	51.72	51.72	55.17

Table A.4: SVM Classification Results of w_2 Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	5	13	13
FP	0	8	8
FN	9	1	1
TN	15	7	7
FPR	0	0.533	0.533
Precision	1	0.619	0.619
Recall	0.357	0.928	0.928
F1−Score	0.526	0.742	0.742
Specificity	1	0.466	0.466
AROC	0.714	0.742	0.742
Accuracy(%)	68.96	68.96	68.96
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	13	14	14
FP	10	10	9
FN	1	0	0
TN	5	5	6
FPR	0.667	0.667	0.6
Precision	0.565	0.583	0.608
Recall	0.928	1	1
F1−Score	0.702	0.736	0.756
Specificity	0.333	0.333	0.4
AROC	0.638	0.7	0.719
Accuracy(%)	62.06	65.51	68.96

Table A.5: SVM Classification Results of w_3 Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	0	0	12
FP	0	0	7
FN	15	15	3
TN	11	11	4
FPR	0	0	0.636
Precision	0	0	0.631
Recall	0	0	0.8
F1−Score	0	0	0.705
Specificity	1	1	0.363
AROC	0.690	0.690	0.660
Accuracy(%)	42.30	42.30	61.53

Table A.6: SVM Classification Results of w_3 Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	0	11	11
FP	0	3	3
FN	15	4	4
TN	11	8	8
FPR	0	0.272	0.272
Precision	0	0.785	0.785
Recall	0	0.733	0.733
F1−Score	0	0.758	0.758
Specificity	1	0.727	0.727
AROC	0.690	0.769	0.769
Accuracy(%)	42.30	73.07	73.07
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	10	14	15
FP	5	9	11
FN	5	1	0
TN	6	2	0
FPR	0.454	0.818	1
Precision	0.667	0.608	0.576
Recall	0.667	0.933	1
F1−Score	0.667	0.736	0.731
Specificity	0.545	0.181	0
AROC	0.696	0.703	0.690
Accuracy(%)	61.53	61.53	57.69

Table A.7: SVM Classification Results of w_4 Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	0	13	13
FP	0	10	12
FN	13	0	0
TN	13	3	1
FPR	0	0.769	0.923
Precision	0	0.565	0.52
Recall	0	1	1
F1−Score	0	0.722	0.684
Specificity	1	0.231	0.076
AROC	0.366	0.521	0.408
Accuracy(%)	50	61.53	53.84

Table A.8: SVM Classification Results of w_4 Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	5	13	13
FP	0	5	5
FN	8	0	0
TN	13	8	8
FPR	0	0.384	0.384
Precision	1	0.722	0.722
Recall	0.384	1	1
F1−Score	0.556	0.838	0.838
Specificity	1	0.615	0.615
AROC	0.562	0.804	0.804
Accuracy(%)	69.23	80.76	80.76
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	12	11	12
FP	6	6	5
FN	1	2	1
TN	7	7	8
FPR	0.461	0.461	0.384
Precision	0.667	0.647	0.705
Recall	0.923	0.846	0.923
F1−Score	0.774	0.733	0.8
Specificity	0.538	0.538	0.615
AROC	0.686	0.639	0.745
Accuracy(%)	73.06	69.23	76.92

Table A.9: SVM Classification Results of w_5 Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	13	13	13
FP	9	9	9
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.591	0.591	0.591
Recall	1	1	1
F1−Score	0.742	0.742	0.742
Specificity	0	0	0
AROC	0.384	0.384	0.384
Accuracy(%)	59.09	59.09	59.09

Table A.10: SVM Classification Results of w_5 Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	13	13	13
FP	9	7	2
FN	0	0	0
TN	0	2	7
FPR	1	0.777	0.222
Precision	0.591	0.65	0.867
Recall	1	1	1
F1−Score	0.742	0.787	0.928
Specificity	0	0.222	0.777
AROC	0.384	0.461	0.821
Accuracy(%)	59.09	68.18	90.90
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	13	11	12
FP	9	1	1
FN	0	2	1
TN	0	8	8
FPR	1	0.111	0.111
Precision	0.591	0.916	0.923
Recall	1	0.846	0.923
F1−Score	0.742	0.88	0.923
Specificity	0	0.888	0.888
AROC	0.384	0.837	0.880
Accuracy(%)	59.09	86.36	90.91

Table A.11: SVM Classification Results of w_6 Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	14	14	14
FP	10	10	10
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.583	0.583	0.583
Recall	1	1	1
F1−Score	0.736	0.736	0.736
Specificity	0	0	0
AROC	0.385	0.385	0.385
Accuracy(%)	58.33	58.33	58.33

Table A.12: SVM Classification Results of w_6 Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	14	14	13
FP	10	6	6
FN	0	0	1
TN	0	4	4
FPR	1	0.6	0.6
Precision	0.583	0.7	0.684
Recall	1	1	0.928
F1−Score	0.736	0.823	0.787
Specificity	0	0.4	0.4
AROC	0.385	0.678	0.642
Accuracy(%)	58.33	75	70.83
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	14	14	14
FP	10	5	4
FN	0	0	0
TN	0	5	6
FPR	1	0.5	0.4
Precision	0.583	0.736	0.778
Recall	1	1	1
F1−Score	0.736	0.848	0.875
Specificity	0	0.5	0.6
AROC	0.385	0.721	0.792
Accuracy(%)	58.33	79.16	83.33

Appendix B

The Performance Measure Metrics of MoRES Dataset

This appendix displays the performance measures, as represented by the confusion matrix, of the proposed SVM on the MoRES dataset, as discussed in Chapter 6.

Table B.1: SVM Classification Results of p_1 Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	320	320	320
FP	280	280	280
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.533	0.533	0.533
Recall	1	1	1
F1-Score	0.695	0.695	0.695
Specificity	0	0	0
AROC	0.519	0.519	0.519
Accuracy(%)	53.33	53.33	53.33

Table B.2: SVM Classification Results of p_1 Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	317	283	289
FP	34	7	5
FN	3	37	31
TN	246	273	275
FPR	0.121	0.025	0.017
Precision	0.903	0.975	0.982
Recall	0.990	0.884	0.903
F1-Score	0.944	0.927	0.941
Specificity	0.878	0.975	0.982
AROC	0.944	0.937	0.949
Accuracy(%)	93.83	92.67	94
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	316	320	320
FP	54	23	12
FN	4	0	0
TN	226	257	268
FPR	0.192	0.082	0.042
Precision	0.854	0.932	0.963
Recall	0.987	1	1
F1-Score	0.915	0.965	0.981
Specificity	0.807	0.917	0.957
AROC	0.902	0.962	0.981
Accuracy(%)	90.33	96.17	98

Table B.3: SVM Classification Results of p_2 Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	300	300	300
FP	300	300	300
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.5	0.5	0.5
Recall	1	1	1
F1-Score	0.667	0.667	0.667
Specificity	0	0	0
AROC	0.504	0.504	0.504
Accuracy(%)	50	50	50

Table B.4: SVM Classification Results of p_2 Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	249	239	244
FP	10	2	1
FN	51	61	56
TN	290	298	299
FPR	0.033	0.006	0.003
Precision	0.961	0.991	0.995
Recall	0.83	0.796	0.813
F1-Score	0.89	0.883	0.895
Specificity	0.966	0.993	0.996
AROC	0.899	0.901	0.909
Accuracy(%)	98.83	89.5	90.5
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	300	300	300
FP	97	74	42
FN	0	0	0
TN	203	226	258
FPR	0.323	0.246	0.14
Precision	0.755	0.802	0.877
Recall	1	1	1
F1-Score	0.86	0.89	0.934
Specificity	0.676	0.753	0.86
AROC	0.837	0.877	0.931
Accuracy(%)	83.83	87.67	93

Table B.5: SVM Classification Results of p_3 Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	331	331	331
FP	269	269	269
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.551	0.551	0.551
Recall	1	1	1
F1-Score	0.711	0.711	0.711
Specificity	0	0	0
AROC	0.504	0.504	0.504
Accuracy(%)	55.17	55.17	55.17

Table B.6: SVM Classification Results of p_3 Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	265	244	247
FP	51	12	11
FN	66	87	84
TN	218	257	258
FPR	0.189	0.044	0.04
Precision	0.838	0.953	0.957
Recall	0.8	0.737	0.746
F1-Score	0.81	0.955	0.959
Specificity	0.81	0.955	0.959
AROC	0.815	0.848	0.854
Accuracy(%)	80.5	83.5	84.17
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	331	331	331
FP	212	195	184
FN	0	0	0
TN	57	74	85
FPR	0.788	0.724	0.684
Precision	0.609	0.629	0.642
Recall	1	1	1
F1-Score	0.757	0.772	0.782
Specificity	0.211	0.275	0.315
AROC	0.611	0.648	0.673
Accuracy(%)	64.67	67.5	69.33

Table B.7: SVM Classification Results of p_4 Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	332	332	332
FP	268	268	268
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.553	0.553	0.553
Recall	1	1	1
F1-Score	0.712	0.712	0.712
Specificity	0	0	0
AROC	0.502	0.502	0.502
Accuracy(%)	55.33	55.33	55.33

Table B.8: SVM Classification Results of p_4 Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	330	311	313
FP	27	1	1
FN	2	21	19
TN	241	267	267
FPR	0.1	0.003	0.003
Precision	0.924	0.996	0.996
Recall	0.993	0.936	0.942
F1-Score	0.957	0.965	0.969
Specificity	0.899	0.996	0.996
AROC	0.944	0.971	0.977
Accuracy(%)	95.17	96.33	96.67
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	330	331	331
FP	70	104	135
FN	2	1	1
TN	198	164	133
FPR	0.261	0.388	0.503
Precision	0.825	0.76	0.71
Recall	0.993	0.996	0.996
F1-Score	0.901	0.863	0.829
Specificity	0.738	0.611	0.496
AROC	0.871	0.81	0.738
Accuracy(%)	88	82.5	77.33

Table B.9: SVM Classification Results of p_5 Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	298	298	298
FP	302	302	302
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.496	0.496	0.496
Recall	1	1	1
F1-Score	0.663	0.663	0.663
Specificity	0	0	0
AROC	0.495	0.495	0.495
Accuracy(%)	49.67	49.67	49.67

Table B.10: SVM Classification Results of p_5 Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	268	264	266
FP	9	1	0
FN	30	34	32
TN	293	301	302
FPR	0.029	0.003	0
Precision	0.967	0.996	1
Recall	0.899	0.885	0.892
F1-Score	0.932	0.937	0.943
Specificity	0.97	0.996	1
AROC	0.931	0.941	0.943
Accuracy(%)	93.5	94.17	94.67
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	298	298	298
FP	134	66	45
FN	0	0	0
TN	168	236	257
FPR	0.443	0.218	0.149
Precision	0.689	0.818	0.868
Recall	1	1	1
F1-Score	0.816	0.9	0.929
Specificity	0.556	0.781	0.85
AROC	0.769	0.88	0.92
Accuracy(%)	77.67	89	92.5

Table B.11: SVM Classification Results of p_{All} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	1619	1619	1619
FP	1381	1381	1381
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.539	0.539	0.539
Recall	1	1	1
F1-Score	0.701	0.701	0.701
Specificity	0	0	0
AROC	0.50	0.50	0.50
Accuracy(%)	53.96	53.96	53.96

Table B.12: SVM Classification Results of p_{All} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	1480	1435	1507
FP	62	16	32
FN	139	184	112
TN	1319	1365	1349
FPR	0.044	0.011	0.023
Precision	0.959	0.988	0.979
Recall	0.914	0.886	0.930
F1-Score	0.936	0.934	0.954
Specificity	0.955	0.988	0.976
AROC	0.939	0.945	0.957
Accuracy(%)	93.3	93.33	95.2
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	1618	1619	1619
FP	316	288	221
FN	1	0	0
TN	1065	1093	1160
FPR	0.228	0.208	0.16
Precision	0.836	0.848	0.879
Recall	0.999	1	1
F1-Score	0.91	0.918	0.936
Specificity	0.771	0.791	0.839
AROC	0.886	0.903	0.924
Accuracy(%)	89.43	90.4	92.63

Appendix C

The Performance Measure Metrics of Museum Dataset

Chapter 7 discusses the performance of the proposed SVM on the Museum dataset, and this appendix presents the performance measures in the form of a confusion matrix. Additionally, the appendix includes various figures related to the SVM performance, such as ROC curves, the area under the ROC curves, and decision boundaries.

Table C.1: SVM Classification Results of x_{f1} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	112	112	112
FP	82	82	82
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.577	0.577	0.577
Recall	1	1	1
F1-Score	0.732	0.732	0.732
Specificity	0	0	0
AROC	0.502	0.502	0.502
Accuracy(%)	57.73	57.73	57.73

Table C.2: SVM Classification Results of x_{f1} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	112	112	112
FP	29	29	27
FN	0	0	0
TN	53	53	55
FPR	0.353	0.353	0.329
Precision	0.794	0.794	0.805
Recall	1	1	1
F1-Score	0.885	0.885	0.892
Specificity	0.646	0.646	0.67
AROC	0.821	0.817	0.825
Accuracy(%)	85.05	85.05	86.08
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	112	112	112
FP	60	31	29
FN	0	0	0
TN	22	51	53
FPR	0.731	0.378	0.353
Precision	0.651	0.783	0.794
Recall	1	1	1
F1-Score	0.788	0.878	0.885
Specificity	0.268	0.621	0.646
AROC	0.639	0.819	0.835
Accuracy(%)	69.07	84.02	85.05

Table C.3: SVM Classification Results of x_{f2} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	343	343	343
FP	251	251	251
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.577	0.577	0.577
Recall	1	1	1
F1-Score	0.732	0.732	0.732
Specificity	0	0	0
AROC	0.49	0.49	0.49
Accuracy(%)	57.74	57.74	57.74

Table C.4: SVM Classification Results of x_{f2} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	338	343	342
FP	84	75	74
FN	5	0	1
TN	167	176	177
FPR	0.334	0.298	0.294
Precision	0.8	0.82	0.822
Recall	0.985	1	0.997
F1-Score	0.883	0.901	0.901
Specificity	0.665	0.701	0.705
AROC	0.829	0.855	0.858
Accuracy(%)	85.01	87.37	87.37
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	343	341	342
FP	160	101	108
FN	0	2	1
TN	91	150	143
FPR	0.637	0.402	0.43
Precision	0.681	0.771	0.76
Recall	1	0.994	0.997
F1-Score	0.81	0.868	0.862
Specificity	0.362	0.597	0.569
AROC	0.668	0.789	0.77
Accuracy(%)	73.06	82.65	81.64

Table C.5: SVM Classification Results of x_{f3} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	89	89	89
FP	62	62	62
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.589	0.589	0.589
Recall	1	1	1
F1-Score	0.741	0.741	0.741
Specificity	0	0	0
AROC	0.543	0.543	0.543
Accuracy(%)	58.94	58.94	58.94

Table C.6: SVM Classification Results of x_{f3} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	87	87	88
FP	24	22	21
FN	2	2	1
TN	38	40	41
FPR	0.387	0.354	0.338
Precision	0.783	0.798	0.807
Recall	0.977	0.977	0.988
F1-Score	0.87	0.878	0.888
Specificity	0.612	0.645	0.661
AROC	0.806	0.817	0.82
Accuracy(%)	82.78	84.1	85.43
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	89	89	89
FP	35	18	18
FN	0	0	0
TN	27	44	44
FPR	0.564	0.29	0.29
Precision	0.717	0.831	0.831
Recall	1	0.994	0.997
F1-Score	0.835	0.908	0.908
Specificity	0.435	0.709	0.709
AROC	0.726	0.882	0.88
Accuracy(%)	76.82	88.07	88.07

Table C.7: SVM Classification Results of x_{f4} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	13	13	13
FP	10	10	10
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.565	0.565	0.565
Recall	1	1	1
F1-Score	0.722	0.722	0.722
Specificity	0	0	0
AROC	0.515	0.515	0.515
Accuracy(%)	56.52	56.52	56.52

Table C.8: SVM Classification Results of x_{f4} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	11	12	12
FP	0	3	3
FN	2	1	1
TN	10	7	7
FPR	0	0.3	0.3
Precision	1	0.8	0.8
Recall	0.846	0.923	0.923
F1-Score	0.916	0.857	0.857
Specificity	1	0.7	0.7
AROC	0.946	0.8	0.8
Accuracy(%)	91.3	82.6	82.6
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	13	13	13
FP	4	2	2
FN	0	0	0
TN	6	8	8
FPR	0.4	0.2	0.2
Precision	0.764	0.866	0.866
Recall	1	1	1
F1-Score	0.866	0.928	0.928
Specificity	0.6	0.8	0.8
AROC	0.823	0.93	0.93
Accuracy(%)	82.6	91.3	91.3

Table C.9: SVM Classification Results of x_{f5} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	335	335	335
FP	241	259	259
FN	0	0	0
TN	18	0	0
FPR	0.93	1	1
Precision	0.581	0.563	0.563
Recall	1	1	1
F1-Score	0.735	0.721	0.721
Specificity	0.06	0	0
AROC	0.518	0.487	0.487
Accuracy(%)	59.42	56.39	56.39

Table C.10: SVM Classification Results of x_{f5} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	300	334	329
FP	18	49	44
FN	35	1	6
TN	241	210	215
FPR	0.06	0.189	0.169
Precision	0.943	0.872	0.882
Recall	0.895	0.997	0.982
F1-Score	0.918	0.93	0.929
Specificity	0.93	0.81	0.83
AROC	0.909	0.897	0.894
Accuracy(%)	91.07	91.58	91.58
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	334	320	331
FP	127	19	13
FN	1	15	4
TN	132	240	246
FPR	0.49	0.07	0.05
Precision	0.724	0.943	0.962
Recall	0.997	0.955	0.988
F1-Score	0.839	0.949	0.974
Specificity	0.509	0.926	0.949
AROC	0.739	0.941	0.97
Accuracy(%)	78.45	94.27	97.13

Table C.11: SVM Classification Results of x_{f6} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	79	79	79
FP	72	72	72
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.523	0.523	0.523
Recall	1	1	1
F1-Score	0.686	0.686	0.686
Specificity	0	0	0
AROC	0.543	0.543	0.543
Accuracy(%)	52.31	52.31	52.31

Table C.12: SVM Classification Results of x_{f6} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	78	77	76
FP	12	11	13
FN	1	2	3
TN	60	61	59
FPR	0.166	0.152	0.18
Precision	0.866	0.875	0.853
Recall	0.987	0.974	0.962
F1-Score	0.923	0.922	0.904
Specificity	0.833	0.847	0.819
AROC	0.923	0.905	0.882
Accuracy(%)	91.39	91.39	89.4
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	79	75	78
FP	37	10	0
FN	0	4	1
TN	35	62	72
FPR	0.513	0.138	0
Precision	0.681	0.882	1
Recall	1	0.949	0.987
F1-Score	0.81	0.914	0.993
Specificity	0.486	0.861	1
AROC	0.764	0.921	0.994
Accuracy(%)	75.49	90.72	99.33

Table C.13: SVM Classification Results of x_{f7} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	9	9	9
FP	14	14	14
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.391	0.391	0.391
Recall	1	1	1
F1-Score	0.562	0.562	0.562
Specificity	0	0	0
AROC	0.46	0.46	0.46
Accuracy(%)	39.13	39.13	39.13

Table C.14: SVM Classification Results of x_{f7} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	8	8	9
FP	3	1	2
FN	1	1	0
TN	11	13	12
FPR	0.214	0.071	0.142
Precision	0.727	0.888	0.818
Recall	0.888	0.888	1
F1-Score	0.8	0.888	0.9
Specificity	0.785	0.928	0.857
AROC	0.769	0.888	0.865
Accuracy(%)	82.6	91.3	91.3
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	9	7	9
FP	7	4	0
FN	0	2	0
TN	7	10	14
FPR	0.5	0.285	0
Precision	0.562	0.636	1
Recall	1	0.777	1
F1-Score	0.72	0.7	1
Specificity	0.5	0.714	1
AROC	0.769	0.73	1
Accuracy(%)	69.56	73.91	100

Table C.15: SVM Classification Results of x_{f8} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	81	81	81
FP	70	70	70
FN	0	0	0
TN	0	0	0
FPR	0.698	0.698	0.698
Precision	0.536	0.536	0.536
Recall	1	1	1
F1-Score	0.698	0.698	0.698
Specificity	0	0	0
AROC	0.548	0.548	0.548
Accuracy(%)	53.64	53.64	53.64

Table C.16: SVM Classification Results of x_{f8} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	80	81	81
FP	17	7	6
FN	1	0	0
TN	53	63	64
FPR	0.242	0.1	0.085
Precision	0.824	0.92	0.931
Recall	0.987	1	1
F1-Score	0.898	0.958	0.964
Specificity	0.757	0.9	0.914
AROC	0.907	0.971	0.969
Accuracy(%)	88.07	95.36	96.02
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	80	78	81
FP	26	6	7
FN	1	3	0
TN	44	64	63
FPR	0.371	0.085	0.1
Precision	0.754	0.928	0.92
Recall	0.987	0.962	1
F1-Score	0.855	0.945	0.958
Specificity	0.628	0.914	0.9
AROC	0.833	0.958	0.972
Accuracy(%)	82.11	94.03	95.36

Table C.17: SVM Classification Results of x_{f9} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	10	10	10
FP	13	13	13
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.434	0.434	0.434
Recall	1	1	1
F1-Score	0.606	0.606	0.606
Specificity	0	0	0
AROC	0.376	0.376	0.376
Accuracy(%)	43.47	43.47	43.47

Table C.18: SVM Classification Results of x_{f9} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	10	9	10
FP	4	2	0
FN	0	1	0
TN	9	11	13
FPR	0.307	0.153	0
Precision	0.714	0.818	1
Recall	1	0.9	1
F1-Score	0.833	0.857	1
Specificity	0.692	0.846	1
AROC	0.7	0.83	1
Accuracy(%)	82.6	86.95	100
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	10	10	10
FP	3	1	1
FN	0	0	0
TN	10	12	12
FPR	0.23	0.076	0.076
Precision	0.769	0.909	0.909
Recall	1	1	1
F1-Score	0.869	0.952	0.952
Specificity	0.769	0.923	0.923
AROC	0.861	0.923	0.923
Accuracy(%)	86.95	95.65	95.65

Table C.19: SVM Classification Results of x_{f10} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	6	12	12
FP	0	11	11
FN	6	0	0
TN	11	0	0
FPR	0	1	1
Precision	1	0.521	0.521
Recall	0.5	1	1
F1-Score	0.666	0.685	0.685
Specificity	1	0	0
AROC	0.696	0.439	0.439
Accuracy(%)	73.91	52.17	52.17

Table C.20: SVM Classification Results of x_{f10} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	10	12	12
FP	0	1	0
FN	2	0	0
TN	11	10	11
FPR	0	0.09	0
Precision	1	0.923	1
Recall	0.833	1	1
F1-Score	0.909	0.96	1
Specificity	1	0.909	1
AROC	0.931	0.909	1
Accuracy(%)	91.3	95.65	100
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	12	10	12
FP	4	1	1
FN	0	2	0
TN	7	10	10
FPR	0.363	0.09	0.09
Precision	0.75	0.909	0.923
Recall	1	0.833	1
F1-Score	0.857	0.869	0.96
Specificity	0.636	0.909	0.909
AROC	0.833	0.856	0.909
Accuracy(%)	82.6	86.95	95.65

Table C.21: SVM Classification Results of x_{m1} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	80	80	80
FP	63	63	63
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.559	0.559	0.559
Recall	1	1	1
F1-Score	0.717	0.717	0.717
Specificity	0	0	0
AROC	0.467	0.467	0.467
Accuracy(%)	55.94	55.94	55.94

Table C.22: SVM Classification Results of x_{m1} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	79	78	76
FP	12	9	10
FN	1	2	4
TN	51	54	53
FPR	0.19	0.142	0.158
Precision	0.868	0.896	0.883
Recall	0.987	0.975	0.95
F1-Score	0.923	0.934	0.915
Specificity	0.809	0.857	0.841
AROC	0.894	0.892	0.863
Accuracy(%)	90.9	92.3	90.2
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	80	80	80
FP	37	16	15
FN	0	0	0
TN	26	47	48
FPR	0.587	0.253	0.238
Precision	0.683	0.833	0.842
Recall	1	1	1
F1-Score	0.812	0.909	0.914
Specificity	0.412	0.746	0.761
AROC	0.668	0.861	0.854
Accuracy(%)	74.12	88.81	89.51

Table C.23: SVM Classification Results of x_{m2} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	254	254	254
FP	224	224	224
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.531	0.531	0.531
Recall	1	1	1
F1-Score	0.693	0.693	0.693
Specificity	0	0	0
AROC	0.487	0.487	0.487
Accuracy(%)	53.13	53.13	53.13

Table C.24: SVM Classification Results of x_{m2} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	250	249	246
FP	55	30	29
FN	4	5	8
TN	169	194	195
FPR	0.245	0.133	0.129
Precision	0.819	0.892	0.894
Recall	0.984	0.98	0.968
F1-Score	0.894	0.934	0.93
Specificity	0.754	0.866	0.87
AROC	0.864	0.916	0.91
Accuracy(%)	87.65	92.67	92.25
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	254	252	254
FP	148	38	42
FN	0	2	0
TN	76	186	182
FPR	0.66	0.169	0.187
Precision	0.631	0.868	0.858
Recall	1	0.992	1
F1-Score	0.774	0.926	0.923
Specificity	0.339	0.83	0.812
AROC	0.66	0.908	0.895
Accuracy(%)	69.03	91.63	91.21

Table C.25: SVM Classification Results of x_{m3} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	55	55	55
FP	45	45	45
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.55	0.55	0.55
Recall	1	1	1
F1-Score	0.709	0.709	0.709
Specificity	0	0	0
AROC	0.543	0.543	0.543
Accuracy(%)	55	55	55

Table C.26: SVM Classification Results of x_{m3} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	52	55	55
FP	0	4	4
FN	3	0	0
TN	45	41	41
FPR	0	0.088	0.088
Precision	1	0.932	0.932
Recall	0.945	1	1
F1-Score	0.971	0.964	0.964
Specificity	1	0.911	0.911
AROC	0.981	0.966	0.966
Accuracy(%)	97	96	96
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	55	55	55
FP	27	3	3
FN	0	0	0
TN	18	42	42
FPR	0.6	0.066	0.066
Precision	0.67	0.948	0.948
Recall	1	1	1
F1-Score	0.802	0.973	0.973
Specificity	0.4	0.933	0.933
AROC	0.73	0.97	0.973
Accuracy(%)	73	97	97

Table C.27: SVM Classification Results of x_{m4} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	14	14	14
FP	8	8	8
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.636	0.636	0.636
Recall	1	1	1
F1-Score	0.777	0.777	0.777
Specificity	0	0	0
AROC	0.526	0.526	0.526
Accuracy(%)	63.63	63.63	63.63

Table C.28: SVM Classification Results of x_{m4} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	12	14	14
FP	0	2	2
FN	2	0	0
TN	8	6	6
FPR	0	0.25	0.25
Precision	1	0.875	0.875
Recall	0.857	1	1
F1-Score	0.923	0.933	0.933
Specificity	1	0.75	0.75
AROC	0.892	0.821	0.821
Accuracy(%)	90.9	90.9	90.9
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	14	13	13
FP	3	0	0
FN	0	1	1
TN	5	8	8
FPR	0.375	0	0
Precision	0.823	1	1
Recall	1	0.928	0.928
F1-Score	0.802	0.973	0.973
Specificity	0.625	1	1
AROC	0.821	0.955	0.955
Accuracy(%)	86.36	95.45	95.45

Table C.29: SVM Classification Results of x_{m5} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	250	250	250
FP	173	228	228
FN	0	0	0
TN	55	0	0
FPR	0.758	1	1
Precision	0.591	0.523	0.523
Recall	1	1	1
F1-Score	0.742	0.686	0.686
Specificity	0.241	0	0
AROC	0.61	0.486	0.486
Accuracy(%)	63.8	52.3	52.3

Table C.30: SVM Classification Results of x_{m5} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	218	246	246
FP	11	37	45
FN	32	4	4
TN	217	191	183
FPR	0.048	0.162	0.197
Precision	0.951	0.869	0.845
Recall	0.872	0.984	0.984
F1-Score	0.91	0.923	0.909
Specificity	0.951	0.837	0.802
AROC	0.902	0.899	0.88
Accuracy(%)	91	91.42	89.74
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	250	249	250
FP	134	47	43
FN	0	1	0
TN	94	181	185
FPR	0.587	0.206	0.188
Precision	0.651	0.841	0.853
Recall	1	0.996	1
F1-Score	0.788	0.912	0.920
Specificity	0.412	0.793	0.811
AROC	0.697	0.887	0.899
Accuracy(%)	71.96	89.95	91

Table C.31: SVM Classification Results of x_{m6} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	49	51	51
FP	1	49	49
FN	2	0	0
TN	48	0	0
FPR	0.02	1	1
Precision	0.98	0.51	0.51
Recall	0.96	1	1
F1-Score	0.97	0.675	0.675
Specificity	0.979	0	0
AROC	0.965	0.533	0.533
Accuracy(%)	97	51	51

Table C.32: SVM Classification Results of x_{m6} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	46	50	51
FP	2	9	10
FN	5	1	0
TN	47	40	39
FPR	0.04	0.183	0.204
Precision	0.958	0.847	0.836
Recall	0.901	0.980	1
F1-Score	0.929	0.909	0.91
Specificity	0.959	0.816	0.795
AROC	0.932	0.899	0.897
Accuracy(%)	93	90	90
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	51	50	51
FP	31	11	11
FN	0	1	0
TN	18	38	38
FPR	0.632	0.224	0.224
Precision	0.621	0.819	0.822
Recall	1	0.98	1
F1-Score	0.766	0.892	0.902
Specificity	0.367	0.775	0.775
AROC	0.707	0.879	0.894
Accuracy(%)	69	88	89

Table C.33: SVM Classification Results of x_{m7} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	0	14	14
FP	0	8	8
FN	14	0	0
TN	8	0	0
FPR	0	1	1
Precision	0	0.636	0.636
Recall	0	1	1
F1-Score	0	0.777	0.777
Specificity	1	0	0
AROC	0.526	0.526	0.526
Accuracy(%)	36.36	63.63	63.63

Table C.34: SVM Classification Results of x_{m7} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	12	14	14
FP	0	1	1
FN	2	0	0
TN	8	7	7
FPR	0	0.125	0.125
Precision	1	0.933	0.933
Recall	0.857	1	1
F1-Score	0.923	0.965	0.965
Specificity	1	0.875	0.875
AROC	0.919	0.892	0.892
Accuracy(%)	90.9	95.45	95.45
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	14	14	14
FP	3	0	1
FN	0	0	0
TN	5	8	7
FPR	0.375	0	0.125
Precision	0.823	1	0.933
Recall	1	1	1
F1-Score	0.903	1	0.965
Specificity	0.625	1	0.875
AROC	0.821	1	0.919
Accuracy(%)	86.36	100	95.45

Table C.35: SVM Classification Results of x_{m8} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	51	51	51
FP	49	49	49
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.51	0.51	0.51
Recall	1	1	1
F1-Score	0.675	0.675	0.675
Specificity	0	0	0
AROC	0.521	0.521	0.521
Accuracy(%)	51	51	51

Table C.36: SVM Classification Results of x_{m8} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	51	50	50
FP	10	2	5
FN	0	1	1
TN	39	47	44
FPR	0.204	0.04	0.102
Precision	0.836	0.961	0.909
Recall	1	0.98	0.98
F1-Score	0.91	0.97	0.943
Specificity	0.795	0.959	0.897
AROC	0.895	0.979	0.94
Accuracy(%)	90	97	94
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	51	51	51
FP	21	7	15
FN	0	0	0
TN	28	42	34
FPR	0.428	0.142	0.306
Precision	0.708	0.879	0.772
Recall	1	1	1
F1-Score	0.829	0.935	0.871
Specificity	0.571	0.857	0.693
AROC	0.796	0.921	0.857
Accuracy(%)	79	93	85

Table C.37: SVM Classification Results of x_{m9} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	14	14	14
FP	8	8	8
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.636	0.636	0.636
Recall	1	1	1
F1-Score	0.777	0.777	0.777
Specificity	0	0	0
AROC	0.526	0.526	0.526
Accuracy(%)	63.63	63.63	63.63

Table C.38: SVM Classification Results of x_{m9} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	14	14	14
FP	1	0	0
FN	0	0	0
TN	7	8	8
FPR	0.125	0	0
Precision	0.933	1	1
Recall	1	1	1
F1-Score	0.965	1	1
Specificity	0.875	1	1
AROC	0.892	1	1
Accuracy(%)	95.45	100	100
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	14	13	14
FP	2	0	1
FN	0	1	0
TN	6	8	7
FPR	0.25	0	0.125
Precision	0.875	1	0.933
Recall	1	0.928	1
F1-Score	0.933	0.962	0.965
Specificity	0.75	1	0.875
AROC	0.901	0.955	0.919
Accuracy(%)	90.9	95.45	95.45

Table C.39: SVM Classification Results of x_{m10} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	11	14	14
FP	0	8	8
FN	3	0	0
TN	8	0	0
FPR	0	1	1
Precision	1	0.636	0.636
Recall	0.785	1	1
F1-Score	0.88	0.777	0.777
Specificity	1	0	0
AROC	0.892	0.526	0.526
Accuracy(%)	86.36	63.63	63.63

Table C.40: SVM Classification Results of x_{m10} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	12	14	14
FP	0	2	2
FN	2	0	0
TN	8	6	6
FPR	0	0.25	0.25
Precision	1	0.875	0.875
Recall	0.857	1	1
F1-Score	0.923	0.933	0.933
Specificity	1	0.75	0.75
AROC	0.892	0.821	0.821
Accuracy(%)	90.9	90.9	90.9
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	14	13	14
FP	3	0	0
FN	0	1	0
TN	5	8	8
FPR	0.375	0	0
Precision	0.823	1	1
Recall	1	0.928	1
F1-Score	0.903	0.962	1
Specificity	0.625	1	1
AROC	0.83	0.937	1
Accuracy(%)	86.36	95.45	100

Table C.41: SVM Classification Results of x_{fm1} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	97	97	97
FP	74	74	74
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.567	0.567	0.567
Recall	1	1	1
F1-Score	0.723	0.723	0.723
Specificity	0	0	0
AROC	0.486	0.486	0.486
Accuracy(%)	56.72	56.72	56.72

Table C.42: SVM Classification Results of x_{fm1} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	96	96	96
FP	35	27	25
FN	1	1	1
TN	39	47	49
FPR	0.472	0.364	0.337
Precision	0.732	0.78	0.79
Recall	0.989	0.989	0.989
F1-Score	0.842	0.872	0.88
Specificity	0.527	0.635	0.662
AROC	0.768	0.804	0.816
Accuracy(%)	78.94	83.62	84.79
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	97	97	97
FP	46	23	16
FN	0	0	0
TN	28	51	58
FPR	0.621	0.31	0.216
Precision	0.678	0.808	0.858
Recall	1	1	1
F1-Score	0.808	0.894	0.923
Specificity	0.378	0.689	0.783
AROC	0.691	0.859	0.922
Accuracy(%)	73.09	86.54	90.64

Table C.43: SVM Classification Results of x_{fm2} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	77	77	77
FP	66	66	66
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.538	0.538	0.538
Recall	1	1	1
F1-Score	0.7	0.7	0.7
Specificity	0	0	0
AROC	0.46	0.46	0.46
Accuracy(%)	53.84	53.84	53.84

Table C.44: SVM Classification Results of x_{fm2} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	74	74	77
FP	15	12	13
FN	3	3	0
TN	51	54	53
FPR	0.227	0.181	0.196
Precision	0.831	0.86	0.855
Recall	0.961	0.961	1
F1-Score	0.891	0.907	0.922
Specificity	0.772	0.818	0.803
AROC	0.856	0.871	0.884
Accuracy(%)	87.41	89.51	90.9
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	77	73	72
FP	36	9	0
FN	0	4	5
TN	30	57	66
FPR	0.545	0.136	0
Precision	0.681	0.89	1
Recall	1	0.948	0.935
F1-Score	0.81	0.918	0.966
Specificity	0.454	0.863	1
AROC	0.691	0.892	0.95
Accuracy(%)	74.825	90.9	96.5

Table C.45: SVM Classification Results of x_{fm3} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	266	266	266
FP	212	212	212
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.556	0.556	0.556
Recall	1	1	1
F1-Score	0.715	0.715	0.715
Specificity	0	0	0
AROC	0.488	0.488	0.488
Accuracy(%)	55.64	55.64	55.64

Table C.46: SVM Classification Results of x_{fm3} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	263	264	264
FP	47	18	17
FN	3	2	2
TN	165	194	195
FPR	0.221	0.084	0.08
Precision	0.848	0.936	0.939
Recall	0.988	0.992	0.992
F1-Score	0.913	0.963	0.965
Specificity	0.778	0.915	0.919
AROC	0.882	0.947	0.951
Accuracy(%)	89.53	95.81	96.02
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	265	260	265
FP	83	23	19
FN	1	6	1
TN	129	189	193
FPR	0.391	0.108	0.089
Precision	0.761	0.918	0.933
Recall	0.996	0.977	0.996
F1-Score	0.863	0.947	0.963
Specificity	0.608	0.891	0.91
AROC	0.798	0.928	0.94
Accuracy(%)	82.42	93.93	95.81

Table C.47: SVM Classification Results of x_{fm4} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	57	57	57
FP	43	43	43
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.57	0.57	0.57
Recall	1	1	1
F1-Score	0.726	0.726	0.726
Specificity	0	0	0
AROC	0.549	0.549	0.549
Accuracy(%)	57	57	57

Table C.48: SVM Classification Results of x_{fm4} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	55	55	56
FP	5	4	4
FN	2	2	1
TN	38	39	39
FPR	0.116	0.093	0.093
Precision	0.916	0.932	0.933
Recall	0.964	0.964	0.982
F1-Score	0.94	0.948	0.957
Specificity	0.883	0.906	0.906
AROC	0.937	0.94	0.946
Accuracy(%)	93	94	95
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	57	53	54
FP	20	1	0
FN	0	4	3
TN	23	42	43
FPR	0.465	0.023	0
Precision	0.74	0.918	1
Recall	1	0.929	0.947
F1-Score	0.85	0.954	0.972
Specificity	0.534	0.976	1
AROC	0.79	0.962	0.974
Accuracy(%)	80	95	97

Table C.49: SVM Classification Results of x_{fm5} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	14	14	14
FP	8	8	8
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.636	0.636	0.636
Recall	1	1	1
F1-Score	0.777	0.777	0.777
Specificity	0	0	0
AROC	0.526	0.526	0.526
Accuracy(%)	63.63	63.63	63.63

Table C.50: SVM Classification Results of x_{fm5} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	13	13	13
FP	6	3	3
FN	1	1	1
TN	2	5	5
FPR	0.75	0.375	0.375
Precision	0.684	0.812	0.812
Recall	0.928	0.928	0.928
F1-Score	0.787	0.866	0.866
Specificity	0.25	0.625	0.625
AROC	0.669	0.767	0.767
Accuracy(%)	68.18	81.81	81.81
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	14	14	14
FP	5	3	3
FN	0	0	0
TN	3	5	5
FPR	0.625	0.375	0.375
Precision	0.736	0.823	0.823
Recall	1	1	1
F1-Score	0.848	0.903	0.903
Specificity	0.375	0.625	0.625
AROC	0.723	0.83	0.794
Accuracy(%)	77.27	86.36	86.36

Table C.51: SVM Classification Results of x_{All} Walk Patterns at $\sigma = 10$.

Performance	$\boldsymbol{\sigma = 10}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	514	514	514
FP	400	400	400
FN	0	0	0
TN	0	0	0
FPR	1	1	1
Precision	0.562	0.562	0.562
Recall	1	1	1
F1-Score	0.719	0.719	0.719
Specificity	0	0	0
AROC	0.5	0.5	0.5
Accuracy(%)	56.23	56.23	56.23

Table C.52: SVM Classification Results of x_{All} Walk Patterns at $\sigma = 0.1$ and at $\sigma = 1$.

Performance	$\boldsymbol{\sigma = 0.1}$		
	$\boldsymbol{C = 0.1}$	$\boldsymbol{C = 1}$	$\boldsymbol{C = 10}$
TP	510	507	508
FP	75	11	19
FN	4	7	6
TN	325	389	381
FPR	0.187	0.027	0.047
Precision	0.871	0.978	0.963
Recall	0.992	0.986	0.988
F1-Score	0.928	0.982	0.975
Specificity	0.812	0.972	0.952
AROC	0.905	0.976	0.968
Accuracy(%)	91.35	98.03	97.26
Performance	$\boldsymbol{\sigma = 1}$		
	C=0.1	**C=1**	**C=10**
TP	491	492	505
FP	57	24	9
FN	23	22	9
TN	343	376	391
FPR	0.142	0.06	0.022
Precision	0.895	0.953	0.982
Recall	0.955	0.953	0.982
F1-Score	0.924	0.955	0.982
Specificity	0.857	0.94	0.977
AROC	0.911	0.949	0.978
Accuracy(%)	91.24	94.96	98.03

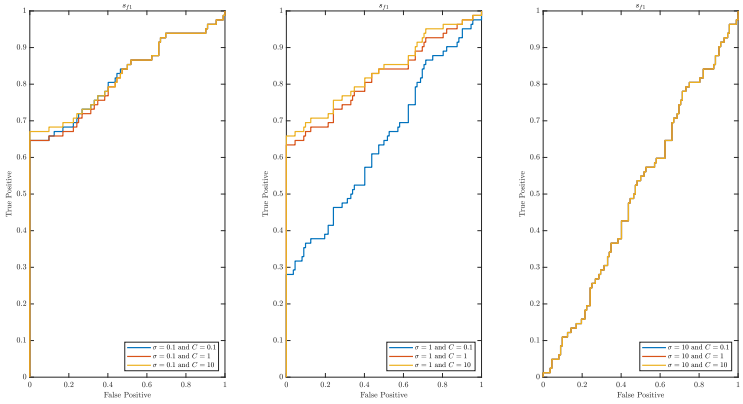


Figure C.1: ROC (Receiver Operating Characteristic) curves of s_{f1} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

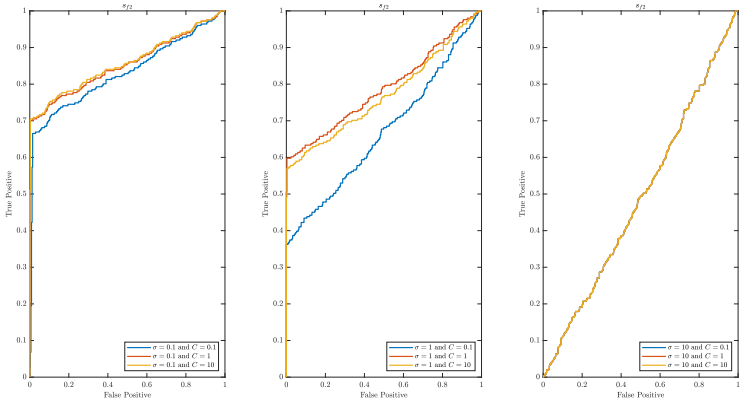


Figure C.2: ROC (Receiver Operating Characteristic) curves of s_{f2} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

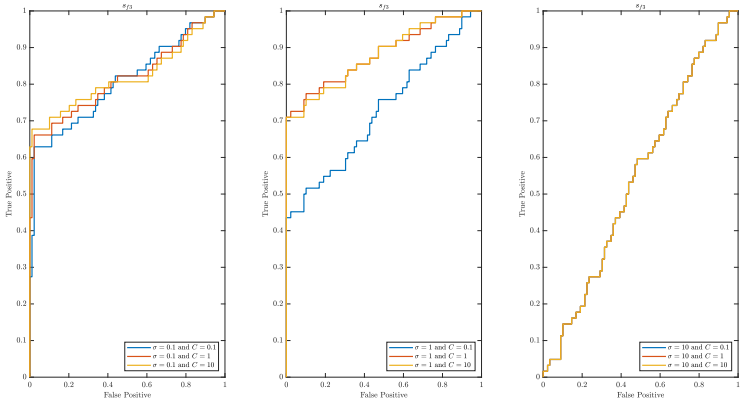


Figure C.3: ROC (Receiver Operating Characteristic) curves of s_{f3} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

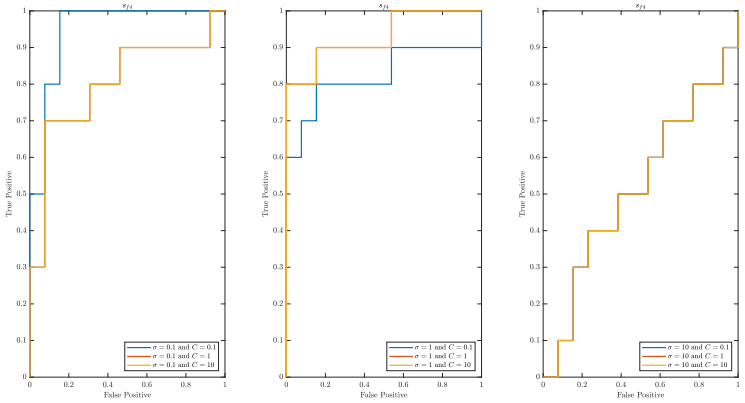


Figure C.4: ROC (Receiver Operating Characteristic) curves of s_{f4} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

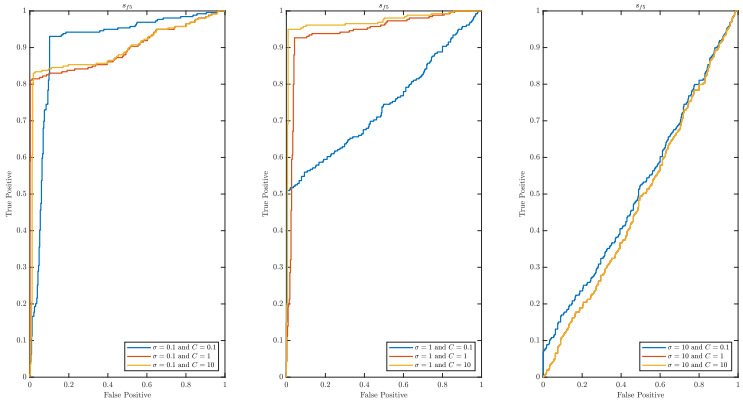


Figure C.5: ROC (Receiver Operating Characteristic) curves of s_{f5} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

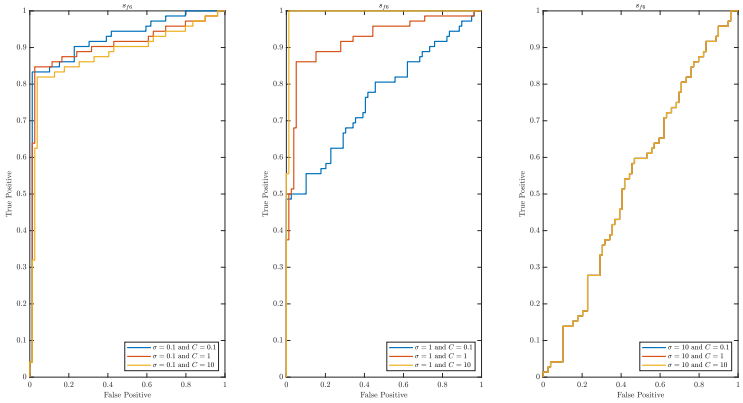


Figure C.6: ROC (Receiver Operating Characteristic) curves of s_{f6} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

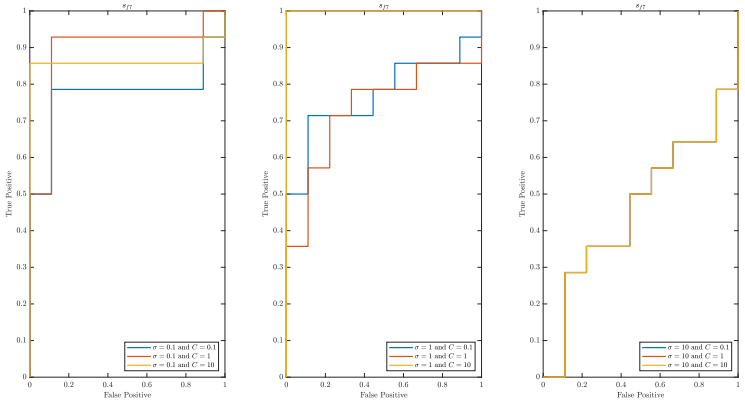


Figure C.7: ROC (Receiver Operating Characteristic) curves of s_{f7} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

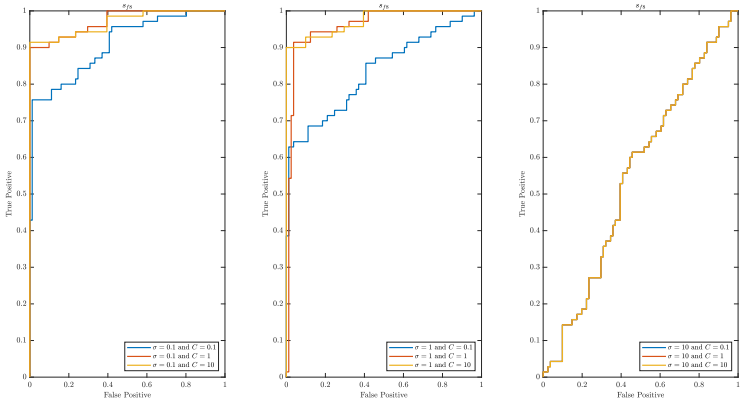


Figure C.8: ROC (Receiver Operating Characteristic) curves of s_{f8} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

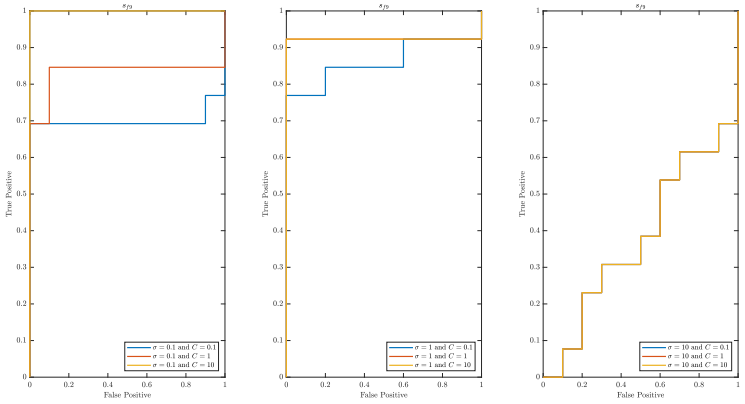


Figure C.9: ROC (Receiver Operating Characteristic) curves of s_{f9} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

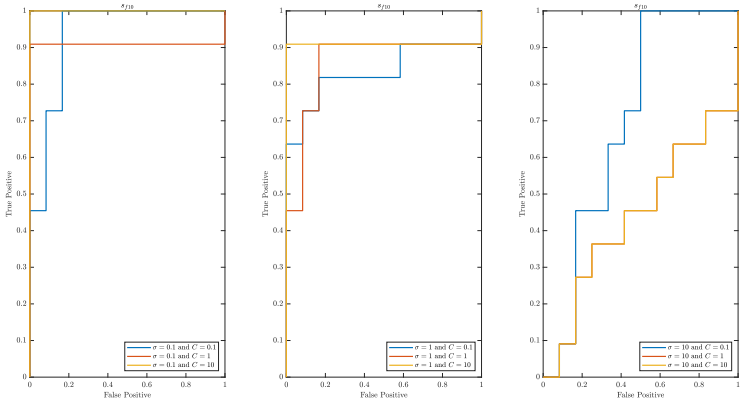


Figure C.10: ROC (Receiver Operating Characteristic) curves of s_{f10} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

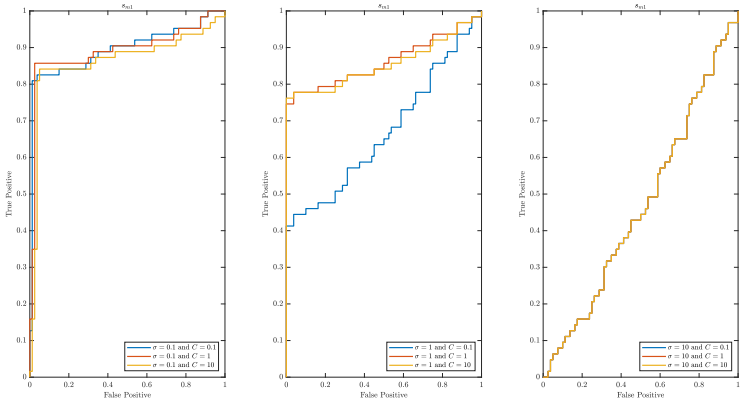


Figure C.11: ROC (Receiver Operating Characteristic) curves of s_{m1} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

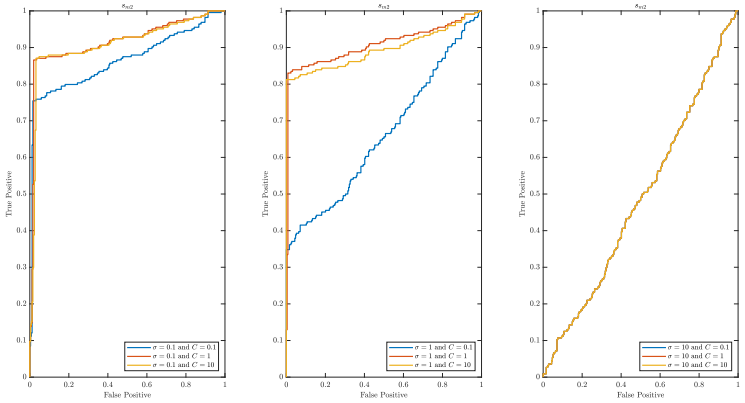


Figure C.12: ROC (Receiver Operating Characteristic) curves of s_{m2} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

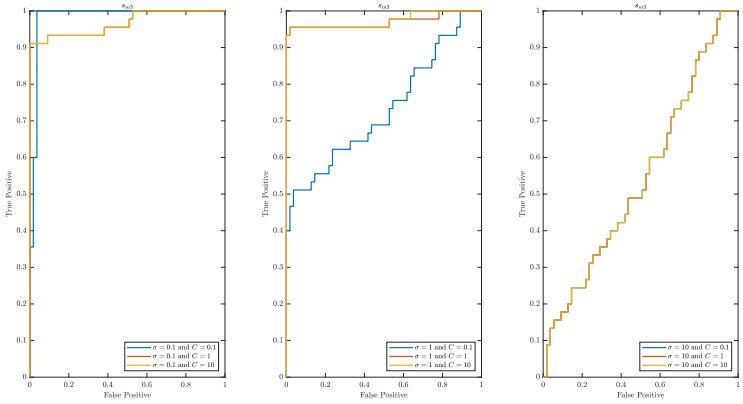


Figure C.13: ROC (Receiver Operating Characteristic) curves of s_{m3} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

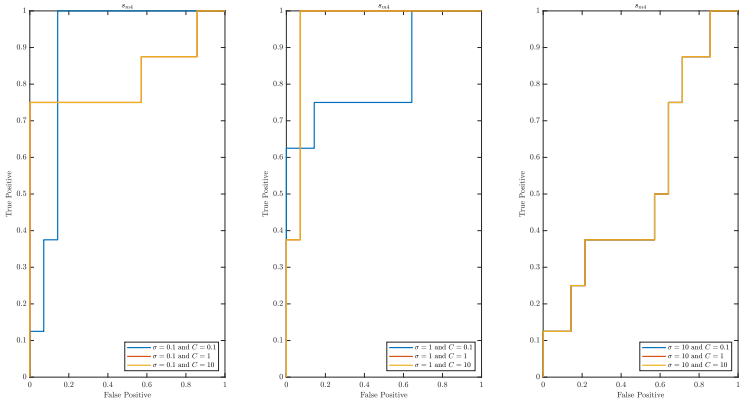


Figure C.14: ROC (Receiver Operating Characteristic) curves of s_{m4} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

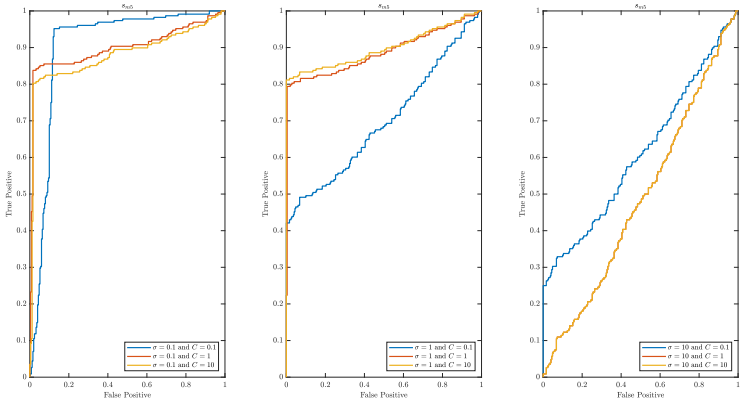


Figure C.15: ROC (Receiver Operating Characteristic) curves of s_{m5} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

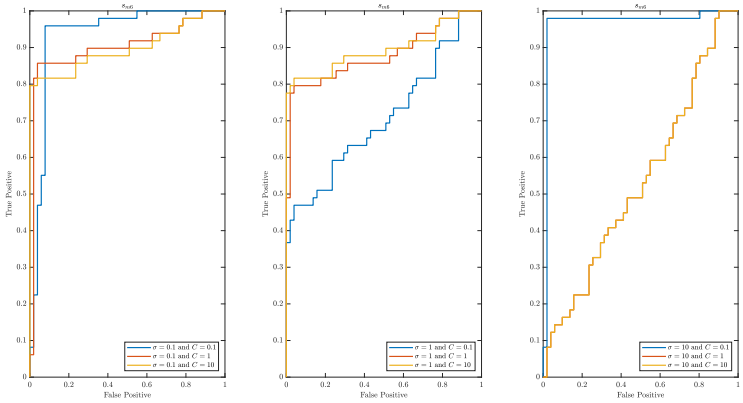


Figure C.16: ROC (Receiver Operating Characteristic) curves of s_{m6} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

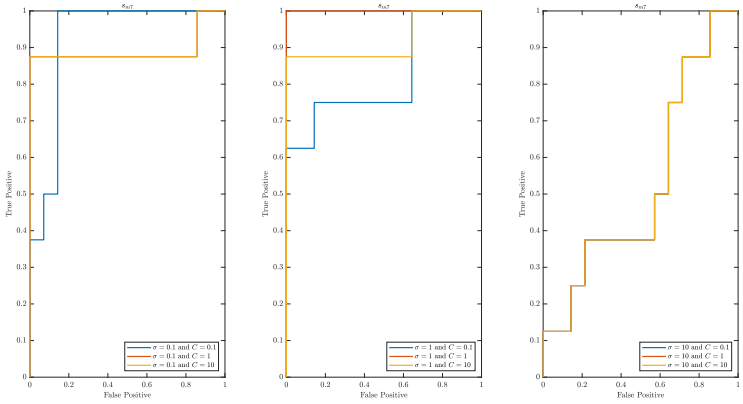


Figure C.17: ROC (Receiver Operating Characteristic) curves of s_{m7} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

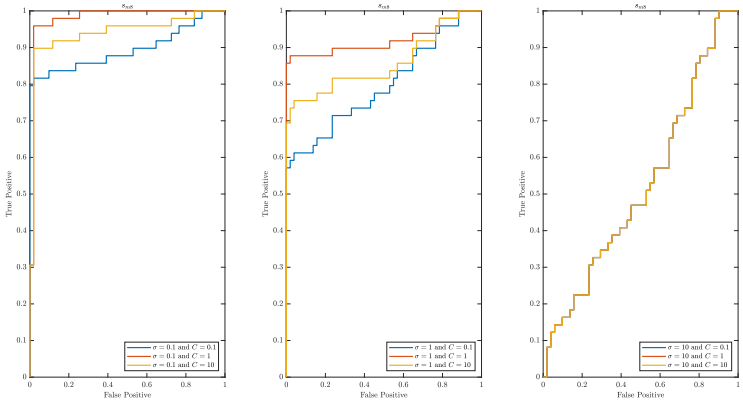


Figure C.18: ROC (Receiver Operating Characteristic) curves of s_{m8} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

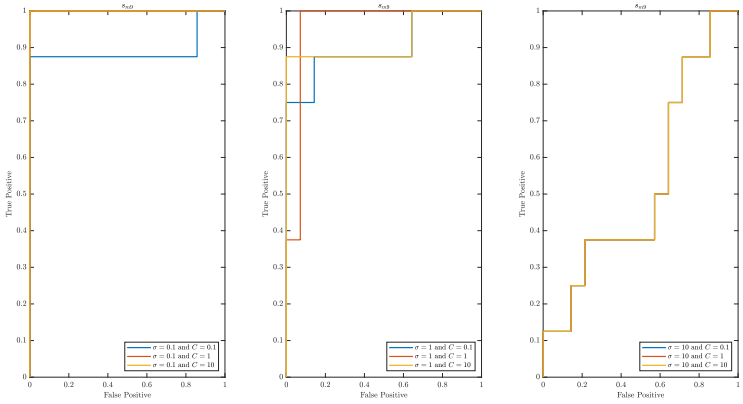


Figure C.19: ROC (Receiver Operating Characteristic) curves of s_{m9} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

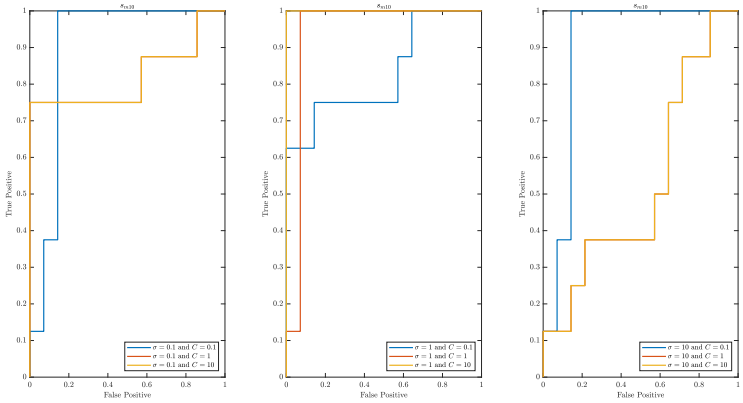


Figure C.20: ROC (Receiver Operating Characteristic) curves of s_{m10} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

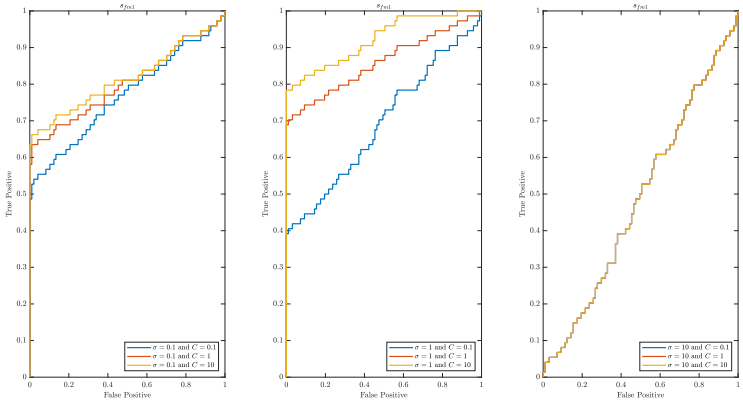


Figure C.21: ROC (Receiver Operating Characteristic) curves of s_{fm1} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

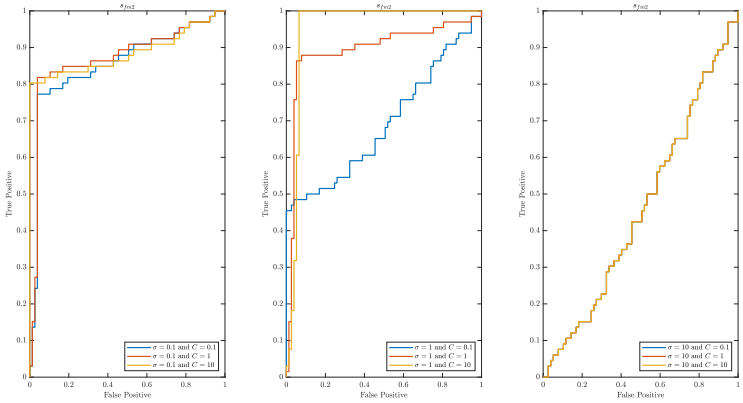


Figure C.22: ROC (Receiver Operating Characteristic) curves of s_{fm2} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

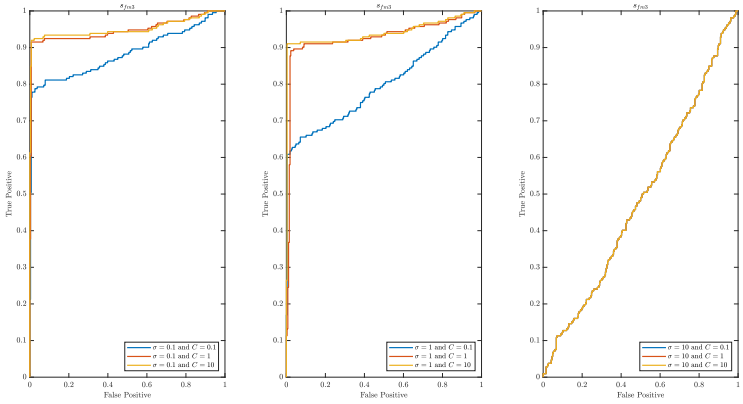


Figure C.23: ROC (Receiver Operating Characteristic) curves of s_{fm3} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

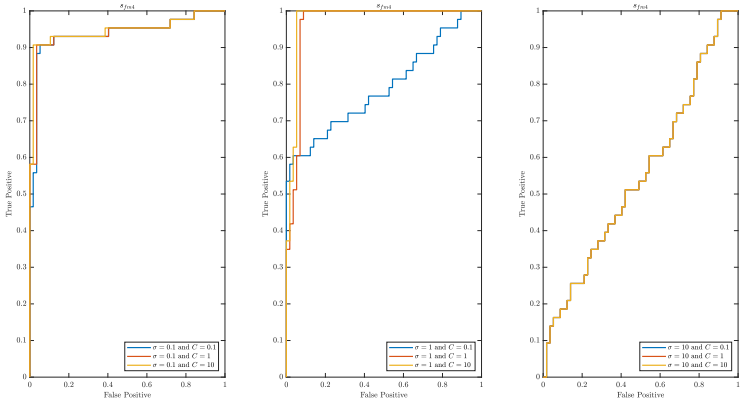


Figure C.24: ROC (Receiver Operating Characteristic) curves of s_{fm4} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

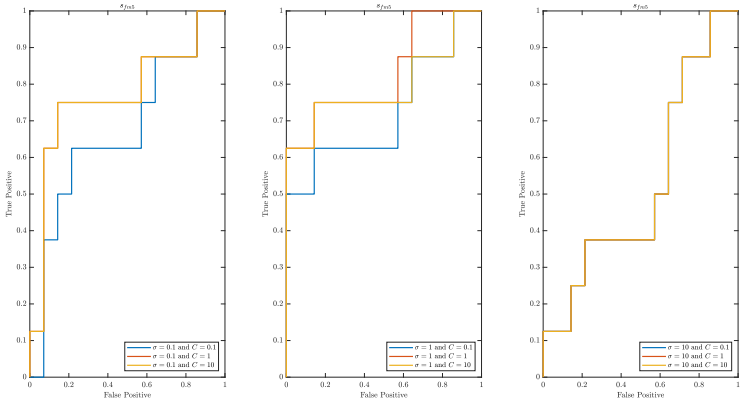


Figure C.25: ROC (Receiver Operating Characteristic) curves of s_{fm5} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

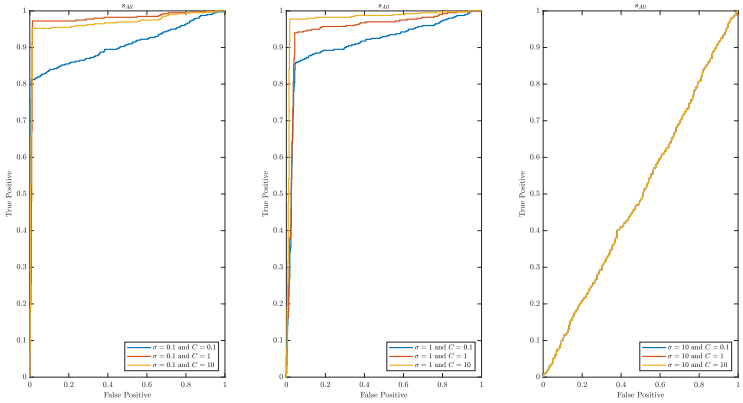


Figure C.26: ROC (Receiver Operating Characteristic) curves of s_{All} show the True Positive (Sensitivity) and False Positive (1-Specificity) for the best different thresholds using kernel property of SVM.

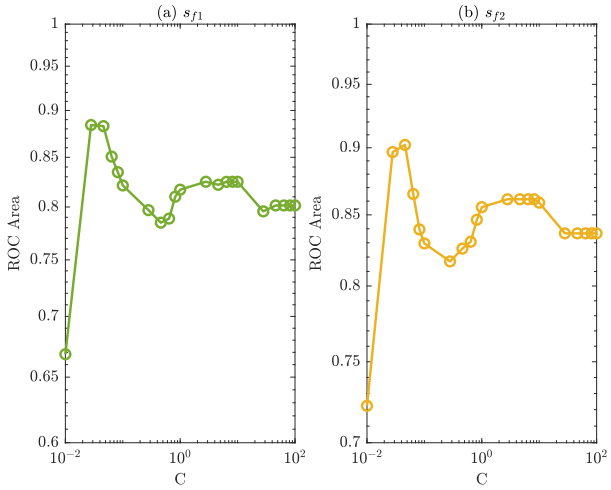


Figure C.27: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when (a) $\sigma = 0.1$ and $C = 10$ and (b) $\sigma = 0.1$ and $C = 10$.

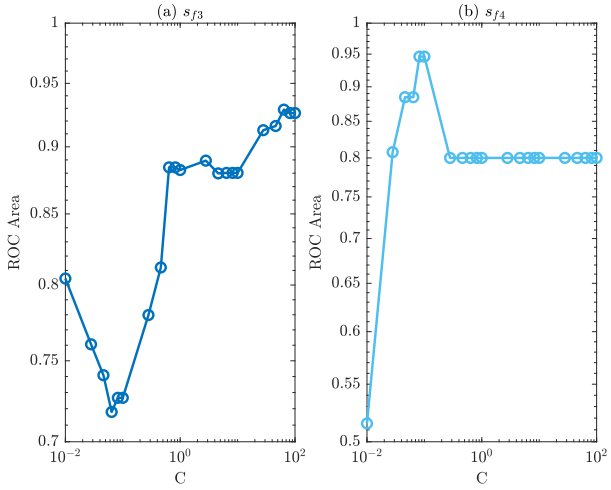


Figure C.28: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when (a) $\sigma = 1$ and $C = 1$ and (b) $\sigma = 0.1$ and $C = 1$.

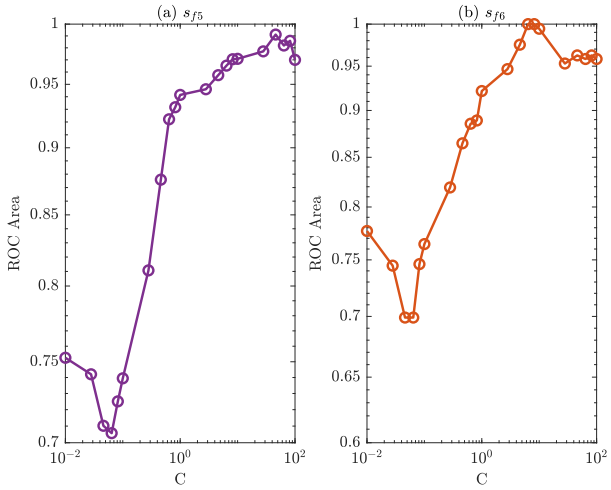


Figure C.29: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when (a) $\sigma = 1$ and $C = 10$ and (b) $\sigma = 1$ and $C = 10$.

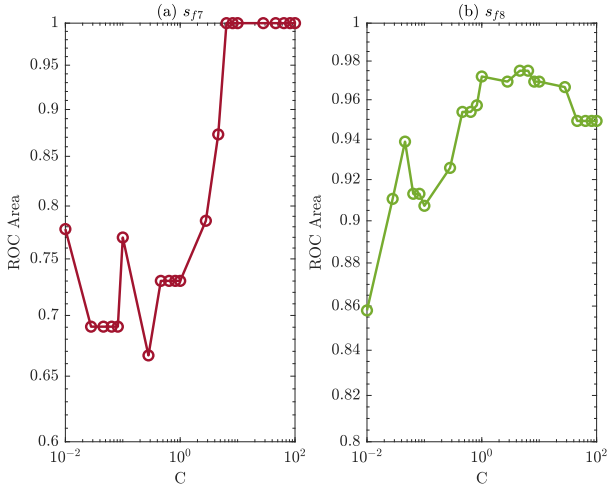


Figure C.30: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when (a) $\sigma = 1$ and $C = 10$ and (b) $\sigma = 0.1$ and $C = 10$.

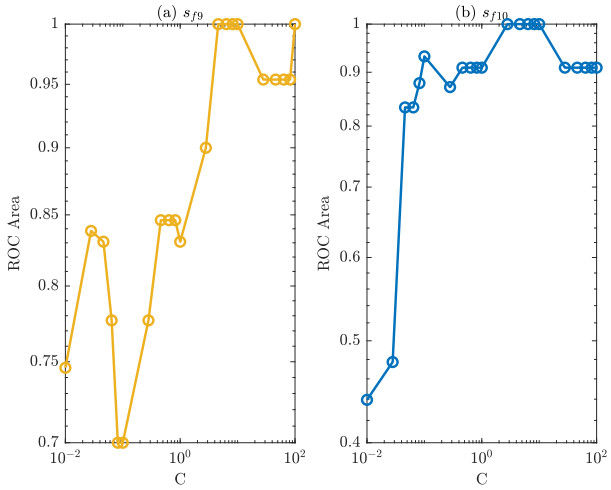


Figure C.31: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when (a) $\sigma = 0.1$ and $C = 10$ and (b) $\sigma = 0.1$ and $C = 10$.

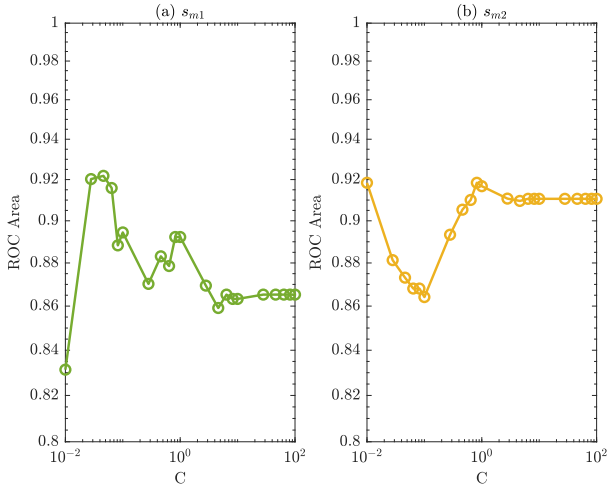


Figure C.32: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when (a) $\sigma = 0.1$ and $C = 1$ and (b) $\sigma = 0.1$ and $C = 1$.

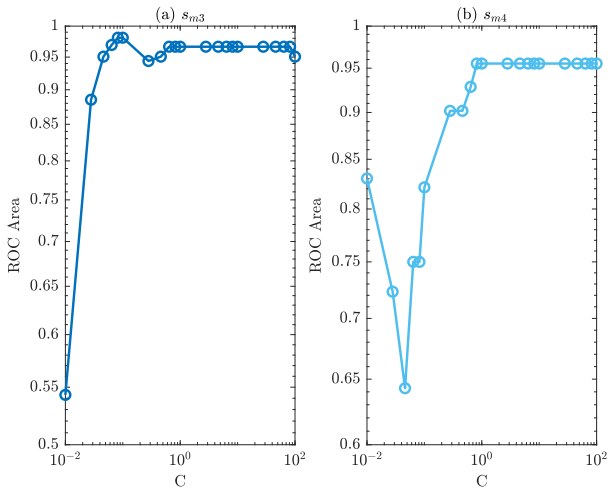


Figure C.33: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when (a) $\sigma = 0.1$ and $C = 0.1$ and (b) $\sigma = 1$ and $C = 10$.

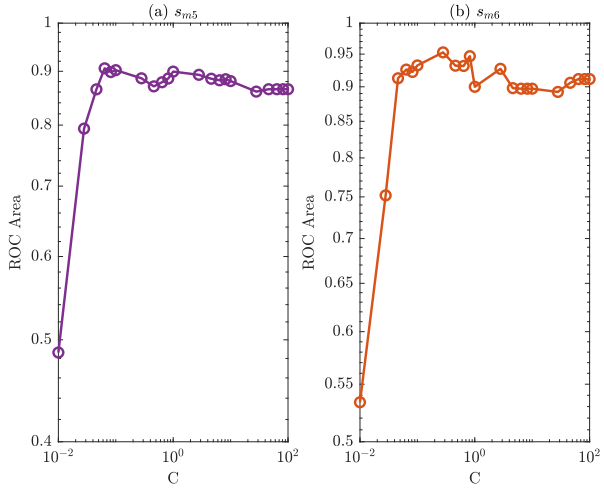


Figure C.34: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when (a) $\sigma = 0.1$ and $C = 1$ and (b) $\sigma = 0.1$ and $C = 0.1$.

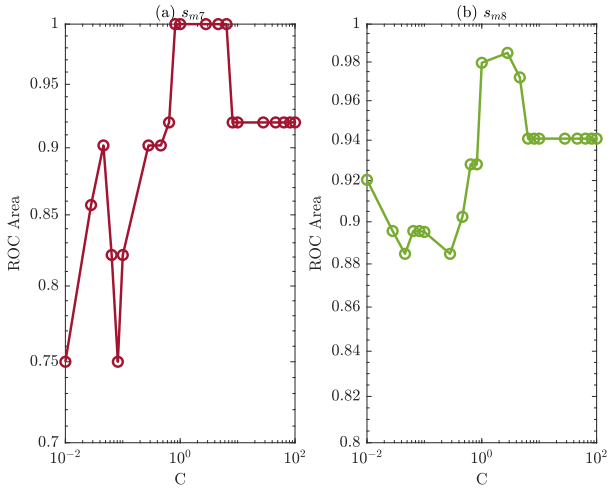


Figure C.35: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when (a) $\sigma = 1$ and $C = 1$ and (b) $\sigma = 0.1$ and $C = 1$.

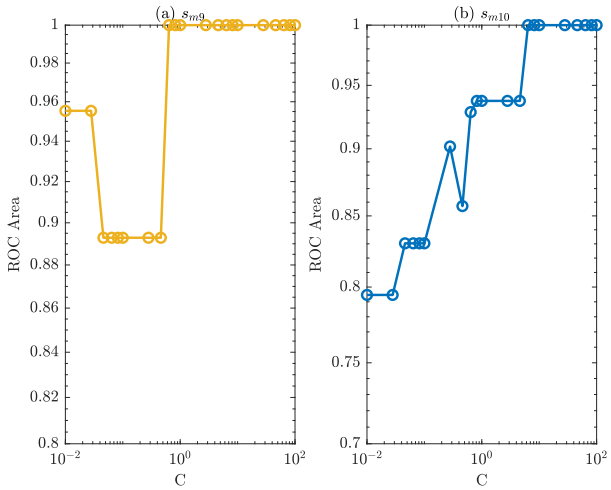


Figure C.36: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when (a) $\sigma = 0.1$ and $C = 10$ and (b) $\sigma = 1$ and $C = 10$.

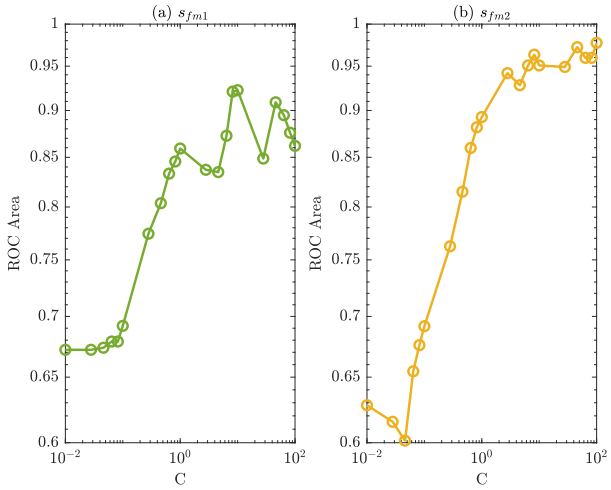


Figure C.37: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when (a) $\sigma = 1$ and $C = 10$ and (b) $\sigma = 1$ and $C = 10$.

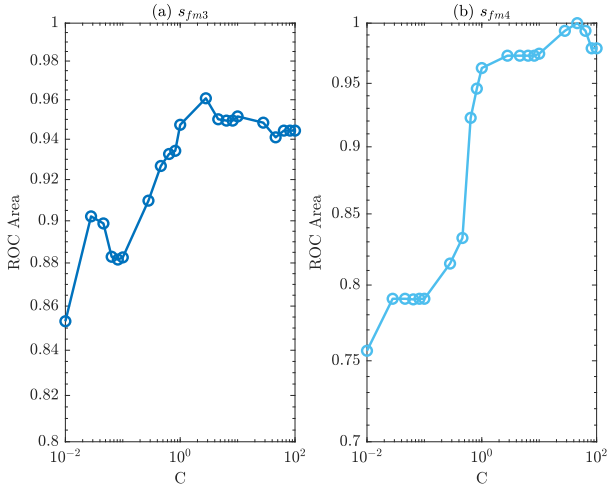


Figure C.38: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when (a) $\sigma = 0.1$ and $C = 10$ and (b) $\sigma = 1$ and $C = 10$.

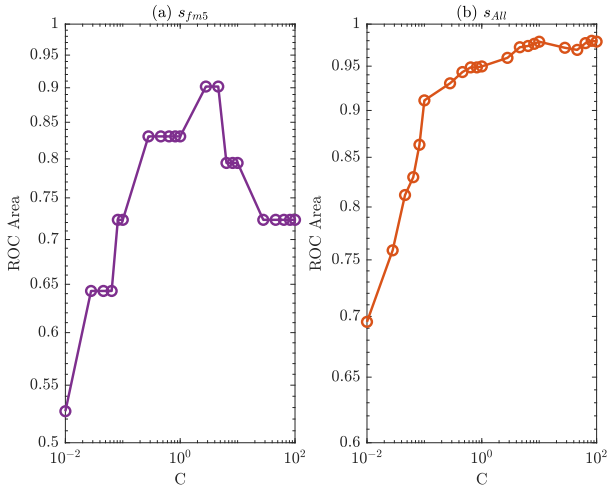
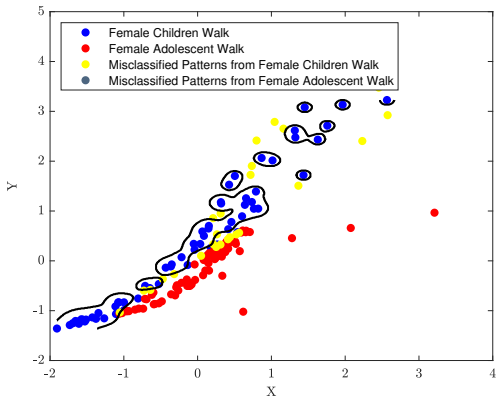
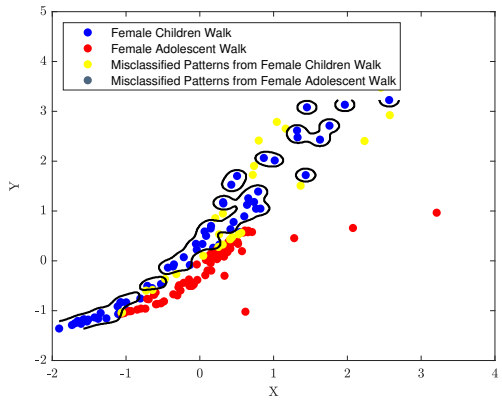


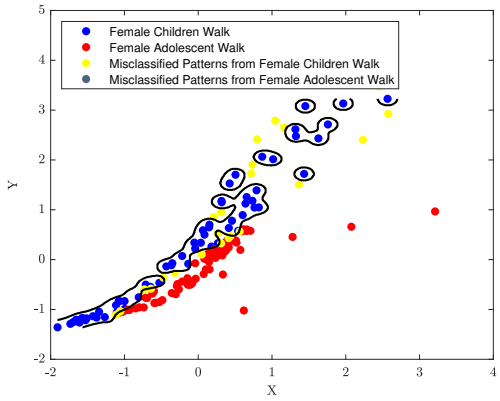
Figure C.39: The ROC area or Area under ROC curve (AROC) versus various values of the regularisation parameter C when (a) $\sigma = 1$ and $C = 1$ and (b) $\sigma = 1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

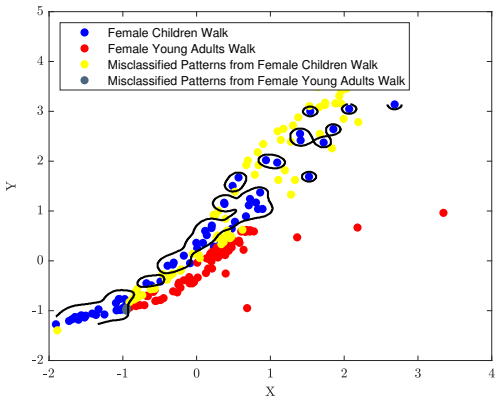


(b) $\sigma = 0.1$ and $C = 1$

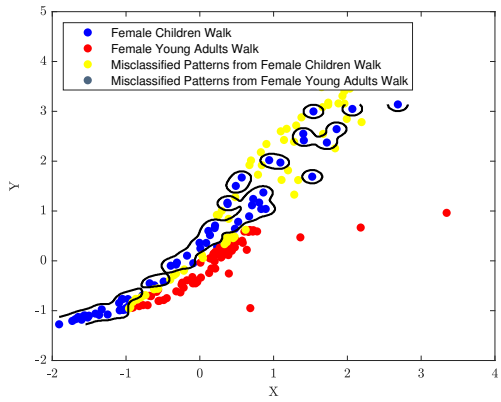


(c) $\sigma = 0.1$ and $C = 10$

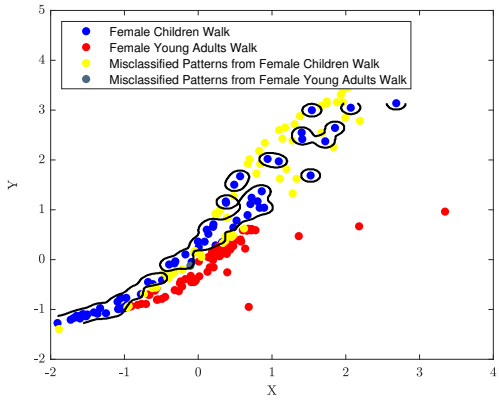
Figure C.40: The decision boundary when the SVM model trained on s_{f1} female group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

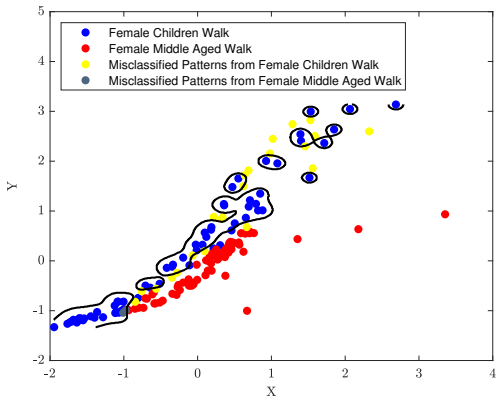


(b) $\sigma = 0.1$ and $C = 1$

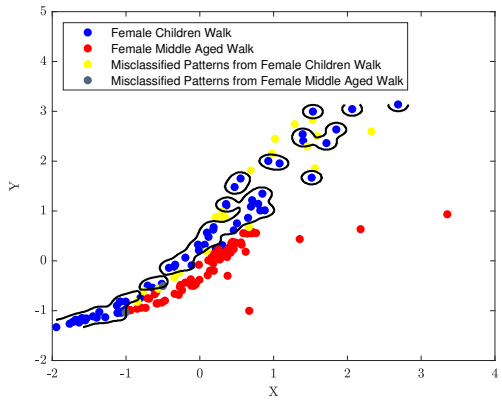


(c) $\sigma = 0.1$ and $C = 10$

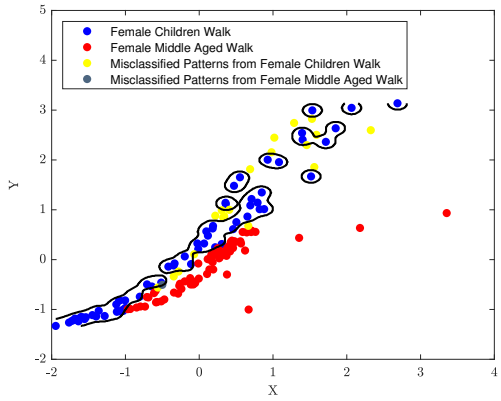
Figure C.41: The decision boundary when the SVM model trained on s_{f2} female group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

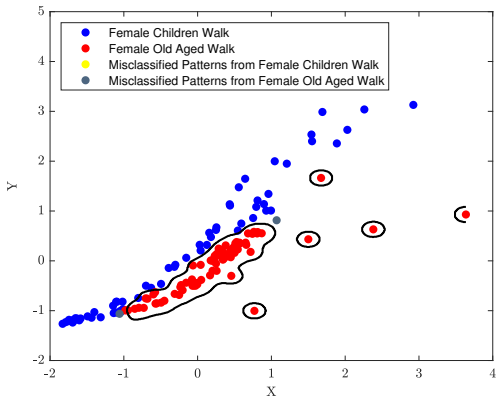


(b) $\sigma = 0.1$ and $C = 1$

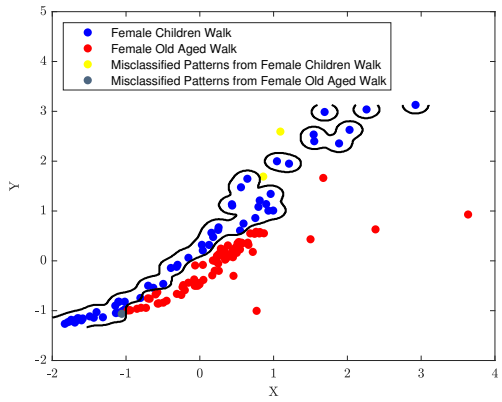


(c) $\sigma = 0.1$ and $C = 10$

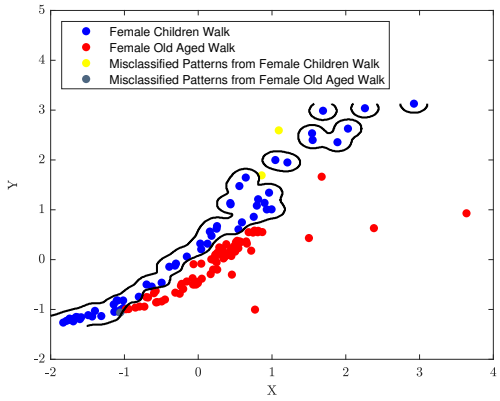
Figure C.42: The decision boundary when the SVM model trained on s_{f3} female group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

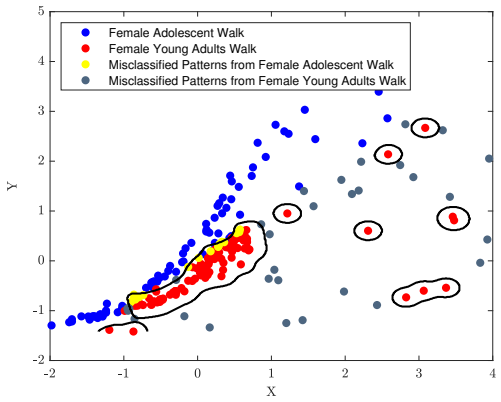


(b) $\sigma = 0.1$ and $C = 1$

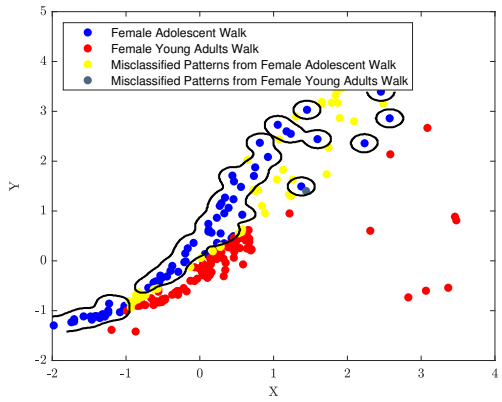


(c) $\sigma = 0.1$ and $C = 10$

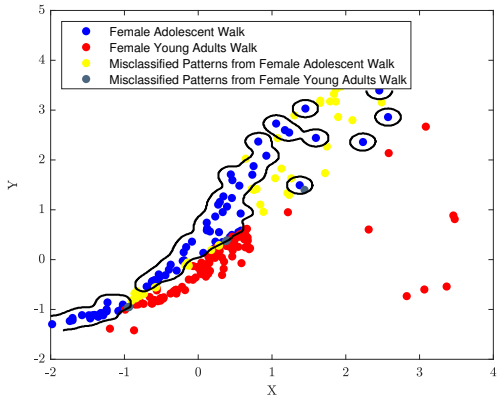
Figure C.43: The decision boundary when the SVM model trained on s_{f4} female group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

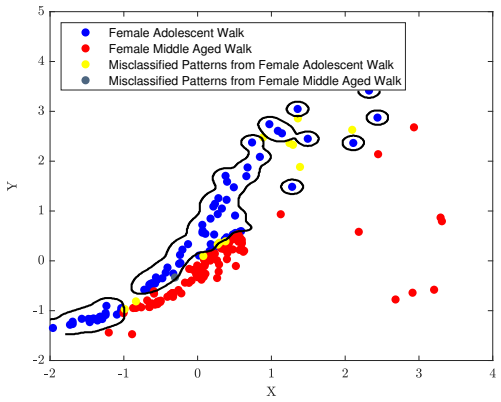


(b) $\sigma = 0.1$ and $C = 1$

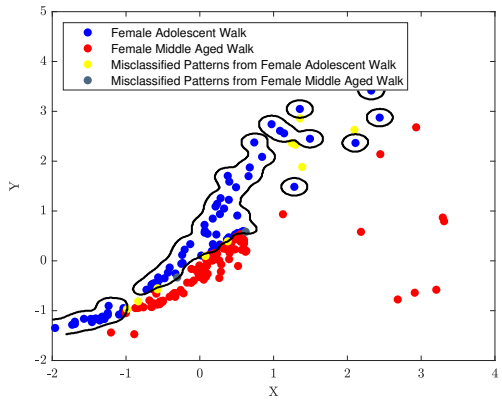


(c) $\sigma = 0.1$ and $C = 10$

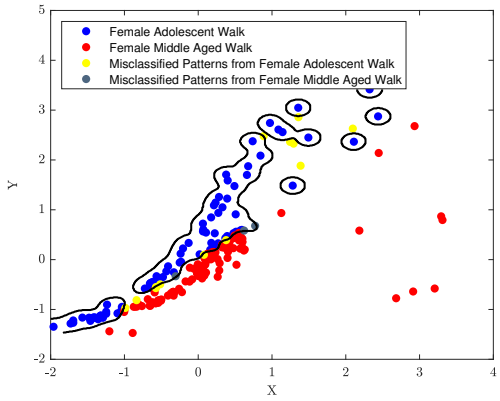
Figure C.44: The decision boundary when the SVM model trained on s_{f5} female group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

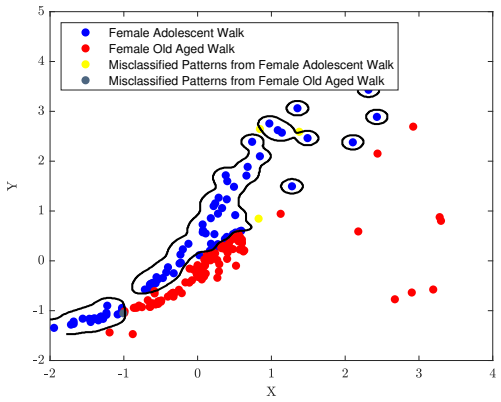


(b) $\sigma = 0.1$ and $C = 1$

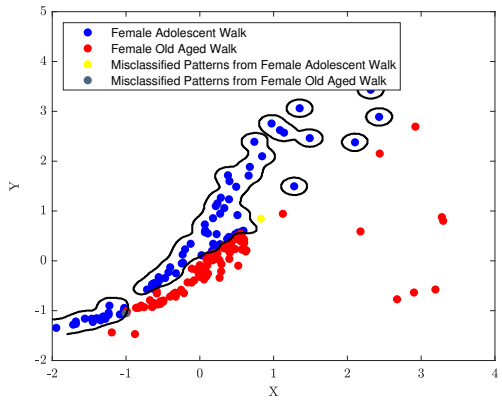


(c) $\sigma = 0.1$ and $C = 10$

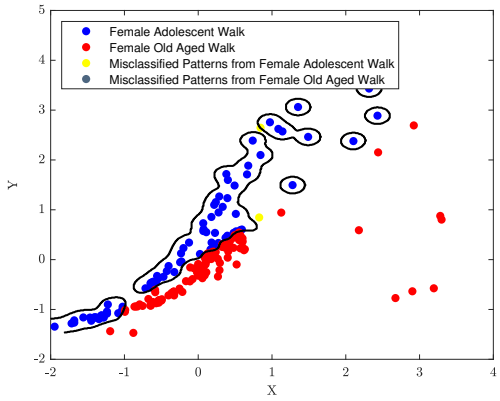
Figure C.45: The decision boundary when the SVM model trained on s_{f6} female group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

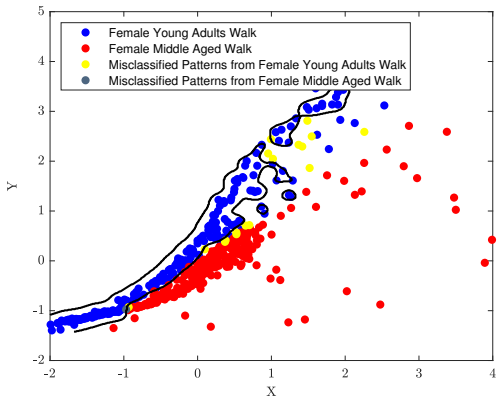


(b) $\sigma = 0.1$ and $C = 1$

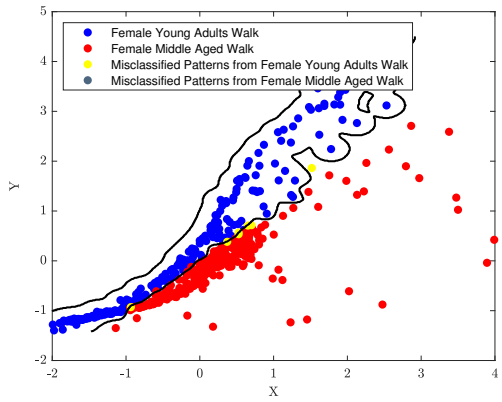


(c) $\sigma = 0.1$ and $C = 10$

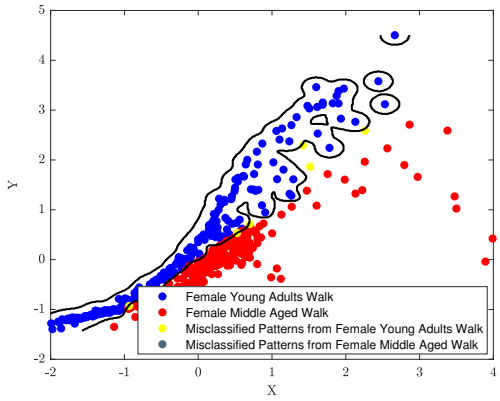
Figure C.46: The decision boundary when the SVM model trained on s_{f7} female group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

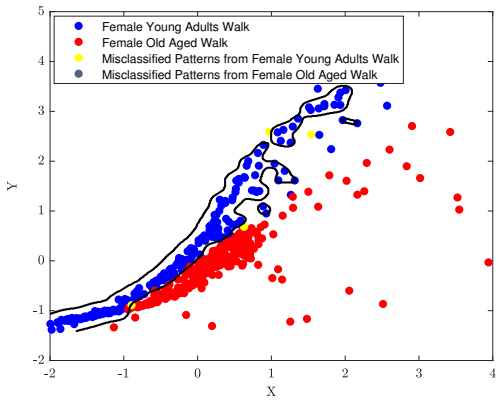


(b) $\sigma = 0.1$ and $C = 1$

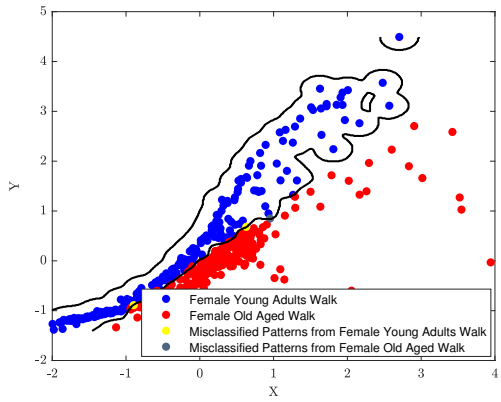


(c) $\sigma = 0.1$ and $C = 10$

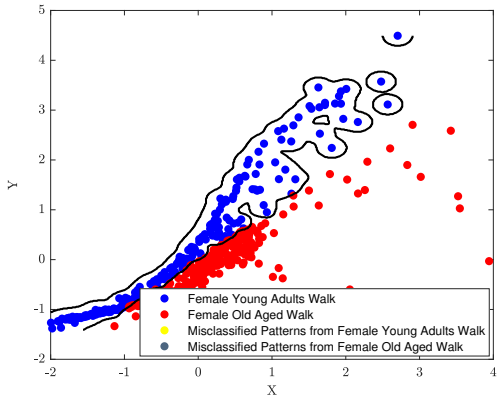
Figure C.47: The decision boundary when the SVM model trained on s_{f8} female group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

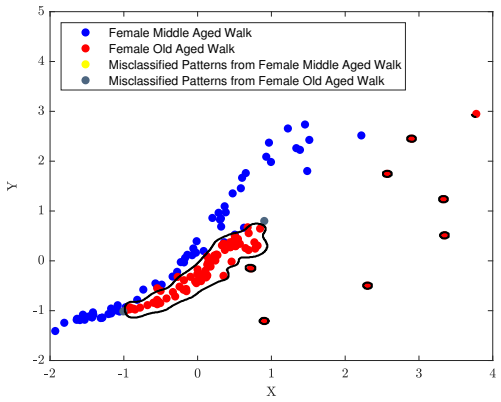


(b) $\sigma = 0.1$ and $C = 1$

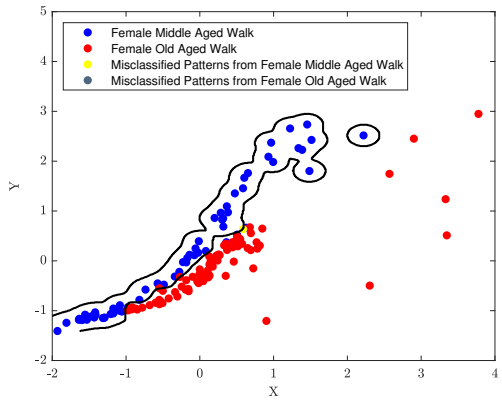


(c) $\sigma = 0.1$ and $C = 10$

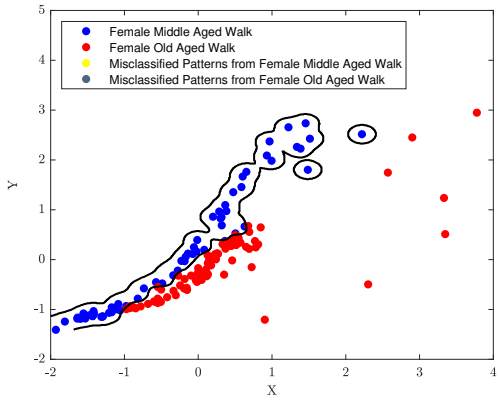
Figure C.48: The decision boundary when the SVM model trained on s_{f9} female group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

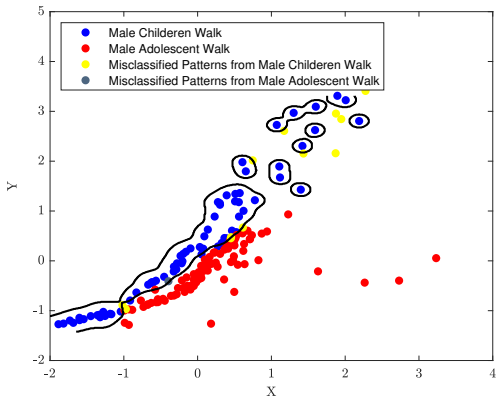


(b) $\sigma = 0.1$ and $C = 1$

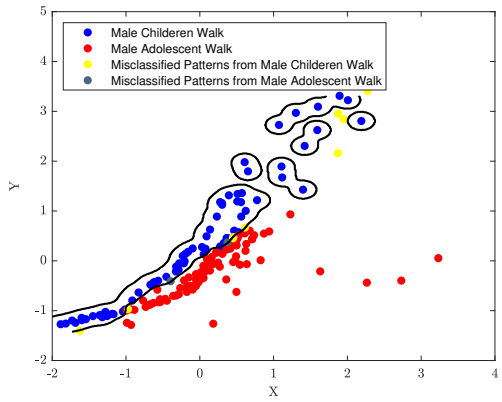


(c) $\sigma = 0.1$ and $C = 10$

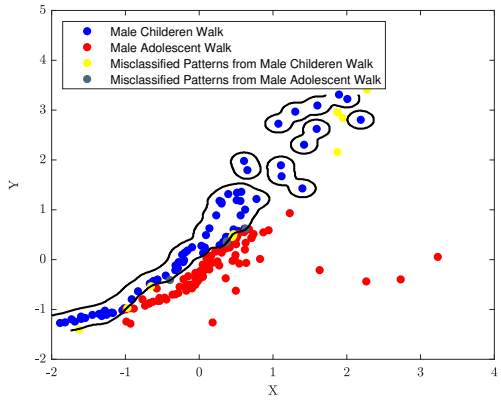
Figure C.49: The decision boundary when the SVM model trained on s_{f10} female group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

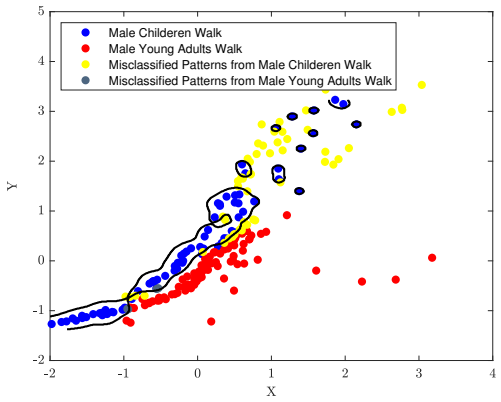


(b) $\sigma = 0.1$ and $C = 1$

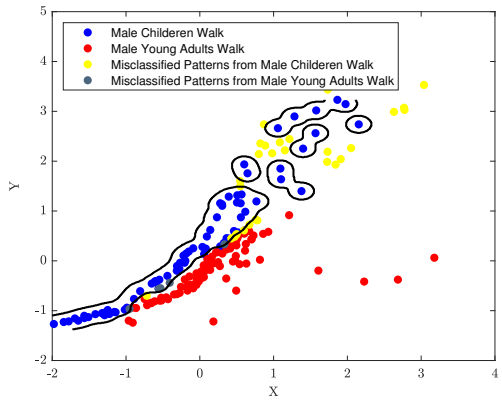


(c) $\sigma = 0.1$ and $C = 10$

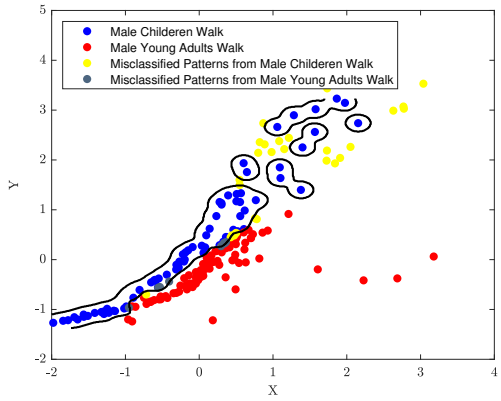
Figure C.50: The decision boundary when the SVM model trained on s_{m1} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

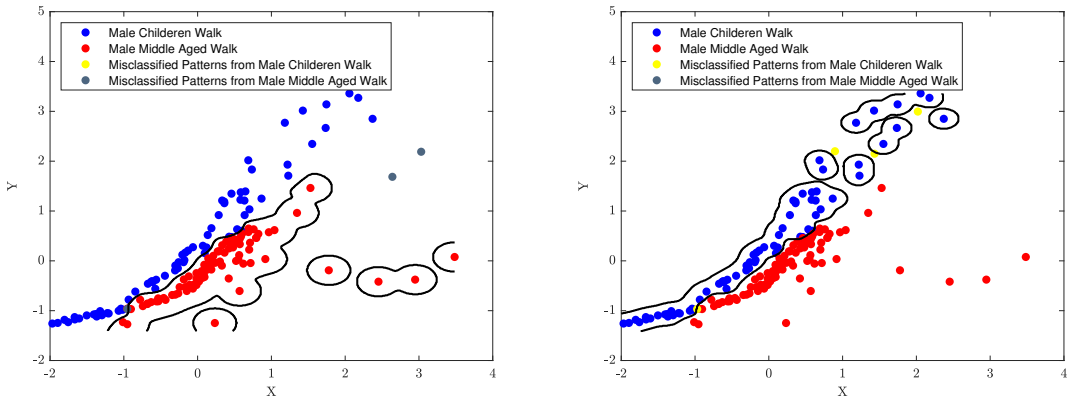


(b) $\sigma = 0.1$ and $C = 1$



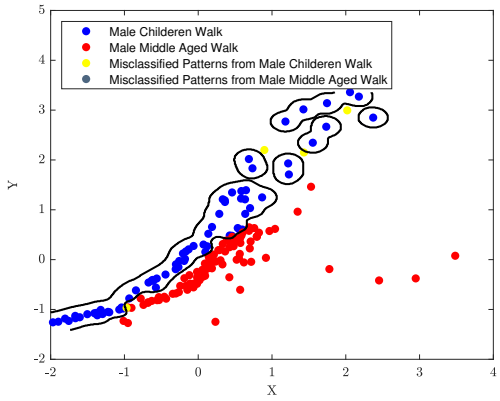
(c) $\sigma = 0.1$ and $C = 10$

Figure C.51: The decision boundary when the SVM model trained on s_{m2} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



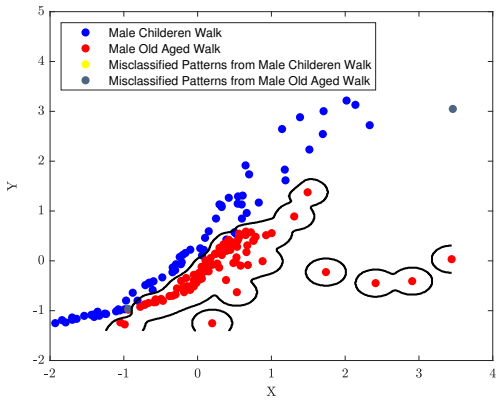
(a) $\sigma = 0.1$ and $C = 0.1$

(b) $\sigma = 0.1$ and $C = 1$

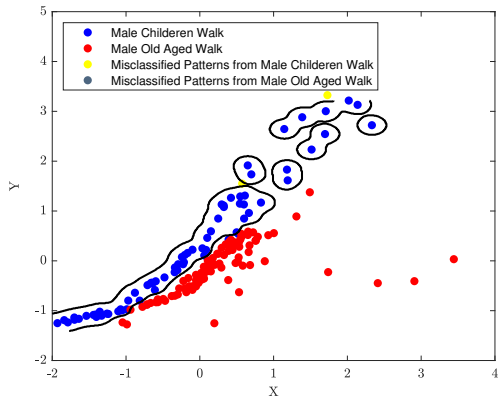


(c) $\sigma = 0.1$ and $C = 10$

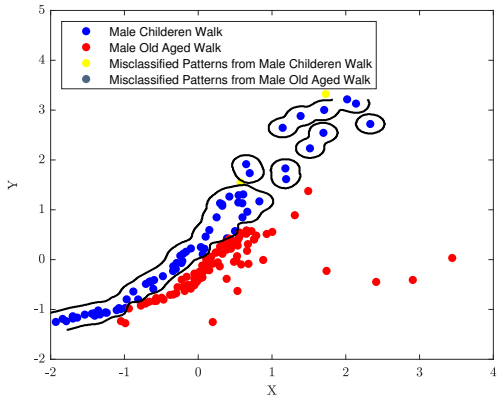
Figure C.52: The decision boundary when the SVM model trained on s_{m3} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

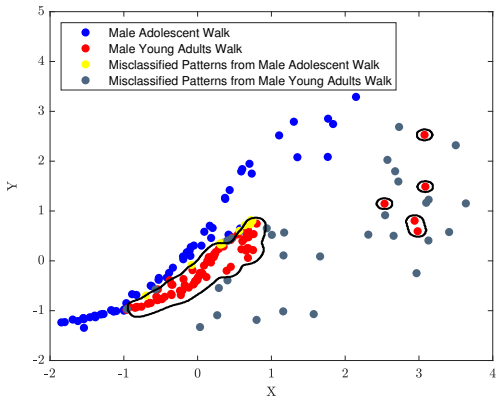


(b) $\sigma = 0.1$ and $C = 1$

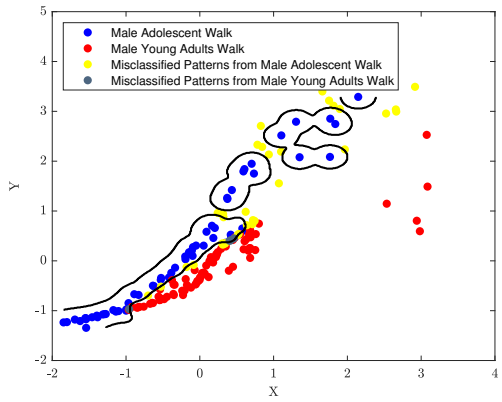


(c) $\sigma = 0.1$ and $C = 10$

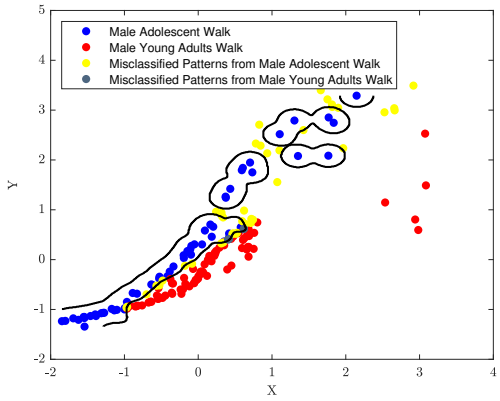
Figure C.53: The decision boundary when the SVM model trained on s_{m4} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

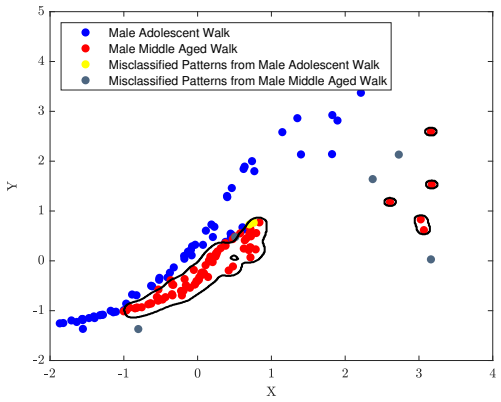


(b) $\sigma = 0.1$ and $C = 1$

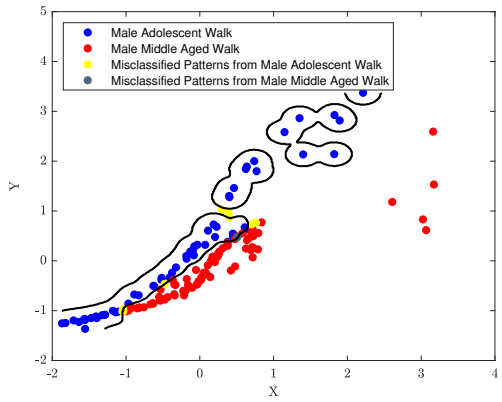


(c) $\sigma = 0.1$ and $C = 10$

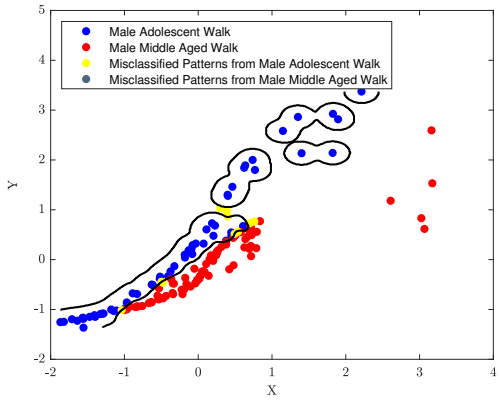
Figure C.54: The decision boundary when the SVM model trained on s_{m5} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

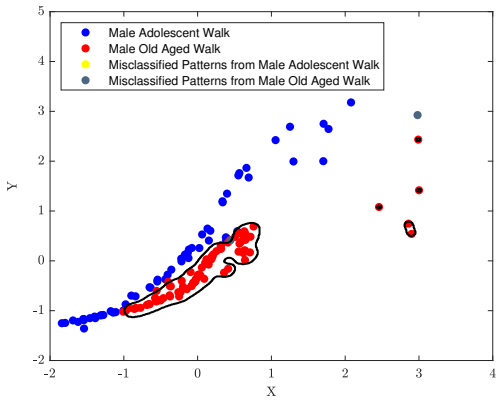


(b) $\sigma = 0.1$ and $C = 1$

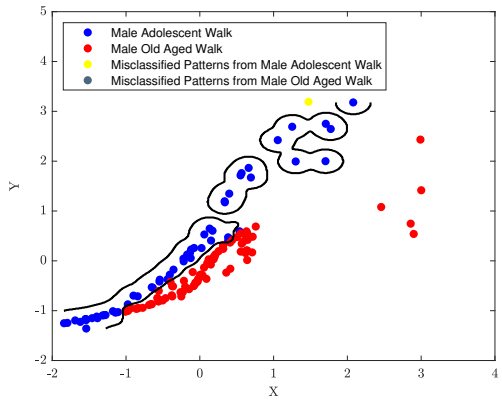


(c) $\sigma = 0.1$ and $C = 10$

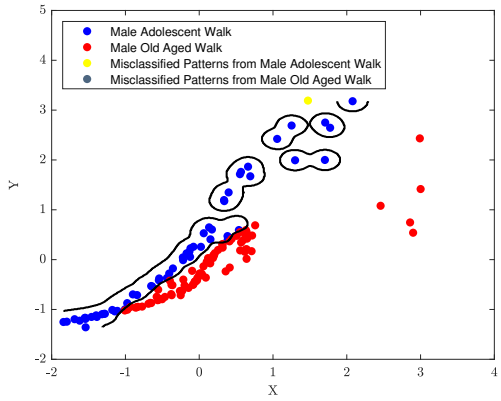
Figure C.55: The decision boundary when the SVM model trained on s_{m6} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

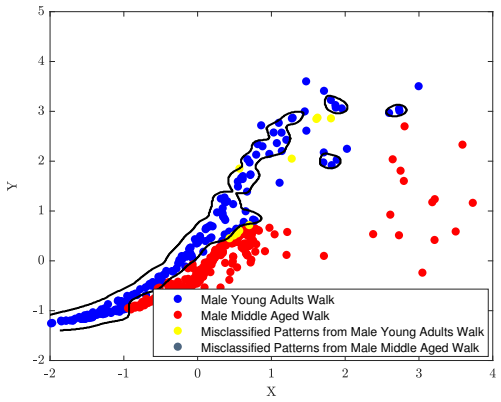


(b) $\sigma = 0.1$ and $C = 1$

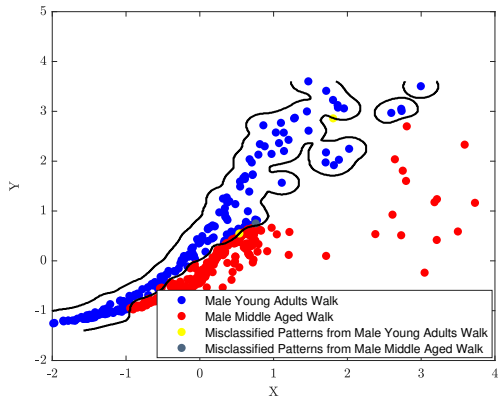


(c) $\sigma = 0.1$ and $C = 10$

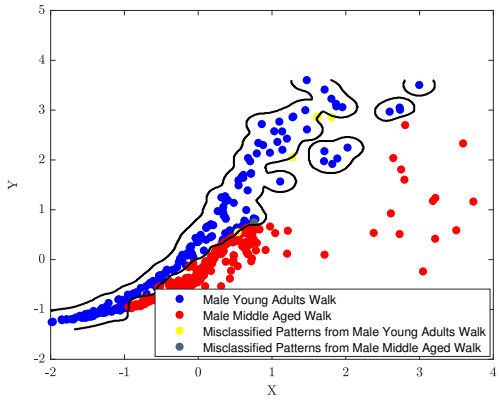
Figure C.56: The decision boundary when the SVM model trained on s_{m7} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

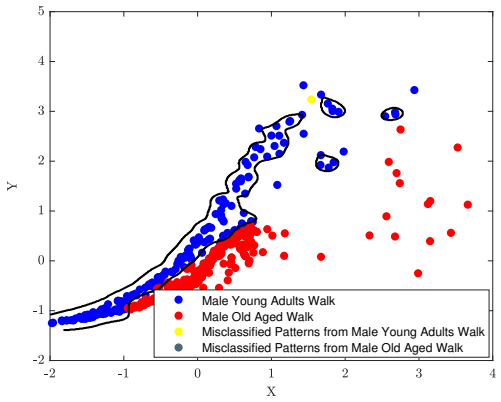


(b) $\sigma = 0.1$ and $C = 1$

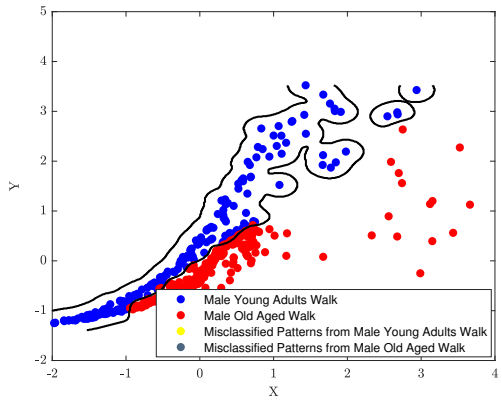


(c) $\sigma = 0.1$ and $C = 10$

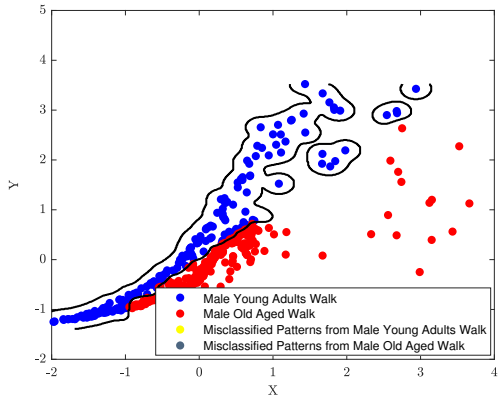
Figure C.57: The decision boundary when the SVM model trained on s_{m8} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

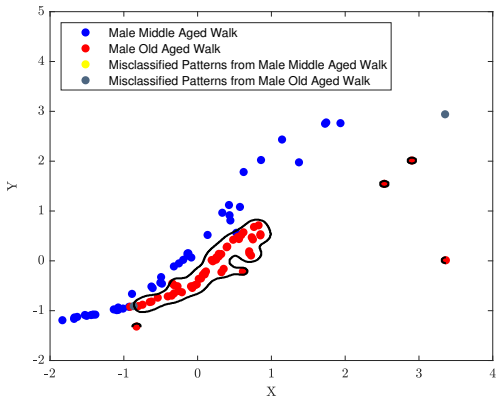


(b) $\sigma = 0.1$ and $C = 1$

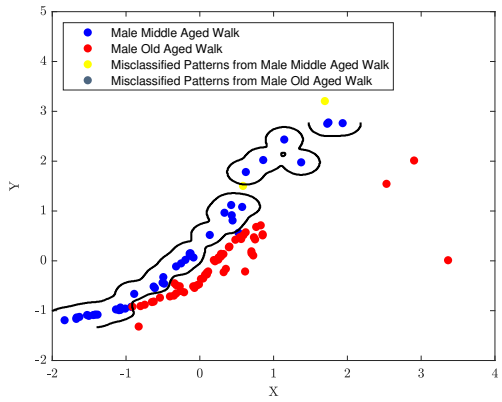


(c) $\sigma = 0.1$ and $C = 10$

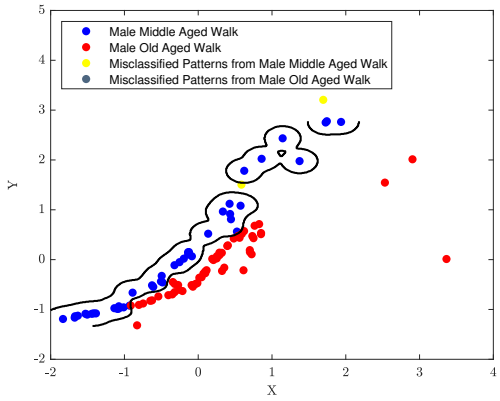
Figure C.58: The decision boundary when the SVM model trained on s_{m9} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

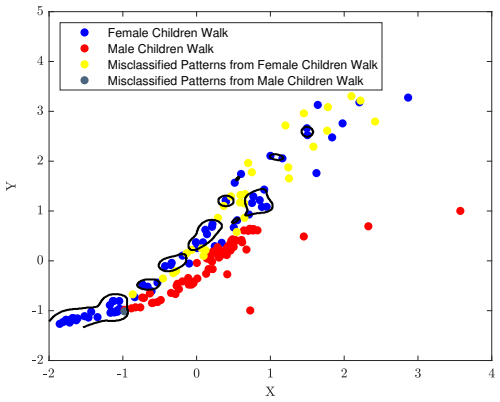


(b) $\sigma = 0.1$ and $C = 1$

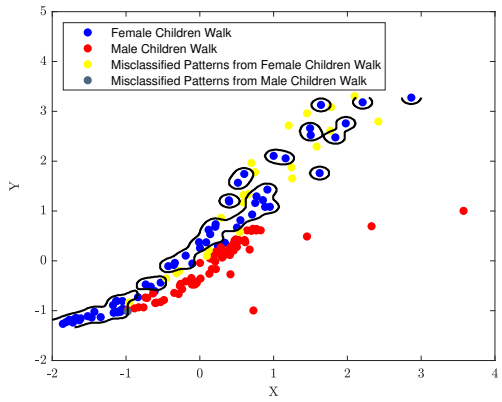


(c) $\sigma = 0.1$ and $C = 10$

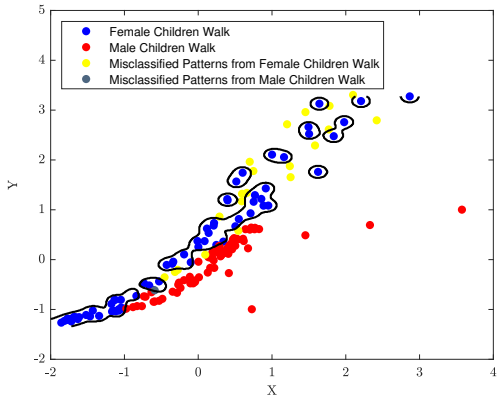
Figure C.59: The decision boundary when the SVM model trained on s_{m10} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

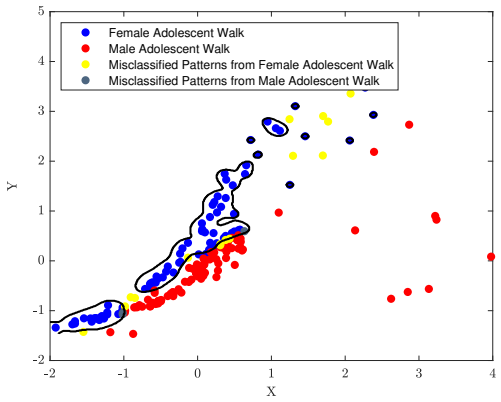


(b) $\sigma = 0.1$ and $C = 1$

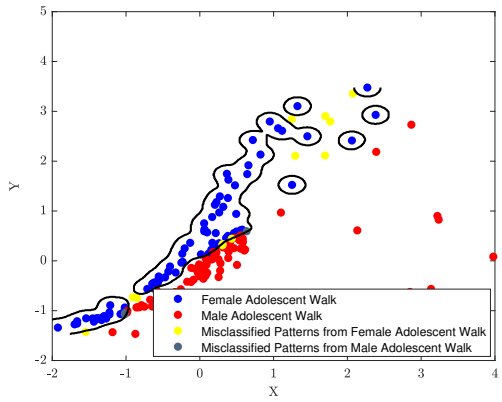


(c) $\sigma = 0.1$ and $C = 10$

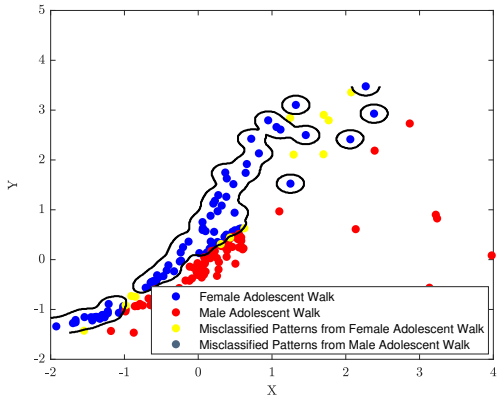
Figure C.60: The decision boundary when the SVM model trained on s_{fm1} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

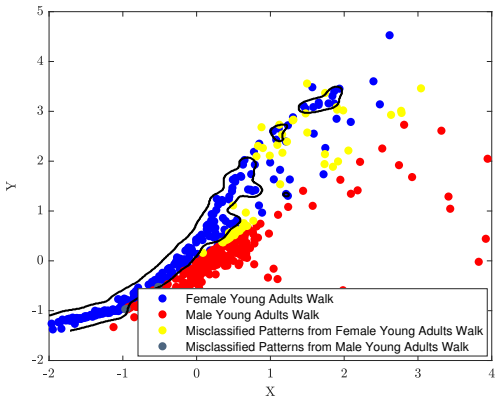


(b) $\sigma = 0.1$ and $C = 1$

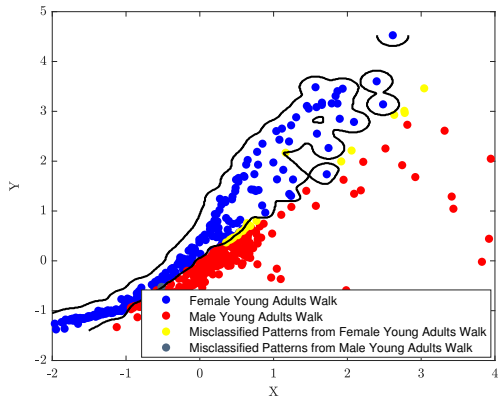


(c) $\sigma = 0.1$ and $C = 10$

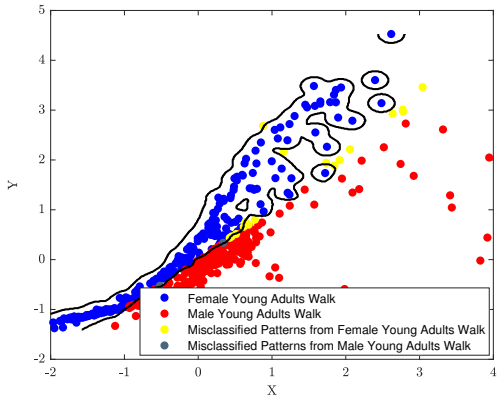
Figure C.61: The decision boundary when the SVM model trained on s_{fm2} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

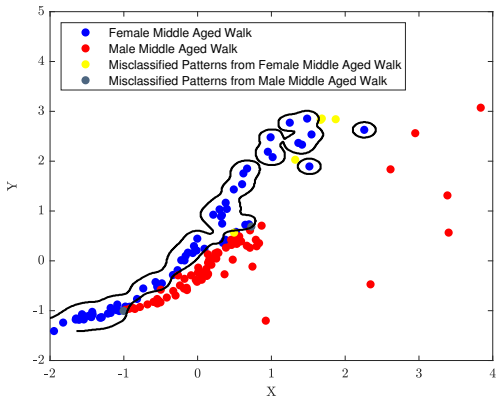


(b) $\sigma = 0.1$ and $C = 1$

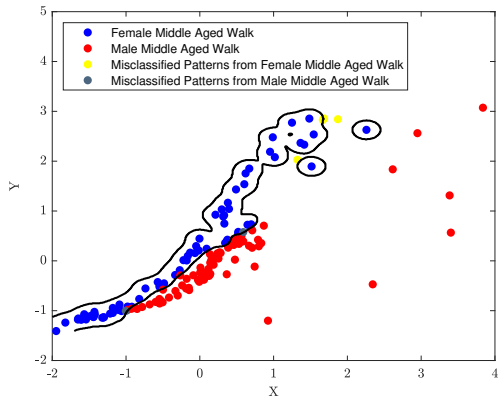


(c) $\sigma = 0.1$ and $C = 10$

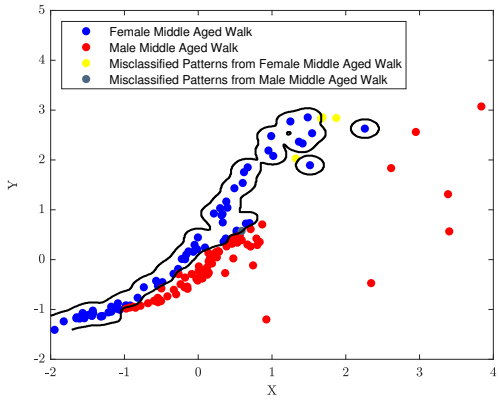
Figure C.62: The decision boundary when the SVM model trained on s_{fm3} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

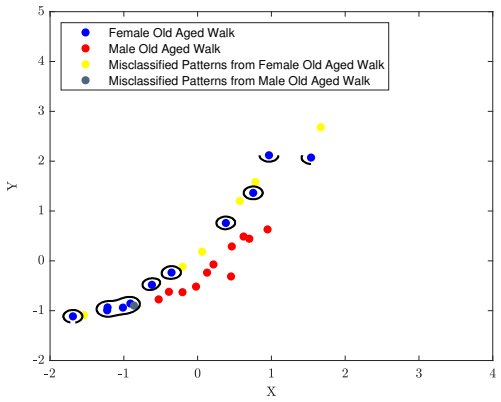


(b) $\sigma = 0.1$ and $C = 1$

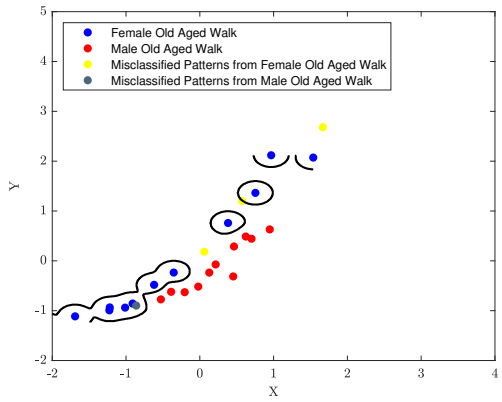


(c) $\sigma = 0.1$ and $C = 10$

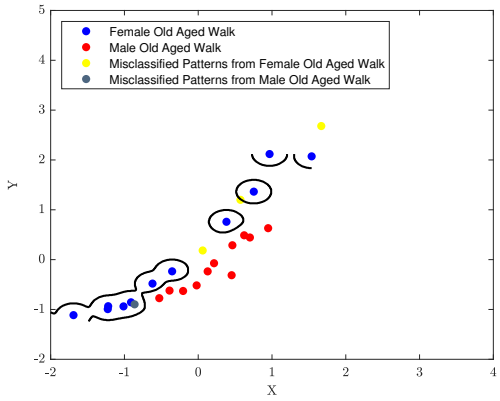
Figure C.63: The decision boundary when the SVM model trained on s_{fm4} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$

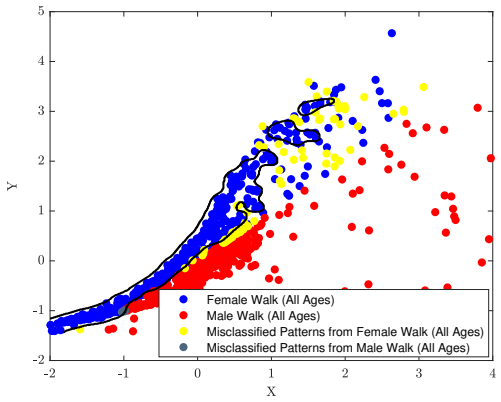


(b) $\sigma = 0.1$ and $C = 1$

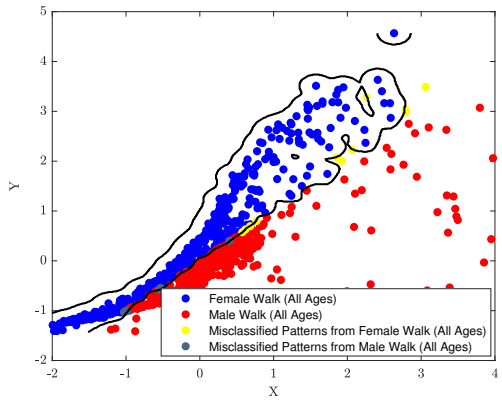


(c) $\sigma = 0.1$ and $C = 10$

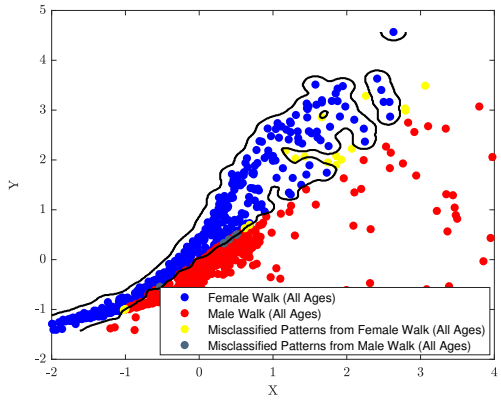
Figure C.64: The decision boundary when the SVM model trained on s_{fm5} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.



(a) $\sigma = 0.1$ and $C = 0.1$



(b) $\sigma = 0.1$ and $C = 1$



(c) $\sigma = 0.1$ and $C = 10$

Figure C.65: The decision boundary when the SVM model trained on s_{All} male group walk patterns, when: (**a**) $\sigma = 0.1$ and $C = 0.1$, (**b**) $\sigma = 0.1$ and $C = 1$, and (**c**) $\sigma = 0.1$ and $C = 10$.

Bibliography

[1] Van Emmerik, R.E., Hamill, J. and McDermott, W.J., 2005. Variability and coordinative function in human gait. Quest, 57(1), pp.102-123.

[2] Rosengren, K.S., McAuley, E. and Mihalko, S.L., 1998. Gait adjustments in older adults: activity and efficacy influences. Psychology and aging, 13(3), p.375.

[3] Henriksen, M., Lund, H., Moe-Nilssen, R., Bliddal, H. and Danneskiod-Samsøe, B., 2004. Test–retest reliability of trunk accelerometric gait analysis. Gait and posture, 19(3), pp.288-297.

[4] Eltanani, S.; Scheper, T. O.; Dawes, H. K. Nearest Neighbour Algorithm: Proposed Solution for Human Gait Data Classification. In Proceedings of the IEEE Symposium on Computers and Communications (ISCC), Athens, Greece, 5–8 September 2021. https://doi.org/10.1109/iscc53001.2021.9631454.

[5] DeLuca PA. Gait analysis in the treatment of the ambulatory child with cerebral palsy. Clin Orthop Relat Res. 1991 Mar;(264):65-75. PMID: 1997253.

[6] Mora, T. and Bialek, W., 2011. Are biological systems poised at criticality?. Journal of Statistical Physics, 144, pp.268-302.

[7] Wrobel, J.S. and Najafi, B., 2010. Diabetic foot biomechanics and gait dysfunction. Journal of diabetes science and technology, 4(4), pp.833-845.

[8] Jordan, K., Challis, J.H. and Newell, K.M., 2006. Long range correlations in the stride interval of running. Gait and posture, 24(1), pp.120-125.

[9] Hausdorff, J.M., 2007. Gait dynamics, fractals and falls: finding meaning in the stride-to-stride fluctuations of human walking. Human movement science, 26(4), pp.555-589.

[10] Van Emmerik, R.E. and Van Wegen, E.E., 2002. On the functional aspects of variability in postural control. Exercise and sport sciences reviews, 30(4), pp.177-183.

[11] Morris, M.E., Huxham, F., McGinley, J., Dodd, K. and Iansek, R., 2001. The biomechanics and motor control of gait in Parkinson disease. Clinical biomechanics, 16(6), pp.459-470.

[12] Tao, W., Liu, T., Zheng, R. and Feng, H., 2012. Gait analysis using wearable sensors. Sensors, 12(2), pp.2255-2283.

[13] Scheper, T.V.S.M.O. and Carnell, A.R., Oxford Brookes University, 2017. Method of controlling a dynamic physical system that exhibits a chaotic behaviour. U.S. Patent 9,740,180.

[14] Patla, A.E. ed., 1991. Adaptability of human gait: implications for the control of locomotion. Elsevier.

[15] Perc, M., 2005. The dynamics of human gait. European journal of physics, 26(3), p.525.

[16] Van Emmerik, R.E., Ducharme, S.W., Amado, A.C. and Hamill, J., 2016. Comparing dynamical systems concepts and techniques for biomechanical analysis. Journal of Sport and Health Science, 5(1), pp.3-13.

[17] Den Hartog, J.P., 1985. Mechanical vibrations. Courier Corporation.

[18] Blickhan, R., 1989. The spring-mass model for running and hopping. Journal of biomechanics, 22(11-12), pp.1217-1227.

[19] Farley, C.T. and Gonzalez, O., 1996. Leg stiffness and stride frequency in human running. Journal of biomechanics, 29(2), pp.181-186.

[20] Farley, C.T., Blickhan, R., Saito, J. and Taylor, C.R., 1991. Hopping frequency in humans: a test of how springs set stride frequency in bouncing gaits. Journal of applied physiology, 71(6), pp.2127-2132.

[21] McMahon, T.A. and Cheng, G.C., 1990. The mechanics of running: how does stiffness couple with speed?. Journal of biomechanics, 23, pp.65-78.

[22] Kram, R., Domingo, A. and Ferris, D.P., 1997. Effect of reduced gravity on the preferred walk-run transition speed. The Journal of experimental biology, 200(4), pp.821-826.

[23] MacKinnon, C.D. and Winter, D.A., 1993. Control of whole body balance in the frontal plane during human walking. Journal of biomechanics, 26(6), pp.633-644.

[24] Cavagna, G.A., Thys, H. and Zamboni, A., 1976. The sources of external work in level walking and running. The Journal of physiology, 262(3), pp.639-657.

[25] Mochon, S. and McMahon, T.A., 1980. Ballistic walking. Journal of biomechanics, 13(1), pp.49-57.

[26] Whittlesey, S.N., van Emmerik, R.E. and Hamill, J., 2000. The swing phase of human walking is not a passive movement. Motor control, 4(3), pp.273-292.

[27] Holt, K.G., Hamill, J. and Andres, R.O., 1990. The force-driven harmonic oscillator as a model for human locomotion. Human Movement Science, 9(1), pp.55-68.

[28] Decroly O, Goldbeter A. From simple to complex oscillatory behaviour: analysis of bursting in a multiply regulated biochemical system. J Theor Biol. 1987 Jan 21;124(2):219-50. doi: 10.1016/s0022-5193(87)80264-3. Erratum in: J Theor Biol 1987 Oct 21;128(4):527. PMID: 3657195.

[29] Berry, H. Chaos in a Bienzymatic Cyclic Model with Two Autocatalytic Loops. Chaos Solitons Fractals 2003, 18, 1001–1014. https://doi.org/10.1016/s0960-0779(03)00070-5.

[30] Ott, E., 2002. Chaos in dynamical systems. Cambridge university press.

[31] Close, C.M., Frederick, D.K. and Newell, J.C., 2001. Modeling and analysis of dynamic systems. John Wiley and Sons.

[32] Gorley, P.N., Horley, P.P. and Tomchuk, P.M., 1999. Trajectory tracing—a new method of studying the evolution of states of dynamic systems. Technical Physics Letters, 25(1), pp.7-9.

[33] Thelen, E. and Smith, L.B., 1998. Dynamic systems theories.

[34] olde Scheper TV. Criticality in Biocomputation. In Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Bruges, Belgium, 26–28 April 2017; pp. 1–6.

[35] Kyrychko, Y. N.; Blyuss, K. B.; Hogan, S. J.; Schöll, E. Control of Spatiotemporal Patterns in the Gray–Scott Model. Chaos Interdiscip. J. Nonlinear Sci. 2009, 19, 043126. https://doi.org/10.1063/1.3270048.

[36] Larreta-Garde, V. and Berry, H., 2002. Modeling extracellular matrix degradation balance with proteinase/transglutaminase cycle. Journal of theoretical biology, 217(1), pp.105-124.

[37] Yamada, K.M., 1997. Integrin signaling. Matrix biology, 16(4), pp.137-141.

[38] Werb, Z., Tremble, P.M., Behrendtsen, O., Crowley, E. and Damsky, C.H., 1989. Signal transduction through the fibronectin receptor induces collagenase and stromelysin gene expression. The Journal of cell biology, 109(2), pp.877-889.

[39] Haroon, Z.A., Lai, T.S., Hettasch, J.M., Lindberg, R.A., Dewhirst, M.W. and Greenberg, C.S., 1999. Tissue transglutaminase is expressed as a host response to tumor invasion and inhibits tumor growth. Laboratory investigation; a journal of technical methods and pathology, 79(12), pp.1679-1686.

[40] Eltanani, Shadi, Tjeerd V. olde Scheper, and Helen Dawes. 2022. "A Novel Criticality Analysis Technique for Detecting Dynamic Disturbances in Human Gait" Computers 11, no. 8: 120. https://doi.org/10.3390/computers11080120.

[41] olde Scheper, T.V. Self-Organised Criticality equation Files [Data set]. Zenodo 2021. https://doi.org/10.5281/zenodo.5727044.

[42] Vapnik, V. Statistical Learning Theory; Wiley: New York, NY, USA, 1998.

[43] Cristianini, N.; Shawe-Taylor, J. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods; Cambridge University Press: Cambridge, UK, 2000.

[44] Hastie, T.; Tibshirani, R.; Friedman, J. The Elements of Statistical Learning; Springer: New York, NY, USA, 2001.

[45] Boyd, S.; Vandenberghe, L. Convex Optimization; Cambridge University Press: Cambridge, UK, 2004.

[46] Shalev-Shwartz, S.; Ben-David, S. Understanding Machine Learning—From Theory to Algorithms; Cambridge University Press: Cambridge, UK, 2014; pp. 1–397.

[47] Que, Q.; Belkin, M. Back to the Future: Radial Basis Function Network Revisited. In IEEE Transactions on Pattern Analysis and Machine Intelligence; IEEE: Piscataway, NJ, USA, 2020; Volume 42, pp. 1856–1867. https://doi.org/10.1109/TPAMI.2019.2906594.

[48] Panchapakesan, C.; Ralph, D.; Palaniswami, M. Effects of Moving the Centers in an RBF Network. In Proceedings of the IEEE International Joint Conference on Neural Networks Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98CH36227), Anchorage, AK, USA, 4–9 May 1998. https://doi.org/10.1109/ijcnn.1998.685954.

[49] AX3 GUI · digitalinteraction/openmovement Wiki. Available online: https://github.com/digitalinteraction/openmovement/wiki/AX3-GUI (accessed on 1 April 2022).

[50] olde Scheper, T.V. Controlled Bio-Inspired Self-Organised Criticality. Plos ONE 2022, 17, e0260016. https://doi.org/10.1371/journal.pone.0260016.

[51] CVX: Matlab Software for Disciplined Convex Programming|CVX Research, Inc. 2012. Available online: http://cvxr.com/cvx (accessed on 15 May 2022).

[52] Sun, L. Multi-Class Associative Classification Based on Intersection Method and Extended Chi-Square Testing. J. Comput. Appl. 2008, 28, 1692–1695. https://doi.org/10.3724/sp.j.1087.2008.01692.

[53] Varrecchia, T.; Castiglia, S.F.; Ranavolo, A.; Conte, C.; Tatarelli, A.; Coppola, G.; Di Lorenzo, C.; Draicchio, F.; Pierelli, F.; Serrao, M. An Artificial Neural Network Approach to Detect Presence and Severity of Parkinson's Disease via Gait Parameters. PLoS ONE 2021, 16, e0244396. https://doi.org/10.1371/journal.pone.0244396.

[54] Wang, F.-C.; Chen, S.-F.; Lin, C.-H.; Shih, C.-J.; Lin, A.-C.; Yuan, W.; Li, Y.-C.; Kuo, T.-Y. Detection and Classification of Stroke Gaits by Deep Neural Networks Employing Inertial Measurement Units. Sensors 2021, 21, 1864. https://doi.org/10.3390/s21051864.

[55] Darbandi, H.; Baniasad, M.; Baghdadi, S.; Khandan, A.; Vafaee, A.; Farahmand, F. Automatic Classification of Gait Patterns in Children with Cerebral Palsy Using Fuzzy Clustering Method. Clin.

Biomech. 2020, 73, 189–194. https://doi.org/10.1016/j.clinbiomech.2019.12.031.

[56] Saleem, F.; Khan, M. A.; Alhaisoni, M.; Tariq, U.; Armghan, A.; Alenezi, F.; Choi, J.-I.; Kadry, S. Human Gait Recognition: A Single Stream Optimal Deep Learning Features Fusion. Sensors 2021, 21, 7584. https://doi.org/10.3390/s21227584.

[57] Trabassi, D.; Serrao, M.; Varrecchia, T.; Ranavolo, A.; Coppola, G.; De Icco, R.; Tassorelli, C.; Castiglia, S. F. Machine Learning Approach to Support the Detection of Parkinson's Disease in IMU-Based Gait Analysis. Sensors 2022, 22, 3700. https://doi.org/10.3390/s22103700.

[58] Rajpurkar, P., Chen, E., Banerjee, O. et al. AI in health and medicine. Nat Med 28, 31–38 (2022). https://doi.org/10.1038/s41591-021-01614-0.

[59] Piccialli, F., Di Somma, V., Giampaolo, F., Cuomo, S. and Fortino, G., 2021. A survey on deep learning in medicine: Why, how and when?. Information Fusion, 66, pp.111-137.

[60] McCalmont, G., Morrow, P., Zheng, H., Samara, A., Yasaei, S., Wang, H. and McClean, S., 2018, December. eZiGait: toward an AI gait analysis and sssistant system. In 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 2280-2286). IEEE.

[61] Christopher M. Bishop. 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag, Berlin, Heidelberg.

[62] Goodfellow, I.; Bengio, Y. & Courville, A. (2016), Deep Learning , MIT Press.

[63] Giuseppe Bonaccorso. 2017. Machine Learning Algorithms: A reference guide to popular algorithms for data science and machine learning. Packt Publishing.

[64] Alpaydin, E. (2014), Introduction to Machine Learning , MIT Press , Cambridge, MA.

[65] Tan, P.-N.; Steinbach, M. & Kumar, V. (2005), Introduction to Data Mining , Addison Wesley .

[66] Yoram Reich, S.V. Barai, Evaluating machine learning models for engineering problems, Artificial Intelligence in Engineering, Volume 13, Issue 3, 1999, Pages 257-272, ISSN 0954-1810.

[67] Majnik, M. and Bosnić, Z., 2013. ROC analysis of classifiers in machine learning: A survey. Intelligent data analysis, 17(3), pp.531-558.

[68] Susmaga, R., 2004. Confusion matrix visualization. In Intelligent information processing and web mining (pp. 107-116). Springer, Berlin, Heidelberg.

[69] Chicco, D., Warrens, M.J. and Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science, 7, p.e623.

[70] Berrar, D., 2019. Cross-Validation.

[71] Yadav, S. and Shukla, S., 2016, February. Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In 2016 IEEE 6th International conference on advanced computing (IACC) (pp. 78-83). IEEE.

[72] Celebi, M.E. and Aydin, K. eds., 2016. Unsupervised learning algorithms. Berlin: Springer International Publishing.

[73] Diday, E. and Simon, J.C., 1976. Clustering analysis. In Digital pattern recognition (pp. 47-94). Springer, Berlin, Heidelberg.

[74] Hand, D. J.; Mannila, H. & Smyth, P. (2001), Principles of Data Mining , MIT Press.

[75] Pham, D.T., Dimov, S.S. and Nguyen, C.D., 2005. Selection of K in K-means clustering. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 219(1), pp.103-119.

[76] Adam Rozumalski, Michael H. Schwartz, Crouch gait patterns defined using k-means cluster analysis are related to underlying clinical pathology, Gait & Posture, Volume 30, Issue 2, 2009, Pages 155-160, ISSN 0966-6362.

[77] Brunton, S. L. and Kutz, J. N. (2019) Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control. Cambridge: Cambridge University Press. doi: 10.1017/9781108380690.

[78] Böhm H, Oestreich C, Rethwilm R, Federolf P, Döderlein L, Fujak A, Dussa CU. Cluster analysis to identify foot motion patterns in children with flexible flatfeet using gait analysis-A statistical approach to detect decompensated pathology? Gait Posture. 2019 Jun;71:151-156. doi: 10.1016/j.gaitpost.2019.04.028. Epub 2019 Apr 26. PMID: 31071537.

[79] Van Engelen, J.E. and Hoos, H.H., 2020. A survey on semi-supervised learning. Machine Learning, 109(2), pp.373-440.

[80] Singh, G., Chowdhary, M., Kumar, A. and Bahl, R., 2020. A personalized classifier for human motion activities with semi-supervised learning. IEEE Transactions on Consumer Electronics, 66(4), pp.346-355.

[81] Jaderberg, M., Mnih, V., Czarnecki, W.M., Schaul, T., Leibo, J.Z., Silver, D. and Kavukcuoglu, K., 2016. Reinforcement learning with unsupervised auxiliary tasks. arXiv preprint arXiv:1611.05397.

[82] Rabbi, M., Ali, S., Choudhury, T., and Berke, E. (2011). Passive and In-Situ Assessment of Mental and Physical Well-Being Using Mobile Sensors.

[83] Zou, Y., Libanori, A., Xu, J., Nashalian, A., and Chen, J. (2020). Triboelectric Nanogenerator Enabled Smart Shoes for Wearable Electricity Generation. Research 2020, 7158953. doi:10.34133/2020/7158953.

[84] Zhang, Z., He, T., Zhu, M., Sun, Z., Shi, Q., Zhu, J., et al. (2020f). Deep Learning-Enabled Triboelectric Smart Socks for IoT-Based Gait Analysis and VR Applications. npj Flexible Electronics 4, 1–12. doi:10.1038/s41528- 020-00092-7.

[85] Totaro, M., Poliero, T., Mondini, A., Lucarotti, C., Cairoli, G., Ortiz, J., et al. (2017). Soft Smart Garments for Lower Limb Joint Position Analysis. Sensors (Basel) 17, 2314. doi:10.3390/s17102314.

[86] Low, J.-H., Chee, P.-S., Lim, E.-H., and Ganesan, V. (2020). Design of a Wireless Smart Insole Using Stretchable Microfluidic Sensor for Gait Monitoring. Smart Mater. Struct. 29, 065003. doi:10.1088/1361-665x/ab802c.

[87] Poniszewska-Maranda, A., Kaczmarek, D., Kryvinska, N., and Xhafa, F. (2019). Studying Usability of AI in the IoT Systems/paradigm through Embedding NN Techniques into mobile Smart Service System. Computing 101, 1661–1685. doi:10.1007/s00607-018-0680-z.

[88] San-Segundo, R., Blunck, H., Moreno-Pimentel, J., Stisen, A., and Gil-Martín, M. (2018). Robust Human Activity Recognition Using Smartwatches and Smartphones. Eng. Appl. Artif. Intelligence 72, 190–202. doi:10.1016/j.engappai.2018.04.002.

[89] Ghose, A., Sinha, P., Bhaumik, C., Sinha, A., Agrawal, A., and Choudhury, A. D. (2013). "UbiHeld: Ubiquitous Healthcare Monitoring System for Elderly and Chronic Patients," in Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication (Zurich, Switzerland: Association for Computing Machinery).

[90] Yang, J., and Yin, Y. (2021). Novel Soft Smart Shoes for Motion Intent Learning of Lower Limbs Using LSTM with a Convolutional Autoencoder. IEEE Sensors J. 21, 1906–1917. doi:10.1109/jsen.2020.3019053.

[91] Xu, W., Xue, W., Lin, Q., Lan, G., Feng, X., Wei, B., et al. (2021). PrivGait: An Energy Harvesting-Based Privacy-Preserving User Identification System by Gait Analysis. IEEE Internet Things J. 99, 1. doi:10.1109/jiot.2021.3089618.

[92] Lan, G., Ma, D., Xu, W., Hassan, M., and Hu, W. (2020). Capacitor-based Activity Sensing for Kinetic-Powered Wearable IoTs. ACM Trans. Internet Things 1, 2. doi:10.1145/3362124.

[93] Xu, W., Lan, G., Lin, Q., Khalifa, S., Bergmann, N., Hassan, M., et al. (2017). "KEH-gait: Towards a Mobile Healthcare User Authentication System by Kinetic Energy Harvesting," in The Network and Distributed System Security Symposium, 4.

[94] Sandhu, M. M., Khalifa, S., Geissdoerfer, K., Jurdak, R., and Portmann, M. (2021). "SolAR: Energy Positive Human Activity Recognition Using Solar Cells," in 2021 IEEE International Conference on Pervasive Computing and Communications (PerCom), 1–10.

[95] Zhang, Z. X., He, T. Y. Y., Zhu, M. L., Sun, Z. D., Shi, Q. F., Zhu, J. X., et al. (2020g). Deep Learning-Enabled Triboelectric Smart Socks for IoT-Based Gait Analysis and VR Applications. Npj Flexible Electronics 4, 29. doi:10.1038/s41528-020-00092-7.

[96] Gao, C., Gehlhar, R., Ames, A. D., Liu, S. C., and Delbruck, T. (2020). "Recurrent Neural Network Control of a Hybrid Dynamical Transfemoral Prosthesis with EdgeDRNN Accelerator," in 2020 IEEE International Conference on Robotics and Automation (ICRA), 5460–5466.

[97] Zhang, Y., Li, S., Nolan, K. J., and Zanotto, D. (2020e). "Reinforcement Learning Assist-As-Needed Control for Robot Assisted Gait Training," in 2020

8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob), 785–790.

[98] Llorente-Vidrio, D., Pérez-San Lázaro, R., Ballesteros, M., Salgado, I., Cruz-Ortiz, D., and Chairez, I. (2020). Event Driven Sliding Mode Control of a Lower Limb Exoskeleton Based on a Continuous Neural Network Electromyographic Signal Classifier. Mechatronics 72, 102451. doi:10.1016/j.mechatronics.2020.102451.

[99] Cook, D. J., Schmitter-Edgecombe, M., and Dawadi, P. (2015). Analyzing Activity Behavior and Movement in a Naturalistic Environment Using Smart Home Techniques. IEEE J. Biomed. Health Inform. 19, 1882–1892. doi:10.1109/jbhi.2015.2461659.

[100] Borelli, E., Paolini, G., Antoniazzi, F., Barbiroli, M., Benassi, F., Chesani, F., et al. (2019). HABITAT: An IoT Solution for Independent Elderly. Sensors (Basel) 19, 1258. doi:10.3390/s19051258.

[101] Hsu, Y. L., Chou, P. H., Chang, H. C., Lin, S. L., Yang, S. C., Su, H. Y., et al. (2017). Design and Implementation of a Smart Home System Using Multisensor Data Fusion Technology. Sensors (Basel) 17, 1631. doi:10.3390/s17071631.

[102] Menegoni, F., Albani, G., Bigoni, M., Priano, L., Trotti, C., Galli, M. and Mauro, A., 2009. Walking in an immersive virtual reality. Annual Review of Cybertherapy and Telemedicine 2009, pp.72-76.

[103] Resnik, L., Etter, K., Klinger, S.L. and Kambe, C., 2011. Using virtual reality environment to facilitate training with advanced upper-limb prosthesis. Journal of Rehabilitation Research and Development, 48(6).

[104] Feigl, T., Gruner, L., Mutschler, C., and Roth, D. (2020). "Real-Time Gait Reconstruction for Virtual Reality Using a Single Sensor," in 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), 84–89.

[105] Bergamin, K., Clavet, S., Holden, D., and Forbes, J. R. (2019). DReCon: Data-Driven Responsive Control of Physics-Based Characters. ACM Trans. Graph. 38, 206. doi:10.1145/3355089.3356536.

[106] Peng, X. B., Kanazawa, A., Malik, J., Abbeel, P., and Levine, S. (2018b). SFV: Reinforcement Learning of Physical Skills from Videos. ACM Trans. Graph. 37, 178. doi:10.1145/3197517.3201311.

[107] Peng, X. B., Abbeel, P., Levine, S., and Panne, M. V. D. (2018a). DeepMimic: Example-Guided Deep Reinforcement Learning of Physics-Based Character Skills. ACM Trans. Graph. 37, 143. doi:10.1145/3197517.3201311.

[108] Huang, Y., Kaufmann, M., Aksan, E., Black, M. J., Hilliges, O., and Pons-Moll, G. (2018). Deep Inertial Poser: Learning to Reconstruct Human Pose from Sparse Inertial Measurements in Real Time. ACM Trans. Graph. 37, 185.

[109] Holden, D., Saito, J., and Komura, T. (2016). A Deep Learning Framework for Character Motion Synthesis and Editing. ACM Trans. Graph. 35, 138. doi:10.1145/2897824.2925975.

[110] Huang, P., Tejera, M., Collomosse, J., and Hilton, A. (2015). Hybrid Skeletal-Surface Motion Graphs for Character Animation from 4D Performance Capture. ACM Trans. Graph. 34, 17. doi:10.1145/2699643.

[111] Ding, M., and Fan, G. (2015). Multilayer Joint Gait-Pose Manifolds for Human Gait Motion Modeling. IEEE Trans. Cybern. 45, 2413–2424. doi:10.1109/tcyb.2014.2373393.

[112] Alvarez-Alvarez, A., Trivino, G., and Cordon, O. (2012). Human Gait Modeling Using a Genetic Fuzzy Finite State Machine. IEEE Trans. Fuzzy Syst. 20, 205–223. doi:10.1109/tfuzz.2011.2171973.

[113] Burdack, J., Horst, F., Giesselbach, S., Hassan, I., Daffner, S., and Schöllhorn, W. I. (2020). Systematic Comparison of the Influence of Different Data Preprocessing Methods on the Performance of Gait Classifications Using Machine Learning. Front. Bioeng. Biotechnol. 8, 260. doi:10.3389/fbioe.2020.00260.

[114] Gage, J. R., Deluca, P. A., and Renshaw, T. S. (1995). Gait Analysis. J. Bone Joint Surg. 77, 1607–1623. doi:10.2106/00004623-199510000-00017.

[115] Di Nardo, F., Morbidoni, C., Cucchiarelli, A., and Fioretti, S. (2020). Recognition of Gait Phases with a Single Knee Electrogoniometer: A Deep Learning Approach. Electronics 9, 355. doi:10.3390/electronics9020355.

[116] Jiang, X., Chu, K. H. T., Khoshnam, M., and Menon, C. (2018). A Wearable Gait Phase Detection System Based on Force Myography Techniques. Sensors (Basel) 18, 1279. doi:10.3390/s18041279.

[117] Vaith, A., Taetz, B., and Bleser, G. (2020). "Uncertainty Based Active Learning with Deep Neural Networks for Inertial Gait Analysis," in 2020 IEEE 23rd International Conference on Information Fusion (FUSION), 1–8. doi:10.23919/fusion45008.2020.9190449.

[118] Pérez-Ibarra, J. C., Siqueira, A. A., Terra, M. H., and Krebs, H. I. (2020). "Hybrid Simulated Annealing and Genetic Algorithm for Optimization of a Rule-Based Algorithm for Detection of Gait Events in Impaired Subjects," in 2020 IEEE/ ASME International Conference on Advanced Intelligent Mechatronics (AIM) (Boston, MA, USA: IEEE), 1167–1171.

[119] Morbidoni, C., Cucchiarelli, A., Fioretti, S., and Di Nardo, F. (2019). A Deep Learning Approach to EMG-Based Classification of Gait Phases during Level Ground Walking. Electronics 8, 894. doi:10.3390/electronics8080894.

[120] Farah, J. D., Baddour, N., and Lemaire, E. D. (2017). "Gait Phase Detection from Thigh Kinematics Using Machine Learning Techniques," in 2017 IEEE International Symposium on Medical Measurements and Applications (MeMeA), 263–268.

[121] Mannini, A., and Sabatini, A. M (2012). Gait Phase Detection and Discrimination Between Walking–Jogging Activities Using Hidden Markov Models Applied to Foot Motion Data From a Gyroscope. Gait Postur 36, 657–661.

[122] Russell, B., Mcdaid, A., Toscano, W., and Hume, P. (2021). Moving the Lab into the Mountains: A Pilot Study of Human Activity Recognition in Unstructured Environments. Sensors (Basel) 21, 654. doi:10.3390/s21020654.

[123] Baghdadi, A., Cavuoto, L. A., Jones-Farmer, A., Rigdon, S. E., Esfahani, E. T., and Megahed, F. M. (2021). Monitoring Worker Fatigue Using Wearable Devices: A Case Study to Detect Changes in Gait Parameters. J. Qual. Technol. 53, 47–71. doi:10.1080/00224065.2019.1640097.

[124] Sedighi Maman, Z., Chen, Y.-J., Baghdadi, A., Lombardo, S., Cavuoto, L. A., and Megahed, F. M. (2020). A Data Analytic Framework for Physical Fatigue Management Using Wearable Sensors. Expert Syst. Appl. 155, 113405. doi:10.1016/j.eswa.2020.113405.

[125] Karvekar, S. B. (2019). Smartphone-based Human Fatigue Detection in an Industrial Environment Using Gait Analysis.

[126] Baghdadi, A., Megahed, F. M., Esfahani, E. T., and Cavuoto, L. A. (2018). A Machine Learning Approach to Detect Changes in Gait Parameters Following a Fatiguing Occupational Task. Ergonomics 61, 1116–1129. doi:10.1080/00140139.2018.1442936.

[127] Janssen, D., Schöllhorn, W. I., Newell, K. M., Jäger, J. M., Rost, F., and Vehof, K. (2011). Diagnosing Fatigue in Gait Patterns by Support Vector Machines and Self-Organizing Maps. Hum. Move. Sci. 30, 966–975. doi:10.1016/j.humov.2010.08.010.

[128] Lu, H., Shao, W., Ngai, E., Hu, X., and Hu, B. (2020). "A New Skeletal Representation Based on Gait for Depression Detection," in 2020 IEEE International Conference on E-health Networking, Application and Services (HEALTHCOM), 1–6.

[129] Khodabandehloo, E., Riboni, D., and Alimohammadi, A. (2021). HealthXAI: Collaborative and Explainable AI for Supporting Early Diagnosis of Cognitive Decline. Future Generation Computer Syst. 116, 168–189. doi:10.1016/j.future.2020.10.030.

[130] Iosa, M., Capodaglio, E., Pelà, S., Persechino, B., Morone, G., Antonucci, G., et al. (2021). Artificial Neural Network Analyzing Wearable Device Gait Data for Identifying Patients with Stroke Unable to Return to Work. Front Neurol. 12, 650542. doi:10.3389/fneur.2021.650542.

[131] Zhu, Y., Xu, W., Luo, G., Wang, H., Yang, J., and Lu, W. (2020). Random Forest Enhancement Using Improved Artificial Fish Swarm for the Medial Knee Contact Force Prediction. Artif. intelligence Med. 103, 101811. doi:10.1016/j.artmed.2020.101811.

[132] Zhou, Y., Romijnders, R., Hansen, C., Campen, J. V., Maetzler, W., Hortobágyi, T., et al. (2020). The Detection of Age Groups by Dynamic Gait Outcomes Using Machine Learning Approaches. Sci. Rep. 10, 4426. doi:10.1038/s41598-020-61423-2.

[133] Zeng, W., Ismail, S. A., and Pappas, E. (2020). Detecting the Presence of Anterior Cruciate Ligament Injury Based on Gait Dynamics Disparity and Neural Networks. Artif. Intell. Rev. 53, 3153–3176. doi:10.1007/s10462-019-09758-9.

[134] Pepa, L., Capecci, M., Andrenelli, E., Ciabattoni, L., Spalazzi, L., and Ceravolo, M. G. (2020). A Fuzzy Logic System for the home Assessment of Freezing of Gait in Subjects with Parkinsons Disease. Expert Syst. Appl. 147, 113197. doi:10.1016/j.eswa.2020.113197.

[135] Kaur, R., Chen, Z., Motl, R., Hernandez, M. E., and Sowers, R. (2020). Predicting Multiple Sclerosis from Gait Dynamics Using an Instrumented Treadmill–A Machine Learning Approach. IEEE Trans. Biomed. Eng. 68, 2666. doi:10.1109/TBME.2020.3048142.

[136] Guo, Y., Deligianni, F., Gu, X., and Yang, G.-Z. (2019). 3-D Canonical Pose Estimation and Abnormal Gait Recognition with a Single RGB-D Camera. IEEE Robot. Autom. Lett. 4, 3617–3624. doi:10.1109/lra.2019.2928775.

[137] Sato, K., Nagashima, Y., Mano, T., Iwata, A., and Toda, T. (2019). Quantifying normal and Parkinsonian Gait Features from home Movies: Practical Application of a Deep Learning-Based 2D Pose Estimator. PloS one 14, e0223549. doi:10.1371/journal.pone.0223549.

[138] Fang, J., Wang, T., Li, C., Hu, X., Ngai, E., Seet, B.-C., et al. (2019). Depression Prevalence in Postgraduate Students and its Association with Gait Abnormality. IEEE Access 7, 174425–174437. doi:10.1109/access.2019.2957179.

[139] Acosta-Escalante, F. D., Beltran-Naturi, E., Boll, M. C., Hernandez Nolasco, J. A., and Pancardo Garcia, P. (2018). Meta-Classifiers in Huntington's Disease Patients Classification, Using iPhone's Movement Sensors Placed at the Ankles. IEEE Access 6, 30942–30957. doi:10.1109/access.2018.2840327.

[140] Pulido-Valdeolivas, I., Gómez-Andrés, D., Martín-Gonzalo, J. A., Rodríguez-Andonaegui, I., López-López, J., Pascual-Pascual, S. I., et al. (2018). Gait Phenotypes in Paediatric Hereditary Spastic Paraplegia Revealed by Dynamic Time Warping Analysis and Random Forests. PLoS One 13, e0192345. doi:10.1371/journal.pone.0192345.

[141] Wan, S., Liang, Y., Zhang, Y., and Guizani, M. (2018b). Deep Multi-Layer Perceptron Classifier for Behavior Analysis to Estimate Parkinson's Disease Severity Using Smartphones. IEEE Access 6, 36825–36833. doi:10.1109/access.2018.2851382.

[142] Arifoglu, D., and Bouchachia, A. (2017). Activity Recognition and Abnormal Behaviour Detection with Recurrent Neural Networks. Proced. Computer Sci. 110, 86–93. doi:10.1016/j.procs.2017.06.121.

[143] Dolatabadi, E., Taati, B., and Mihailidis, A. (2017). An Automated Classification of Pathological Gait Using Unobtrusive Sensing Technology. IEEE Trans. Neural Syst. Rehabil. Eng. 25, 2336–2346. doi:10.1109/tnsre.2017.2736939.

[144] Shetty, S., and Rao, Y. S. (2016). "SVM Based Machine Learning Approach to Identify Parkinson's Disease Using Gait Analysis," in 2016 International Conference on Inventive Computation Technologies (ICICT), 1–5.

[145] Taborri, J., Molinaro, L., Santospagnuolo, A., Vetrano, M., Vulpiani, M. C., and Rossi, S. (2021). A Machine-Learning Approach to Measure the Anterior Cruciate Ligament Injury Risk in Female Basketball Players. Sensors (Basel) 21, 3141. doi:10.3390/s21093141.

[146] Johnson, W. R., Mian, A., Robinson, M. A., Verheul, J., Lloyd, D. G., and Alderson, J. A. (2021). Multidimensional Ground Reaction Forces and Moments from Wearable Sensor Accelerations via Deep Learning. IEEE Trans. Biomed. Eng. 68, 289–297. doi:10.1109/tbme.2020.3006158.

[147] Nguyen, M. D., Mun, K.-R., Jung, D., Han, J., Park, M., Kim, J., et al. (2020). "IMU-based Spectrogram Approach with Deep Convolutional Neural Networks for Gait Classification," in 2020 IEEE International Conference on Consumer Electronics (ICCE) (Las Vegas, NV, USA: IEEE), 1–6. doi:10.1109/icce46568.2020.9042999.

[148] Guo, Y., and Wang, X. (2021). Applying TS-DBN Model into Sports Behavior Recognition with Deep Learning Approach. J. Supercomput 77, 12192 12208. doi:10.1007/s11227-021-03772-x.

[149] Gholami, M., Napier, C., and Menon, C. (2020). Estimating Lower Extremity Running Gait Kinematics with a Single Accelerometer: A Deep Learning Approach. Sensors 20, 2939. doi:10.3390/s20102939.

[150] Kang, X., Huang, B., and Qi, G. (2018). A Novel Walking Detection and Step Counting Algorithm Using Unconstrained Smartphones. Sensors (Basel) 18, 297. doi:10.3390/s18010297.

[151] Onodera, A. N., Gavião Neto, W. P., Roveri, M. I., Oliveira, W. R., and Sacco, I. C. (2017). Immediate Effects of EVA Midsole Resilience and Upper Shoe Structure on Running Biomechanics: a Machine Learning Approach. PeerJ 5, e3026. doi:10.7717/peerj.3026.

[152] Lin, F., Wu, Y., Zhuang, Y., Long, X. and Xu, W., 2016. Human gender classification: a review. International Journal of Biometrics, 8(3-4), pp.275-300.

[153] Kwon, B., and Lee, S. (2021). Joint Swing Energy for Skeleton-Based Gender Classification. IEEE Access 9, 28334–28348. doi:10.1109/access.2021.3058745.

[154] Jain, A., and Kanhangad, V. (2018). Gender Classification in Smartphones Using Gait Information. Expert Syst. Appl. 93, 257–266. doi:10.1016/j.eswa.2017.10.017.

[155] Lee, L. and Grimson, W.E.L., 2002, May. Gait analysis for recognition and classification. In Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition (pp. 155-162). IEEE.

[156] Deng, Q., Xu, Y., Wang, J. and Sun, K., 2015, November. Deep learning for gender recognition. In 2015 International Conference on Computers, Communications, and Systems (ICCCS) (pp. 206-209). IEEE.

[157] Harris, E.J., Khoo, I.H. and Demircan, E., 2022. A survey of human gait-based artificial intelligence applications. Frontiers in Robotics and AI, 8, p.749274.

[158] Castro, F. M., Marín-Jiménez, M. J., Guil, N., and De La Blanca, N. P. (2017). "Automatic Learning of Gait Signatures for People Identification," in International Work-Conference on Artificial Neural Networks (Berlin, Germany: Springer), 257–270.

[159] Zhang, S., Wang, Y., and Li, A. (2019b). "Gait-Based Age Estimation with Deep Convolutional Neural Network," in 2019 International Conference on Biometrics (ICB), 1–8. doi:10.1109/icb45273.2019.8987240.

[160] Kwon, B., and Lee, S. (2021). Joint Swing Energy for Skeleton-Based Gender Classification. IEEE Access 9, 28334–28348. doi:10.1109/access.2021.3058745.

[161] Lu, J., Wang, G., and Moulin, P. (2014). Human Identity and Gender Recognition from Gait Sequences with Arbitrary Walking Directions. IEEE Trans.Inform.Forensic Secur. 9, 51–61. doi:10.1109/tifs.2013.2291969.

[162] Khera, P. and Kumar, N., 2020. Role of machine learning in gait analysis: a review. Journal of Medical Engineering and Technology, 44(8), pp.441-467.

[163] Kidziński Ł, Delp S, Schwartz M. Automatic real-time gait event detection in children using deep neural networks. PLoS One. 2019 Jan 31;14(1):e0211466. doi: 10.1371/journal.pone.0211466. PMID: 30703141; PMCID: PMC6354999.

[164] W. Kim and Y. Kim, "Abnormal Gait Recognition based on Integrated Gait Features in Machine Learning," 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), Madrid, Spain, 2021, pp. 1683-1688, doi: 10.1109/COMPSAC51774.2021.00251.

[165] I. Tien, S. D. Glaser and M. J. Aminoff, "Characterization of gait abnormalities in Parkinson's disease using a wireless inertial sensor system," 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina, 2010, pp. 3353-3356, doi: 10.1109/IEMBS.2010.5627904.

[166] Nukala, B.T., Shibuya, N., Rodriguez, A., Tsay, J., Lopez, J., Nguyen, T., Zupancic, S. and Lie, D.Y.C., 2015. An efficient and robust fall detection system using wireless gait analysis sensor with artificial neural network (ANN) and support vector machine (SVM) algorithms. Open journal of applied biosensor, 3(04), p.29.

[167] N. Shibuya et al., "A real-time fall detection system using a wearable gait analysis sensor and a Support Vector Machine (SVM) classifier," 2015 Eighth International Conference on Mobile Computing and Ubiquitous Networking (ICMU), Hakodate, Japan, 2015, pp. 66-67, doi: 10.1109/ICMU.2015.7061032.

[168] Huang, J., Di Troia, F. and Stamp, M., 2018, January. Acoustic Gait Analysis using Support Vector Machines. In ICISSP (pp. 545-552).

[169] Hayashi, H., Toribatake, Y., Murakami, H., Yoneyama, T., Watanabe, T. and Tsuchiya, H., 2015. Gait analysis using a support vector machine for lumbar spinal stenosis. Orthopedics, 38(11), pp.e959-e964.

[170] Yoo, J.H., Hwang, D. and Nixon, M.S., 2005, September. Gender classification in human gait using support vector machine. In International Conference on Advanced Concepts for Intelligent Vision Systems (pp. 138-145). Springer, Berlin, Heidelberg.

[171] Begg, R. and Kamruzzaman, J., 2003, October. A comparison of neural networks and support vector machines for recognizing young-old gait patterns.

In TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region (Vol. 1, pp. 354-358). IEEE.

[172] Si, W., Yang, G., Chen, X. and Jia, J., 2019. Gait identification using fractal analysis and support vector machine. Soft Computing, 23(19), pp.9287-9297.

[173] Kamruzzaman, J. and Begg, R.K., 2006. Support vector machines and other pattern recognition approaches to the diagnosis of cerebral palsy gait. IEEE Transactions on Biomedical Engineering, 53(12), pp.2479-2490.

[174] Horst, F., Kramer, F., Schäfer, B., Eekhoff, A., Hegen, P., Nigg, B.M. and Schöllhorn, W.I., 2016. Daily changes of individual gait patterns identified by means of support vector machines. Gait and posture, 49, pp.309-314.

[175] Deutsch, A. and Dormann, S., 2005. Mathematical modeling of biological pattern formation (pp. 45-56). Birkhäuser Boston.

[176] Beek, P.J., Peper, C.E. and Stegeman, D.F., 1995. Dynamical models of movement coordination. Human movement science, 14(4-5), pp.573-608.

[177] van Emmerik, R.E. and van Wegen, E.E., 2000. On variability and stability in human movement. Journal of Applied Biomechanics, 16(4), pp.394-406.

[178] Taga, G., 1995. A model of the neuro-musculo-skeletal system for human locomotion: I. Emergence of basic gait. Biological cybernetics, 73(2), pp.97-111.

[179] Dingwell, J.B. and Cusumano, J.P., 2000. Nonlinear time series analysis of normal and pathological human walking. Chaos: An Interdisciplinary Journal of Nonlinear Science, 10(4), pp.848-863.

[180] Piorek, M., Josinski, H., Michalczuk, A., Switonski, A. and Szczesna, A., 2017. Quaternions and joint angles in an analysis of local stability of gait for different variants of walking speed and treadmill slope. Information Sciences, 384, pp.263-280.

[181] Maxwell, A.E., Warner, T.A. and Fang, F., 2018. Implementation of machine-learning classification in remote sensing: An applied review. International Journal of Remote Sensing, 39(9), pp.2784-2817.

[182] Higginson, B.K., 2009. Methods of running gait analysis. Current sports medicine reports, 8(3), pp.136-141.

[183] Hu, L.Y., Huang, M.W., Ke, S.W. and Tsai, C.F., 2016. The distance function effect on k-nearest Neighbour classification for medical datasets. SpringerPlus, 5(1), pp.1-9.

[184] Cristianini, N. and Scholkopf, B., 2002. Support vector machines and kernel methods: the new generation of learning machines. Ai Magazine, 23(3), pp.31-31.

[185] Khera, P. and Kumar, N., 2020. Role of machine learning in gait analysis: a review. Journal of Medical Engineering and Technology, 44(8), pp.441-467.

[186] Stitson, M.O., Weston, J.A.E., Gammerman, A., Vovk, V. and Vapnik, V., 1996. Theory of support vector machines. University of London, 117(827), pp.188-191.

[187] Zhou, Y., Romijnders, R., Hansen, C., Campen, J.V., Maetzler, W., Hortobágyi, T. and Lamoth, C.J., 2020. The detection of age groups by dynamic gait outcomes using machine learning approaches. Scientific reports, 10(1), p.4426.

[188] P. Patil, K. S. Kumar, N. Gaud and V. B. Semwal, "Clinical Human Gait Classification: Extreme Learning Machine Approach," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), 2019, pp. 1-6, doi: 10.1109/ICASERT.2019.8934463.

[189] M. W. Rahman and M. L. Gavrilova, "Kinect gait skeletal joint feature-based person identification," 2017 IEEE 16th International Conference on Cognitive Informatics and Cognitive Computing (ICCI*CC), 2017, pp. 423-430, doi: 10.1109/ICCI-CC.2017.8109783.

[190] J. P. Félix et al., "A Parkinson's Disease Classification Method: An Approach Using Gait Dynamics and Detrended Fluctuation Analysis," 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), 2019, pp. 1-4, doi: 10.1109/CCECE.2019.8861759.

[191] P. Tahafchi and J. W. Judy, "Freezing-of-Gait Detection Using Wearable Sensor Technology and Possibilistic K-Nearest-Neighbour Algorithm," 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 4246-4249, doi: 10.1109/EMBC.2019.8856480.

[192] H. Huang, X. Li and Y. Sun, "A triboelectric motion sensor in wearable body sensor network for human activity recognition," 2016 38th Annual International

Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2016, pp. 4889-4892, doi: 10.1109/EMBC.2016.7591823.

[193] Y. Lin and J. Le Kernec, "Performance Analysis of Classification Algorithms for Activity Recognition Using Micro-Doppler Feature," 2017 13th International Conference on Computational Intelligence and Security (CIS), 2017, pp. 480-483, doi: 10.1109/CIS.2017.00111.

[194] P. Kasebzadeh, K. Radnosrati, G. Hendeby and F. Gustafsson, "Joint Pedestrian Motion State and Device Pose Classification," in IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 8, pp. 5862-5874, Aug. 2020, doi: 10.1109/TIM.2019.2958005.

[195] Z. Liu, W. Lin, Y. Geng and P. Yang, "Intent pattern recognition of lower-limb motion based on mechanical sensors," in IEEE/CAA Journal of Automatica Sinica, vol. 4, no. 4, pp. 651-660, 2017, doi: 10.1109/JAS.2017.7510619.

[196] W. Kim and Y. Kim, "Abnormal Gait Recognition based on Integrated Gait Features in Machine Learning," 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), 2021, pp. 1683-1688, doi: 10.1109/COMPSAC51774.2021.00251.

[197] H. Li, A. Mehul, J. Le Kernec, S. Z. Gurbuz and F. Fioranelli, "Sequential Human Gait Classification With Distributed Radar Sensor Fusion," in IEEE Sensors Journal, vol. 21, no. 6, pp. 7590-7603, 15 March15, 2021, doi: 10.1109/JSEN.2020.3046991.

[198] D. Guffanti, A. Brunete and M. Hernando, "Non-Invasive Multi-Camera Gait Analysis System and its Application to Gender Classification," in IEEE Access, vol. 8, pp. 95734-95746, 2020, doi: 10.1109/ACCESS.2020.2995474.

[199] P. Zou et al., "Wearable Iontronic FMG for Classification of Muscular Locomotion," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 7, pp. 2854-2863, July 2022, doi: 10.1109/JBHI.2022.3173968.

[200] The Math Works, Inc. MATLAB. Version R2020b, The Math Works, Inc., 2020. Computer Software. http://www.mathworks.com/.

[201] Shawe-Taylor, J. and Cristianini, N. (2004) Kernel Methods for Pattern Analysis. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511809682.

[202] Ebersbach G, Sojer M, Müller J, Heijmenberg M, Poewe W. Sociocultural differences in gait. Mov Disord. 2000 Nov;15(6):1145-7. doi: 10.1002/1531-8257(200011)15:6<1145::aid-mds1013>3.0.co;2-c. PMID: 11104198.

[203] Mahlknecht P, Kiechl S, Bloem BR, Willeit J, Scherfler C, Gasperi A, Rungger G, Poewe W, Seppi K. Prevalence and burden of gait disorders in elderly men and women aged 60-97 years: a population-based study. PLoS One. 2013 Jul 24;8(7):e69627. doi: 10.1371/journal.pone.0069627. PMID: 23894511; PMCID: PMC3722115.

[204] Richardson K, Bennett K, Kenny RA. Polypharmacy including falls risk-increasing medications and subsequent falls in community-dwelling middle-aged and older adults. Age Ageing. 2015 Jan;44(1):90-6. doi: 10.1093/ageing/afu141. Epub 2014 Oct 12. PMID: 25313240.

[205] Uranga RM, Keller JN. The Complex Interactions Between Obesity, Metabolism and the Brain. Front Neurosci. 2019 May 24;13:513. doi: 10.3389/fnins.2019.00513. PMID: 31178685; PMCID: PMC6542999.

[206] Esser, P.; Dawes, H.; Collett, J.; Howells, K. IMU: Inertial Sensing of Vertical CoM Movement. J. Biomech. 2009, I, 1578–1581. https://doi.org/10.1016/j.jbiomech.2009.03.049.

[207] Esser, P.; Dawes, H.; Collett, J.; Howells, K. Insights into Gait Disorders: Walking Variability Using Phase Plot Analysis, Parkinson's Disease. Gait Posture 2013, 38, 648–652. https://doi.org/10.1016/j.gaitpost.2013.02.016.

[208] olde Scheper, T. V. Biologically Inspired Rate Control of Chaos. Chaos: An Interdiscip. J. Nonlinear Sci. 2017, 27, 103122. https://doi.org/10.1063/1.5008892.

[209] Sudarsky L. Gait disorders: prevalence, morbidity, and etiology. Adv Neurol. 2001;87:111-7. PMID: 11347214.

[210] P. B. Shelke and P. R. Deshmukh, "Gait Based Gender Identification Approach," 2015 Fifth International Conference on Advanced Computing and Communication Technologies, 2015, pp. 121-124, doi: 10.1109/ACCT.2015.66.

[211] Osoba MY, Rao AK, Agrawal SK, Lalwani AK. Balance and gait in the elderly: A contemporary review. Laryngoscope Investig Otolaryngol. 2019 Feb 4;4(1):143-153. doi: 10.1002/lio2.252. PMID: 30828632; PMCID: PMC6383322.

[212] Esser, P., Dawes, H., Collett, J., Feltham, M. G. and Howells, K. Assessment of spatio-temporal gait parameters using inertial measurement units in neurological populations. Gait Posture 34, 558–560. https://doi.org/10.3791/51878 (2011).

[213] Zijlstra, W. and Hof, A. L. Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. Gait Posture 18, 1-10 (2003).

[214] Gonzalez, R. C., Alvarez, D., Lopez, A. M. and Alvarez, J. C. Modified pendulum model for mean step length estimation. Conf Proc IEEE Eng Med Biol Soc, 1371-1374 (2007).

[215] Han, T. R., Paik, N. J. and Im, M. S. Quantification of the path of center of pressure (COP) using an F-scan in-shoe transducer. Gait and Posture 10, 248-254, doi:S0966636299000405 [pii] (1999).

[216] Schmid, M., Beltrami, G., Zambarbieri, D. and Verni, G. Centre of pressure displacements in trans-femoral amputees during gait. Gait and Posture 21, 255-262 (2005).

[217] Ghani F, Rachele JN, Washington S, Turrell G. Gender and age differences in walking for transport and recreation: Are the relationships the same in all Neighbour hoods? Prev Med Rep. 2016 May 16;4:75-80. doi: 10.1016/j.pmedr.2016.05.001. PMID: 27413664; PMCID: PMC4929074.

[218] Sprott, J.C. and Hoover, W.G., 2017. Harmonic oscillators with nonlinear damping. International Journal of Bifurcation and Chaos, 27(11), p.1730037.

[219] Acebrón, J.A., Bonilla, L.L., Vicente, C.J.P., Ritort, F. and Spigler, R., 2005. The Kuramoto model: A simple paradigm for synchronization phenomena. Reviews of modern physics, 77(1), p.137.

[220] Hong, H. and Strogatz, S.H., 2011. Conformists and contrarians in a Kuramoto model with identical natural frequencies. Physical Review E, 84(4), p.046202.

[221] Kalay, Z., 2007. A review on synchronization through the Kuramoto model. Physics, 87131(2), pp.1-8.

[222] Maistrenko, Y.L., Popovych, O.V. and Tass, P.A., 2005. Desynchronization and chaos in the Kuramoto model. In Dynamics of coupled map lattices and of related spatially extended systems (pp. 285-306). Springer, Berlin, Heidelberg.

[223] Ahmedt-Aristizabal, D., Armin, M.A., Denman, S., Fookes, C. and Petersson, L., 2021. Graph-based deep learning for medical diagnosis and analysis: past, present and future. Sensors, 21(14), p.4758.