Objective Assessment of Neurological Conditions using Machine Learning

Siti Anizah Muhamed

PhD

University of York

Electronic Engineering

March 2019

ABSTRACT

Movement disorders are a subset of neurological conditions that are responsible for a significant decline in the health of the world's population, having multiple negative impacts on the lives of patients, their families, societies and countries' economy. Parkinson's disease (PD), the most common of all movement disorders, remains idiopathic (of unknown cause), is incurable, and without any confirmed pathological marker that can be extracted from living patients. As a degenerative condition, early and accurate diagnosis is critical for effective disease management in order to preserve a good quality of life. It also requires an in-depth understanding of clinical symptoms to differentiate the disease from other movement disorders. Unfortunately, clinical diagnosis of PD and other movement disorders is subject to the subjective interpretation of clinicians, resulting in a high rate of misdiagnosis of up to 25%. However, computerised methods can support clinical diagnosis through objective assessment. The major focus of this study is to investigate the use of machine learning approaches, specifically evolutionary algorithms, to diagnose, differentiate and characterise different movement disorders, namely PD, Huntington disease (HD) and Essential Tremor (ET). In the first study, movement features of three standard motor tasks from Unified Parkinson's Disease Rating Scale (UPDRS), finger tapping, hand opening-closing and hand pronation-supination, were used to evolve the high-performance classifiers. The results obtained for these conditions are encouraging, showing differences between the groups of healthy controls, PD, HD and ET patients. Findings on the most discriminating features of the best classifiers provide insight into different characteristics of the neurological disorders under consideration. The same algorithm has also been applied in the second study on Dystonia patients. A differential classification between Organic Dystonia and Functional Dystonia patients is less convincing, but positive enough to recommend future studies.

List of content

LIST OF TABLE

LIST OF FIGURES

ACKNOWLEDGEMENTS

In the name of Allah, most Gracious, most Compassionate. The One who continually showers all of creation with blessings and prosperity without any disparity. The One who is most kind, loving and merciful. All praise be to Allah, with your blessings and mercy I managed to complete this thesis.

I would like to show my sincere gratitude to Professor Stephen Smith for his supervision, guidance, encouragement, supports and literal helps throughout the years of research until the submission of this thesis. His enthusiasm and critical view on research has made a deep impression on me and I have learned so much from him. I am very grateful to my collaborator, Dr Rachel Newby, from Leeds Teaching Hospitals NHS Trust, UK for kindly sharing her data and her clinical expertise.

A special thanks to the person who without him I would not have finished this thesis, the one who supported me as much as he could, my husband, Muhammad Sharif Abd Rani. My children Sirhan, Suhail and Fahima. Your laughs and smiles gave me the strength to pursue this study to the end. I would like to express my sincere gratitude to my mother, Hazziah Lazim for her continuous love and Dua. And my thanks to all my brothers for their supports.

Thank you to Amir and Chiara for all your helps. Thank you to all Malaysian York family.

Finally, I would like to acknowledge the Ministry of Higher Education Malaysia for providing me with the financial supports to pursue this study.

AUTHOR'S DECLARATION

I, Siti Anizah Muhamed, declare that this thesis titled, ‘Objective Assessment of Neurological Conditions using Machine Learning' is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

CHAPTER 1

Introduction

This chapter provides an overview of the thesis, motivations for the work, the hypothesis and the contributions made.

1.1 Overview

Human body movement is a result of a complex system comprising our brain, spinal cord, nerves and muscles. When a human has a desire or need to move, a specific part of the brain sends signals through the spinal cord and nerves to the muscles. Muscles that are attached to the bones of the skeletal system along with other moving parts such as skin and soft tissues are responsible for human movement. If there are problems in the part of the brain responsible for sending the signals or failure of the signals to reach the muscles as intended, movement disorders can occur. The term 'movement disorder' was coined in late 1960's and since then has become a distinct domain. Movement disorders are a sub-set of neurological conditions which in turn are the biggest cause of loss of healthy life in the world's population. Examples of common and well-known movement disorders are Parkinson's disease, dystonia, essential tremor and a range of diseases known

collectively as atypical parkinsonism. Some of these diseases worsen over time and, hence, are known as neurodegenerative disorders. Some are hereditary, such as Huntington's disease and Wilson's disease. Movement disorders symptoms vary from the slowing of movements, spontaneous muscle contractions, to repetitive and involuntary movements such as grimacing or eye blinking. They also come with non-motor symptoms such as depression, dementia, sleep disorders and pain.

Being able to move our body normally is a most important part of every-day living. Inability to do so has many implications for quality of life and the care required. Movement disorders cause multiple impacts to a patient, family, society and the country's economy. One of the important efforts to reduce this impact is an early and accurate diagnosis. Unfortunately, clinical diagnosis of movement disorders is subject to the interpretation of clinicians, causing a high rate of misdiagnosis. Computerised methods can support clinical diagnosis through objective assessment. However, many of the computerised methods can be expensive, need experts to operate or are too specific to a certain type of disorder. Therefore, this research proposes a computerised method that can be used on a range of movement disorders at relatively low cost, that is reliable in a conventional clinical setting and provides an objective assessment to support a clinical diagnosis. This method manipulates real-time kinematic data of movement disorders from patients performing basic motor tasks, acquired using electromagnetic tracking sensors to build classifiers employing Evolutionary Algorithms (EAs). EAs are a type of a machine learning method that is inspired by natural evolution. The classifiers evolved have been trained to identify a specific movement disorder with the added benefit of characterising the features of the disorder. Movement disorders of interest in this research are Parkinson's disease (PD), essential tremor (ET), Huntington's disease (HD), dystonia and functional (psychogenic) dystonia.

1.2 Motivation

Movement disorders are a category of neurological disorders. Neurological disorders, also known as central nervous disorders, are conditions caused by an abnormality in the human central nervous system. The symptoms vary according to which part of the brain is affected. In Europe, neurological disorders are classed as the highest contributor to global burden of disease (GBD) that assesses mortality and disability (World Health Organization, 2006). In the latest analysis of 2015, the worldwide GBD study found that neurological disorders as a group is the largest cause of loss of healthy life and second only as causes of global death (Naghavi et al., 2015).

With the exception of Alzheimer's disease, movement disorders such as Parkinson's disease, essential tremor, restless legs syndrome and dystonia, are the most common neurological conditions (Wenning et al., 2000). In 2005, over one million people in Europe were diagnosed with Parkinson's disease, with around 127000 of them living in the United Kingdom. It has been projected that the number will double by 2030 (Dorsey et al., 2007). Approximately 2.2% of the US population have essential tremor, some 6.38 to 7.63 million people (Louis & Ottman, 2014). These disorders afflicted not only the patients; they pose a challenge to families, caregivers, communities, the health care system and the economy.

Cost

The economic cost of movement disorders is high, consisting of public health delivery activities that include formal health care, personal medical attention, home care and self-care, health promotion and disease prevention (World Health Organization, 2006). It was estimated in 2011 European countries spent 13.9 billion euros on public health deliveries for Parkinson's disease alone. While in the United States, the national economic burden

of PD exceeds $14.4 billion in 2010 (EPDA, 2012; Kowal et al., 2013). Apart from direct costs, there are indirect costs that result from job loss, change in worker productivity, reduced earning ability and missed work days caused by the disease. In fact, indirect cost became the bigger part of the economic burden especially in advance stages of the disease (Keränen et al., 2003). The high cost of economic burden is linked to not only to the financial burden but also to the quality of life of patients and their families (Dowding et al., 2006).

Quality of life (QOL)

The symptoms of movement disorders are many-fold. A few examples are slow decreased movement, spontaneous muscle contractions, repetitive and involuntary movements such as grimacing or eye blinking and involuntary rhythmic shaking of parts of the body. Apart from motor symptoms, there are also non-motor components, such as olfactory dysfunction (disturbances of smell and taste), depression, dementia, sleep disorders, bowel and bladder problems, fatigue, apathy (lack of emotion, passion, interest or motivation) and pain. Although these abnormalities and disturbances in movement have a major impact on health, they are rarely a direct cause of death. Instead, the conditions result in disabilities that affect the patient's functioning and quality of life. The symptoms cause a decrease in quality of living (QOL) in many ways including loss of independence, negative impact on relationships and work, and difficulties in performing day-to-day activities (Dowding et al., 2006; Soh et al., 2013; World Health Organization, 2006). The disorders affect QOL not only of the patients, but the caregivers as well (Carter et al., 2008; Martínez-Martín et al., 2005, 2007). The degrading quality of life of patients and caregivers is especially true for neurodegenerative disorders such as PD and HD where the burden increases as the disease progresses. In addition, recent research also indicates that the most common movement disorder, essential tremor, might also worsen over time

(Louis et al., 2016). In order to reduce the burden caused by movement disorders, well-planned, efficient disease management programs are critical. However, the first step of such program, must be an accurate diagnosis.

1.3 Clinical diagnosis of movement disorders

One of the critical parts of the management of a disease is the diagnosis. Diagnostic tests are imperative to provide information about the patient's condition, to influence the health care provider's plan for managing the patient and to understand disease mechanism and its history through research. However, for most movement disorders, diagnosis proves to be difficult and requires experts in the field to achieve higher accuracy. For many disorders, no known pathological markers are available to confirm the diagnosis. Although many diagnostic tests are available, some of these, such as single photon emission computed tomography (SPECT) and positron emission tomography (PET), are expensive and invasive (National Collaborating Centre for Chronic Conditions., 2006). Hence, most diagnosis of movement disorders remains clinical. Clinical diagnosis is based on clinical features taken from examination, patient history and response to medication. Unfortunately, the clinical diagnosis is usually complex, subjective and hence, often unreliable. The accurate diagnosis of movement disorders is particularly difficult at the early stages. Clinical misdiagnosis rates in movement disorders is high. For example, in Parkinson's Disease, it is estimated that the misdiagnosis rate remains as high as 25% since the 1970's (Ali H Rajput & Rajput, 2014). About one in three patients with other tremor conditions were misdiagnosed as having ET, with the most frequent incorrect diagnoses being Parkinson's disease and dystonia (Jain et al., 2006). This means that there are ET patients being treated as PD patients (and vice versa) when the prognosis of the two diseases are clearly different. Moreover, since there are disorders that are

genetically inherited (such as Huntington Disease and Wilson's disease), the diagnosis has serious implications for patient's family. Apart from that, due to their sometimes unusual presentations, patients with movement disorders may be diagnosed as having disorders caused by psychological problems (psychogenic movement disorders). More objective diagnosis is needed to confirm the clinical diagnosis.

1.4 Objective diagnosis of movement disorders

There are many potential objective assessment methods in movement disorders based on motor symptoms. Works include using electronic sensors to collect movement data. For example, using the patient's performance in drawing an Archimedes spiral, handwriting, rapid alternate movement or doing an everyday task such as walking. Different technologies have been used in clinical settings to obtain kinematic data, such as digitising pads for 2D drawing data, gyroscopes for rapid alternating movements, motion capture using cameras and optical devices, and electromagnetic sensors for motion capture in 6 degrees of freedom (Allen et al., 2007; Daneault et al., 2013; D. a. Heldman et al., 2014; D. A. Heldman et al., 2011; Á. Jobbágy et al., 2005; Rovini et al., 2017; Spasojević et al., 2017; Stamatakis et al., 2013; Yokoe et al., 2009a). Using the acquired kinematic data, quantitative studies demonstrate that movement features belonging to different groups of movement disorders or healthy controls can be differentiated statistically. Significant features, such as speed and acceleration of movement, can be used to train algorithms that are able to automatically classify these groups. The classification task is part of what is known as predictive modelling. Predictive modelling is basically the process of developing a mathematical tool or model that generates an accurate prediction by considering relationships between the available data sets. Previous work has demonstrated that using machine learning methods can successfully achieve

high accuracy in the classification of movement disorders. However, many of these works focused on classifying movement disorders of patients from healthy controls as in in the reviews of machine learning classification of PD patients (Ahlrichs & Lawo, 2013). Yet, to address the problem of misdiagnosis, the real challenge is to differentiate between groups of different movement disorders that are commonly mistaken for each other.

1.5 Predictive modelling of small and imbalanced data

In a classification task, as is widely understood, the bigger the sample size, the higher the probability of finding a significant result. However, a large sample size is not always obtainable in any field of research, biomedical data included. A small sample size can be the consequence of research protocols, a small research population, or when it is ethically and morally unjustified to gather a larger sample. The number of training samples may also be limited because obtaining samples in a form suitable for learning may be costly or impractical (Weiss & Provost, 2003). Another well-known challenge in classification tasks is when the data in one group or class is very small compared to the other class(s). The problem is called an imbalanced data distribution or imbalanced dataset. There are many natural situations with imbalanced datasets such as analysing financial risk, predicting technical equipment failures, managing network intrusion and information filtering (He & Garcia, 2009; He & Ma, 2013). It is also common in the medical field, such as the occurrence of a rare disease, or newly found version of known medical conditions, as in case of psychogenic movement disorders. There are many machine learning methods tested to combat imbalanced dataset problems. Some use standard learning algorithms, others propose new methods to address the specific problems and some, combinations of the two. However, most of these solutions were tested in isolation usually with artificial datasets, and their solutions are not easily pertinent to the real-world

problems (Napierala & Stefanowski, 2016). The challenge of binary classification on small and imbalanced real datasets of movement disorders is yet to be tested using EAs.

1.6 Evolutionary Algorithms

Evolutionary algorithms (EAs) are a computational intelligence technique inspired by natural evolution and survival of plants and animals. There are several different types and versions of EAs with the major ones being genetic algorithms, evolutionary programming and genetic programming. EAs and genetic computation are known for their significant contributions to medical applications (Smith, 2011; Smith & Cagnoni, 2011). EAs have not only proven effective in classification tasks for medical diagnosis but offer one great advantage – an ability to provide "white box" analysis in which a complete description of how features in the data have been used to classify between the groups (Lacy, 2015; Lones, Alty, et al., 2014; Lones et al., 2013). With a reliable amount of data, the description of these features can be generalised to inform clinical assessment and support clinical diagnosis (Lones et al., 2013).

1.7 Summary

When individuals lose the ability to move normally (along with other symptoms that present with movement disorders) it impacts their quality of life, as well as the people around them, their community and public health providers. Early and accurate diagnosis plays an important role in reducing the impact and prolonging a good quality of living. Clinical diagnosis of movement disorders are often subjective, prove to be complex and contribute to high rates of misdiagnosis. EAs, applied to predictive modelling using data from motion sensors offer an objective diagnosis of movements disorders. Additionally, they can give a deeper understanding of movement characteristics of different disorders

while performing the same motor tasks. Further analysis of the movement characteristic open possibilities to inform clinical assessment. However, the reality is biomedical data from patients are hard to obtain, time-consuming and expensive; this is true of data obtained from movement disorders patients as well. Furthermore, certain types of movement disorders, such as psychogenic movement disorders, are rare and naturally small in number, but need to be differentiated from their organic counterparts. Motivated by these challenges, the work presented in this thesis employs EAs to evolve classifiers that can objectively support clinical diagnosis and characterise a range of movement disorders.

1.8 Hypothesis

Based on the evidences that (1) movement disorders are common conditions with high global morbidity and mortality, causing a significant impact to the economy and negatively affecting quality of living of involved parties, (2) current clinical diagnosis of movement disorders is complex and largely subjective which results in a high rate of misdiagnosis, and (3) EAs offer great potential in objective diagnosis and characterisation of movement disorders, it is the hypothesized that:

> *"Evolutionary Algorithms offer a means of differentiating and characterising a range of movement disorders using digitised kinematic data from common conventional clinical tasks."*

The work in this thesis considers using kinematic data from three motor tasks taken from a conventional and often used clinical assessment, the Movement Disorders Society Unified Parkinson's Disease Rating Scale, MDS-UPDRS namely finger tapping, hand opening-closing and hand-pronation-supination.

The objectives are to:

- Apply EA predictive modelling to classify Parkinson's disease patients from healthy controls, essential tremor and Huntington's disease patients using available datasets.

- Observe characteristics of different movement disorders performing the same motor tasks by analysing the EA classifiers evolved.

- Use EAs to discriminate patients of functional (psychogenic) dystonia from healthy controls, and its organic counterpart, primary dystonia

1.9 Thesis structure

This thesis is organised as follows:

This chapter provides an introduction and outline of work presented in the thesis. It also states the motivations, objectives and hypothesis of the research. Chapter 2 gives an overview of Parkinson's disease and other movement disorders describing their history, anatomy, causes, main symptoms and the cognitive impairment associated with them. Chapter 3 provides an overview of evolutionary algorithms, examining the different program representations that define the different kinds of algorithm. The focus is on Cartesian Genetic Programming, the evolutionary algorithm chosen for this work, which is explained in detail in section 3.4. Chapter 4 contains a description of the motor tasks assessments, equipment used and the methodology used in this study. Chapter 5 presents results of classification and characterisation of movement features of four groups: healthy elderly people, PD, Huntington's disease and essential tremor patients. Chapter 6 gives an overview of dystonia, describing its history, the different kinds of dystonia with their signs and symptoms, the cause, the diagnosis and the two kind of dystonia considered:

functional and organic dystonia. The results of classification experiments are also discussed. Chapter 7 presents the conclusions of the work, revisits the hypothesis and offers suggestions for future work.

CHAPTER 2

Movement Disorders Conditions

The term ‘movement disorders’ was coined in late 1960’s by Stanley Fahn and Lewis P. Rowland in an effort to extend a Parkinson’s disease clinic at the University of Pennsylvania into a clinic that also covered other movement abnormalities (Klein, 2005). Since then, the term has been widely used in research related to neurological problems. Movement disorders are neurologic syndromes that show excess of movement or a paucity of voluntary and automatic movements, but not related to muscle weakness or spasticity (Fahn, 2011). However, the disorders that were included under this term are characterised by many other complex signs and symptoms (Fernandez et al., 2015).

Historically, whereas the nineteenth century can be viewed as the century that established neurology as a speciality in medicine, the twentieth century,specifically the second half, marked the evolution of movement disorders as a distinct domain (Goetz et al., 2001). Although many conditions that are today known as movement disorders had already been discovered since the last century, it is only in late 1960’s that the field of movement disorders was born. It was started with observations and efforts by C. David Marsden and Stanley Fahn (Goetz et al., 2001). In the beginning, the classification of the disorders was according to localisation of anatomical parts of body where the movement was affected.

Since then, other classifications were introduced, such as according to aetiology (causes), phenomenology (experience), and pathophysiology (Fahn, 2011). Perhaps the modern classification that is most accepted is the one developed and updated by Marsden, Fahn and Jankovic, published in Fahn et al. (2011). In this classification, the two major groups are the hypokinesias and the hyperkinesias. Hypokinesia conditions are where the movement is slower than normal, and hyperkinesia is the opposite. Other modes of classification are by clinical criteria such as age of onset and response to treatment, by post-mortem criteria such as presence of Lewy bodies in PD patients, and by genetic/molecular criteria such as a defect in protein function (Klein 2005).

Table 2.1 : Classification of movement disorders based on speed of the abnormal movements. Content of the table from (Fahn et al., 2011)

Hypokinesias	Hyperkinesias
Akinesia/bradykinesia (parkinsonism) Apraxia Blocking (holding) tics Cataplexyanddropattacks Catatonia, psychomotor depression, and obsessional slowness Freezing phenomenon Hesitant gaits Hypothyroid slowness Rigidity Stiff muscles	Abdominal dyskinesias Akathitic movements Ataxia/asynergia/dysmetri a Athetosis Ballism Chorea Dystonia Hemifacial Spasm Hyperekplexia Hypnogenic dyskinesias Jumping disorders Jumpy stumps Moving toes and fingers Myoclonus Myokymia and synkinesis Myorhythmia Paroxysmal dyskinesias Periodic movements in sleep REM sleep behavior disorder Restless legs Stereotypy Tics Tremor

2.1 Parkinson's Disease

Parkinsonism syndrome or Parkinsonism is a group of neurological conditions that share common symptoms clinically defined by the presence of certain motor features; tremor at rest, rigidity, bradykinesia, and gait and postural abnormalities. Parkinson's disease (PD) is the idiopathic version of this syndrome where a person slowly develops difficulty with the control of movements without any known cause or confirmed risk factors. Simple daily activities such as walking and reaching for an object become challenging. Parkinson's disease is chronic and progressively worsening over time. The difficulties are largely caused by loss of dopamine-generating neurones in a part of the brain called the Substantia Nigra. Both motor and non-motor deficits caused by Parkinson's disease substantially impact on a person's quality of life (Martinez-Martin et al., 2011). Although Parkinson's disease is one of the most well-researched movement disorders it does not bring us closer to a complete understanding of this complex disease. Significant progress has been made in the treatment of PD motor symptoms, though the mechanisms underlying treatment success are often not well understood (Sulzer et al., 2015). Parkinson's disease cannot be confirmed while the patient is alive. The diagnosis can only be made with certainty during post mortem. Low diagnostic accuracies have been linked to PD for more than 20 years where autopsies revealed about 25% of patients with PDas final clinical diagnosis did not meet required pathological markers (Ali H Rajput & Rajput, 2014).

2.1.1 Signs and Symptoms of Parkinson's disease

There are four main signs of Parkinson's disease: resting tremor, rigidity, bradykinesia and postural instability.

Tremor

Tremor is a rhythmic, involuntary shaking of a part of the body. Rest tremor is the most common and easily recognised symptom of PD. It is obvious when the affected body part is supported against gravity (and is often absent or disappears during voluntary movement). The frequency of the rest tremor is between 4 and 6 Hz, and usually initially showing in one hand or foot without any regular pattern (Stanley Fahn et al., 2011). The term ‘‘unilateral onset’’ is used to describe the usual pattern of onset in Parkinson’s disease, meaning that the rest tremor begins on one side of the body and in one limb.

Rigidity

Rigidity is an increase in resistance to passive movements that occurs in the limbs. For any movement to happen, there needs to be a pattern of contraction and relaxation of muscles. Contraction is stimulated by signals from brain that prompts the movement of the muscle, whereas relaxation follows when the signals end, and the muscle eases back to its natural state. Rigidity in Parkinson’s disease develops when the natural contraction and relaxation of opposing muscles fails to take place, caused by the failure of the brain signals to trigger the muscle movement.

Bradykinesia

The term akinesia literally means absence of movement; bradykinesia, meaning slowness; and hypokinesia, meaning decreased amplitude; are all used interchangeably to describe the most prominent phenomena of Parkinsonism. The conditions they describe are usually referred to collectively as bradykinesia (Figure 2.1). The complex nature of Bradykinesia itself is one of the reasons that make it difficult for clinicians or neurologists to be certain of its existence at an early stage of Parkinson’s disease. Clinicians look for signs of

bradykinesia by observing a patient's ability to perform rapid, repetitive, alternating movements of the hand such as finger taps, hand grips and hand pronation– supination (Jankovic, 2008).

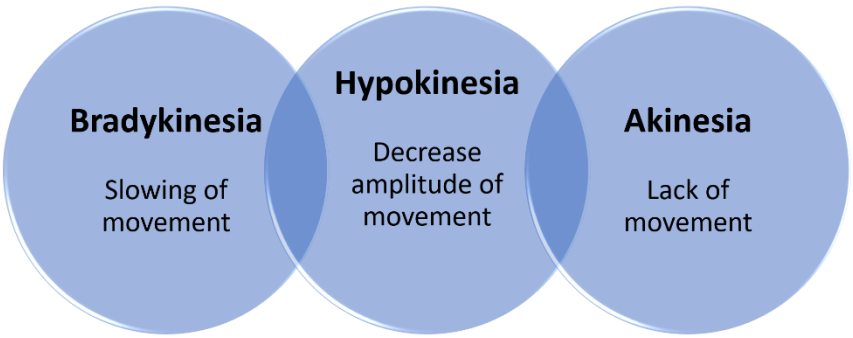

Figure 2.1: Spectrum of Bradykinesia. Figure edited from Fernandez et al.(2014)

Postural instability

Postural instability is the inability of a person to maintain posture and balance caused by impaired postural reflexes. People with Parkinson's disease tend to fall backward, although falls are unusual in the early stages. Postural instability can be present even at diagnosis in some patients but when prominent in early stages of the disease it should suggest the possibility of an atypical parkinsonian disorder, such as progressive supranuclear palsy (Egerton et al., 2012). A normal individual should be able to recover balance spontaneously with a quick backward step, but patients with PD often take several steps to recover balance and may even fall if not caught in time. Although the four main symptoms described above are the most well-known signs of PD, several secondary signs and symptoms, including non-motors symptoms, have been detected. Note that not everyone with Parkinson's disease develops the same symptoms. Other reported symptoms include masked facies (loss of facial expression and a decreased rate of eye blinking), speech deficits, dysphagia (difficulty in swallowing), drooling, gastrointestinal complications, micrographia (handwriting becomes small because of difficulty with fine

motor movement), disturbance of the spatiotemporal aspects of gait, sleep disorders, and mood fluctuations. Non-motor symptom such as depression is the most common psychological problem in Parkinson's disease (Sharma, 2008). These symptoms will eventually severely impair patients of PD quality of life such as shown in advanced Parkinson's disease patients.

2.1.2 The anatomy of Parkinson's Disease

In the 1960s, scientists discovered that deficiency of a compound named dopamine in the specific part of the brain called basal ganglia is at the root of Parkinson's disease (Hornykiewicz, 2006). The basal ganglia refers to a group of structures linked to the thalamus in the base of the brain and involved in coordination of movement. Parkinson's disease is associated with damage to the basal ganglia and has several distinct outward symptoms.

Symptoms of Parkinson's disease are caused by the death of neural cells in one of the nuclei (group of neurons) of the basal ganglia called the Substantia nigra; the cause of the cell death is unknown. The Substantia nigra naturally produces a neurotransmitter (chemical messenger) called dopamine that helps in the transmission of signals in the brain. Dopamine is vital to regulate movement throughout the body.

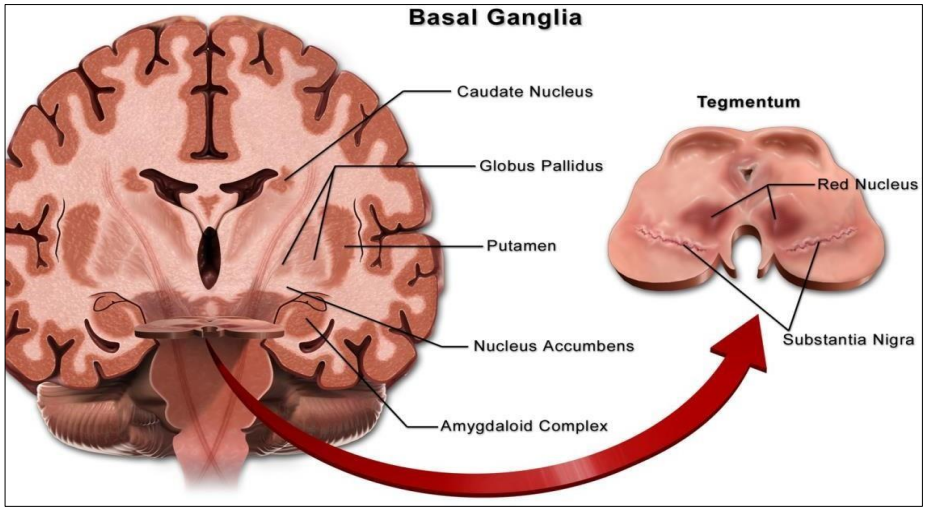

Figure 2.2: Location of the substantia nigra within the basal ganglia (Figure from https://en.wikiversity.org)

As the disease progresses and less dopamine is delivered, outgoing signals become undependable and movements become irregular and uncontrolled. Figure 2.3 shows a normal midbrain compared to Parkinsonian midbrains with pigmentation loss.

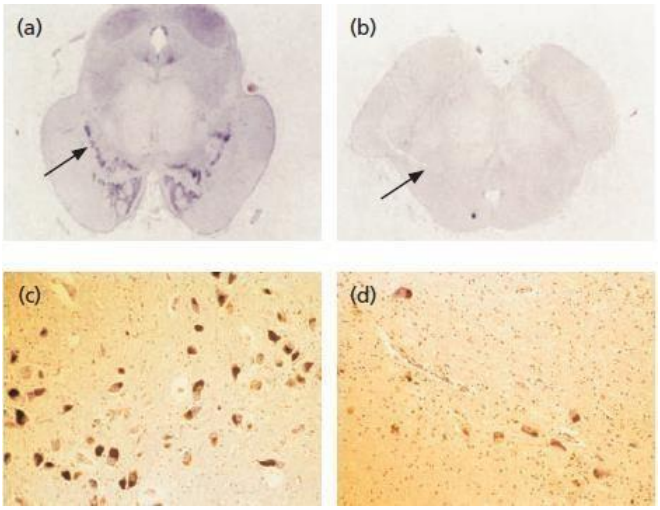

Figure 2.3: Pigmentation loss in the substantia nigra a) Normal b) Parkinsonian midbrain showing a characteristic of pigmentation loss in the substantia nigra (arrowed). Micrographs of the substantia nigra reveal c) normal pigmented neurons in normal brain and d) the loss of pigmented neurons in a brain affected by Parkinson's disease. Figure from Chaudhuri & Ondo (2009).

Apart from the cell loss in the Substantia nigra, pathological changes of PD also include the presence of Lewy Bodies (LBs) in surviving cells. Lewy bodies are abnormal masses of protein that develop inside nerve cells found in PD and occur in all areas of neuronal degeneration. Although LBs are a pathological hallmark of idiopathic PD, unfortunately, they are not unique to PD; other conditions in which they may be seen in degenerating neurons include corticobasal degeneration, progressive supranuclear palsy, motor neurone disease, ataxia telangiectasia and Hallevorden-Spattz disease (Gibb & Lees, 1988), classic Pick's disease, argyrophilic grain disease, Alzheimer's disease, and dementia with Lewy bodies (Popescu et al., 2004).

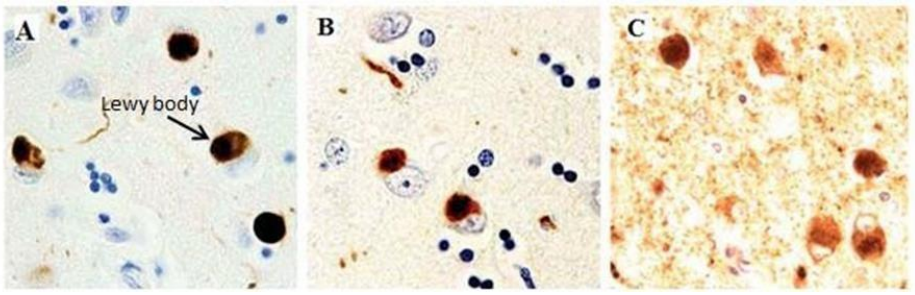

Figure 2.4: Lewy bodies presence in neurological conditions LBs are numerous in the amygdala in Alzheimer disease (A), dementia with LB (B), and classic Pick 's disease (C). Figure from Popescu et al., (2004).

Other than pathological overlap with these diseases, PD also shares some clinical features with a number of other neurodegenerative conditions which make it even difficult for a clear-cut diagnosis. Conditions showing similar symptoms as PD are discussed further in Section 2.1.6. The success of determining what is happening to PD patients still leaves us puzzling what cause it to happen. The idiopathic nature of PD means that scientists still cannot comprehend what is actually causing the loss of neurons cells and presence of LBs in affected areas of the brain. Research is ongoing to find out possible causes of PD and who the people are that have a higher chance of developing the disease.

2.1.3 The diagnosis of Parkinson's Disease

As for most movement disorders, the diagnosis of Parkinson's disease is a 'clinical' diagnosis since the methodology depends heavily on the observation for manifestations of the Parkinsonism symptoms and history taking of the subject by the neurologist or other qualified medical practitioners (National Collaborating Centre for Chronic Conditions., 2006). In the face of huge resources devoted to learn more about Parkinson's disease, there are still no laboratory tests such as a tissue diagnostic or other conclusive biomarkers test that can confirm the diagnosis of Parkinson's disease while a patient is alive.

2.1.4 Clinical assessment of Parkinson's disease

Demonstration of Parkinsonism requires evidence of bradykinesia and at least one of the other cardinal features: tremor at rest, rigidity and postural instability. These criteria have been summarised in the UK Parkinson's Disease Society Brain Bank Criteria (UKBBDC) (Gibb & Lees, 1988). UKBBDC is the most widely accepted clinical criteria for the diagnosis of PD and may improve clinicopathological correlation when strictly applied (Hughes et al., 2001). There are many scales that have been developed for the purpose of quantifying the severity of PD. There are scales to assess PD manifestation (impairment), disability, and scales that assess impairment and disability (Perlmutter, 2010). Table 2.2 summarise the scales. Among these scales, the Unified Parkinson's Disease Rating Scale (UPDRS) has been the most accepted and popular rating scale (Martinez-Martin et al., 1994). However, several highlighted limitations of UPDRS (Hilten et al., 1994; Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease, 2003) have led to a new revised version of UPDRS by the Movement Disorders Society: MDS-UPDRS (Goetz et al., 2008)

Table 2.2: Clinical rating scales used to assess Parkinson's disease

Assessment	Scales
Motor symptoms	• Webster Scale • Columbia University Rating Scale • PD Impairment Scale
Disability	• Hoehn & Yahr • Schwab and England
Motor symptoms & disability	• UPDRS, MDS-UPDRS • New York University Scale • Short PD Evaluation Scale • University of California Los Angeles PD Disability Scale

MDS-UPDRS

The Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) has four sections:

Part I: non-motor experiences of daily living

Part II: motor experiences of daily living

Part III: motor examination, and

Part IV: motor complications.

In each part, there are items that assess different parts of PD. This rating scale evaluates the severity of PD symptoms in a 5-point scoring system, 0 for no symptom and 4 for a marked severity of the symptom. Items from parts I, II and IV are rated by the patient themselves, either by filling in the questionnaire or by answering questions (from the scale) from the clinicians. Part III items are rated by clinicians, through observation and motor examination. In part III, instructions are given on motor tasks to be executed by patients and guides for clinicians on how to rate the performance. The full MDS-UPDRS scale is provided in Appendix A of this thesis. Examinations by clinicians comprise items

to test symptoms such as rigidity, speech impairments and leg agility. Motor tasks include finger tapping, hand movements (opening-closing), pronation-supination movement of hands, toe-tapping and gait testing.

Part III of the scales (UPDRS and MDS-UPDRS) have been tested and proven to provide good quality of measurements (Goetz & Stebbins, 2004; Goetz et al., 2008). However, there is no confirmation on how strictly the criteria in any guide are followed during the diagnosis process. The clinical diagnosis is considered as not definitive. The gold standard for definitive diagnosis of PD is pathologic findings which is neuronal loss in substantia nigra and present of Lewy bodies in the brain of patient. This confirmation is done during autopsy. Comparison has been made between clinical diagnosis and findings of autopsy on PD patients to determine the diagnostic accuracy of Parkinson's disease.

2.1.5 Measurement of diagnostic accuracy

Accurate diagnosis is very important in medicine, usually determining what additional diagnostic tests are needed, help in choosing more targeted, effective and often less invasive treatments or interventions, to eventually improve patient care management and reduce healthcare costs. The accuracy of a diagnosis is measured by comparing the test results to the true condition status of the patient or the gold standard. In case of PD, the gold standard is autopsy report of the patient. Two basic measures of diagnostic accuracy are sensitivity and specificity. Their definitions are best illustrated by a table with 2 rows and 2 columns, or table of frequency (Table 2.3 from Petrie & Sabin (2000)).

Table 2.3: Table of frequencies

	Gold Standard Test		
Test result	Disease	No Disease	Total
Positive	a	b	a + b
Negative	c	d	c + d
Total	a + c	b + d	n = a + b + c+ d

If n is number of individuals studied, then a + c individuals have the disease. Of the a + c individuals who have the disease, a have positive test results (true positives, TP) and c negative have test results (false negatives, FN). Of the b + d individuals who do not have the disease, d have negative test results (true negatives, TN) and b have positive test results (false positives, FP).

Sensitivity is the proportion of individuals with the disease who are correctly identified by the test. A test that is able to detect all the individuals with disease is said as having perfect sensitivity or 100% sensitivity or a sensitivity of 1.0. It can be calculated as:

$$\text{Sensitivity} = \frac{a}{a+c}$$

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

Specificity is proportion of individuals without the disease who are correctly identified by the test, calculated as following:

$$\text{Specificity} = \frac{d}{b+d}$$

$$\text{Specificity} = \frac{TN}{TN+FP}$$

A test with 100% specificity will detect all the healthy individuals and give a negative result to them (negative of having the disease). The best test would be one that has a

sensitivity and specificity that are both as close to 1 (or 100%) as possible. However, in practice, sensitivity may be gained at the expense of specificity, and vice versa. Whether the aim is for a high sensitivity or high specificity depends on the condition or disease, also what the consequences of FN or FP test result on the individuals tested are. For conditions that are easily treatable, high sensitivity is preferable; for those that are serious and untreatable, high specificity is important to avoid making a false positive diagnosis (Petrie & Sabin, 2000). In the case of Parkinson's disease, false positive diagnosis results in the unnecessary administration of drugs, a false negative diagnosis delays the initiation of drug therapy or other types of therapies.

Another way diagnostic accuracy is commonly described are predictive values which highlights the consequences associated with the test results. These predictive values provide information about how likely it is that the individual is correctly diagnosed. Positive predictive value (PPV) is the proportion of individuals with a positive test result who have the disease. Negative predictive value (NPV) is proportion of individuals with a negative test result who do not have the disease. These values can be calculated as following:

$$\mathrm{PPV} = \frac{TP}{TP+FP}$$

$$\mathrm{NPV} = \frac{TN}{TN+FP}$$

Predictive values are dependent on how common the disease in the population is being studied or the disease prevalence. In populations where the disease is rare, the NPV value will be much lower than in populations where the disease is common and vice versa. In certain cases, there are needs to make a diagnosis on the basis of a continuous measurement. It means that there is no certain definitive threshold above (or below) that

implies the disease is positively detected. In these cases, the upper (or lower) limit of the reference interval can be used. A cut-off value choice will change its related sensitivity, specificity and predictive values. The objective is to choose a cut-off value that will optimise these measures as desired.

2.1.6 Differential diagnosis of Parkinson's disease

The low diagnostic accuracy is one of the issues that have been associated with Parkinson's disease for years. The accuracy rates were observed by comparing final clinical diagnosis of a patient to autopsy result on brain matters. This type of study is known as clinicopathological. Perhaps the first significant clinicopathological study that presented a quite shocking result about the diagnostic accuracy of Parkinson's disease was published in 1991. In Rajput et al.(1991) , autopsy was performed on 65 patients, 41 of them having PD as a final clinical diagnosis after an average of 11.7 (range 2 -39) years of illness. A 22 year study by Rajput et al. (1991) found out that out of the 41 patients, Lewy bodies' presence was detected in only 31 of them which indicated only a 76% rate of diagnostic accuracy. Soon later, another clinicopathological study with a bigger number of subjects was published; the study on 100 cases of clinically diagnosed PD shown same rate of accuracy at 76% (Hughes et al., 1992). In contrast, very high percentages of diagnostic accuracy (99%) have been reported in tertiary hospitals by movement disorder specialists (Hughes et al., 2002). It was speculated that the higher accuracy may be due different methodology (Rajput & Rajput, 2014) or the fact that patients were evaluated by specialist neurologist in a specialist movement disorder service. This may have been proven by a later study by Joutsa et al. (2014) that investigated diagnostic accuracy of parkinsonism syndromes by general neurologists in Finland. Only 58 (75.3%) of the 77 patients that had been diagnosed as Parkinson's disease (PD) were confirmed after the neuropathological examination. Another recent

study by Adler et al. (2014) reports results of diagnostic accuracy of an ongoing clinical-neuropathologic research. Taking into account all PD patients examined in this study, the final clinical diagnosis of PD is accurate in only 77% of the subjects. It can be seen that both studies, Adler et al.(2014) and Joutsa et al.(2014) showed percentage of diagnosis accuracies remarkably similar with 1990's studies, indicating that PD is still facing serious problems in this area. It is rather sobering that there is no marked improvement in accuracy of diagnosis since more than 20 years ago (Ali H Rajput & Rajput, 2014).

Diagnostic inaccuracy can cause patients to receive treatment as PD patients up to their demise when they actually had other Parkinsonism syndromes with similar features with PD. There are many possible answers to low diagnostic accuracy of Parkinson's disease. One possibility is the difficulty of rejecting other Parkinsonism conditions with the different degrees of overlap especially in the early course of the disease. Parkinsonism is an umbrella term used to describe many conditions that share the main symptoms of slow movement, sometimes with tremor, rigidity and problems with walking. A variety of conditions may cause Parkinsonism syndromes (also called Parkinsonian) independent of the idiopathic loss of substantia nigra neurons; these include degenerative conditions, drug-induced and several other types of movement disorders. Several conditions that have been misdiagnose as PD include Alzheimer's disease, progressive supranuclear palsy, multiple system atrophy, dementia with Lewy bodies, corticobasal ganglionic degeneration, vascular Parkinsonism (Adler et al., 2014; Hughes et al., 2002; Joutsa et al., 2014) and drug-induced parkinsonism (Rajput et al., 1991). Apart from Alzheimer's disease, they are all rarer than PD but may be clinically indistinct in the early stages of disease.

Multiple system atrophy (MSA)

Multiple system atrophy, MSA is a condition often confused with Parkinson's disease because of its Parkinsonism symptoms. There is up to 55% of cases where MSA patients die with PD as final clinical diagnosis (Quinn, 2005). Two variants of MSA are MSA-P and MSA-C. "P" in the first variant is for parkinsonian where there is predominating Parkinson's –like symptoms, and clinically the MSA-P phenotype presents the most difficulty in differentiating it from Parkinson's disease (Chaudhuri & Ondo, 2009).

The parkinsonian features of MSA include progressive bradykinesia, which is usually symmetrical, rigidity and postural instability. MSA-C is a cerebellar dysfunction variant and characterised by progressive ataxia (an inability to coordinate voluntary muscular movements) of the gait and arms and dysarthria (difficulty in pronouncing words).

Compared to PD, autonomic failure in the form of sympathetic dysfunction occurs early and more severe in MSA. Genitourinary problems such as erectile failure and urinary symptoms are also common in MSA. In contrast to PD, it is also unusual for a patient with MSA to develop dementia or hallucinations.

Progressive Supranuclear Palsy (PSP)

PSP is also known as Steele-Richardson-Olszewski syndrome. PSP can have a wide clinical spectrum, with different clinical variants. Classic PSP variant is characterised by supranuclear gaze palsy (predominantly vertical gaze), Parkinsonism, pseudobulbar affect, prominent frontal lobe syndrome, axial symptoms (neck and trunk) with early falls in the first year after onset, and symmetric symptoms. Resting tremor is uncommon. Gait is broad-based and unsteady, unlike the typical small-stepped shuffling gait in PD, with the typical "surprised" facies, spastic dysarthria, and retrocollis. Parkinsonian variant has parkinsonian features very similar to those seen in PD or MSA, but without much response to levodopa and a predominance of postural instability and gait dysfunction

(Fernandez et al., 2014). Even with strict clinical diagnostic criteria applied, autopsy revealed 22% of patients with PSP as final clinical diagnosis turned out to be PD patients (Respondek et al., 2013).

Dementia with Lewy bodies (DLB)

DLB is characterised by early onset progressive dementia with behavioural abnormalities that may interfere with normal social and occupational function, visual hallucinations, visual defects fluctuation in cognition and attention, even psychosis in the drug-naive state and Parkinsonism. DLB is very difficult to differentiate from PD dementia because patients with either condition can exhibit fluctuations in alertness, frequent falls, hallucinations, and sensitivity to PD medications (Fernandez et al., 2014). During autopsy, cases of DLB cannot be differentiated from cases of dementia occurring late in Parkinson's disease (Chaudhuri & Ondo, 2009).

Corticobasal Ganglionic Degeneration (CBD)

CBD is a slowly progressive rare condition with patients aged 60's to 70's. Unilateral development of tremor, apraxia and rigidity in an upper limb are among the first symptoms shown. Patients then develop progressive gait disturbances, cortical sensory loss, motor apraxia and stimulus-sensitive myoclonus, which results in a jerky, uncontrollable hands. In 50% of patients, a jerky, uncontrollable movement called alien limb phenomenon may have occurred. Along the course of the disease, the condition becomes bilateral with progressive cognitive impairment. Patients do not respond to Levodopa or other forms of dopaminergic treatment (Chaudhuri & Ondo, 2009).

Alzheimer's disease

Alzheimer's disease is the most common progressive late-onset disorder. It is a degenerative condition that leads to a condition called dementia. Dementia is a general term used to describe the loss of memory and mental abilities severe enough to affect daily life. Sixty to eighty per cent of all dementias are Alzheimer's disease cases that make it the most common type of dementia. With progressive loss of memory as its main feature, there are cases where it is mistaken as Parkinson's disease, especially PD with dementia variant (Adler et al., 2014; Joutsa et al., 2014).

Vascular Parkinsonism

Vascular Parkinsonism symptoms include rest tremor, bradykinesia, rigidity and postural instability but unlike in Parkinson's disease, the motor symptoms are usually bilateral. Most patients have predominant lower-extremity symptoms, such as a gait disturbance (freezing, gait initiation failure, turning difficulties), with minimal upper-extremity symptoms, or walk with small steps. These common features might make it challenging to differentiate the condition with Parkinson's disease that showing postural instability and gait difficulty (Chaudhuri & Ondo, 2009).

Drug-induced Parkinsonism (DIP)

Dopamine receptor-blocking antipsychotic drugs such as thioridazine, chlorpromazine and haloperidol commonly known to cause Parkinsonism, although any kind medication have a potential of doing so. Due to many overlaps of clinical features such as rigidity, bradykinesia, tremor and gait disturbance, Drug-induced parkinsonism may be misdiagnosed as Parkinson's Disease. Although DIP symptoms are usually more bilateral or symmetrical compared to PD and might have a slightly faster tremor, symptoms

between the two diseases are almost impossible to be clinically distinguished (Chaudhuri & Ondo, 2009).

2.2 Huntington Disease

Huntington's disease (HD) is a genetically inherited neurodegenerative disease, resulting in death within 15 to 20 years after diagnosis. The prevalence of diagnosed HD in the UK has doubled between 1990 and 2010. By extrapolation, it is estimated that there are more than 5700 diagnosed adult HD patients in UK (Evans et al., 2013). Huntington's disease has autosomal dominant inheritance. In this type of inheritance pattern, each children of an affected individual have a 50% risk of inheriting the disorder. Individuals at-risk are not gender dependent and it does not skip generations (Figure 2.5). Apart from apparent effects to quality of life of patients, the fact of dominant inheritance touches entire families: at-risk individuals and even genetically normal family members.

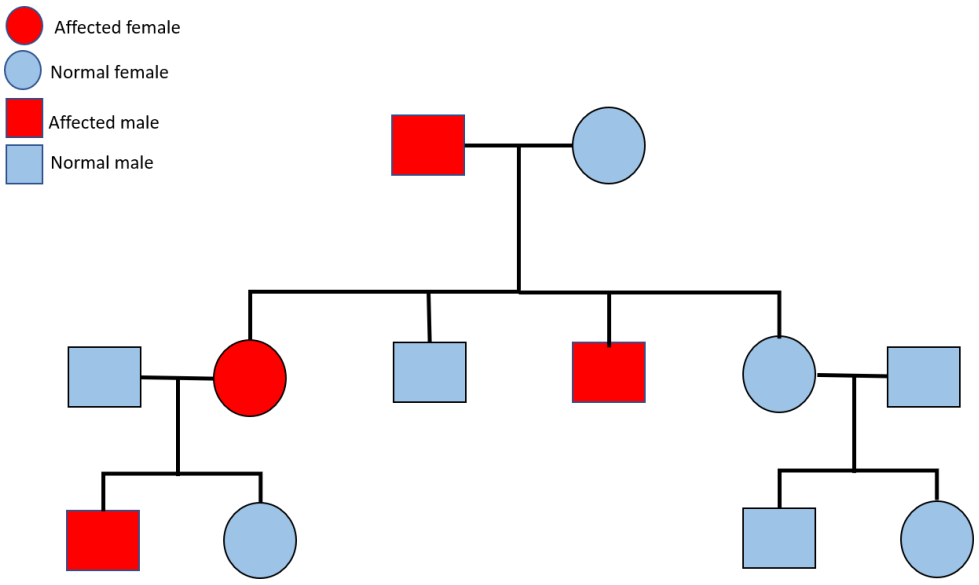

Figure 2.5: Chart redraw from Bates et al. (2002) to show autosomal dominant transmission in HD.

The accepted clinical feature of HD are psychiatric abnormalities, cognitive decline and movement disorders (Bates et al., 2002). Early signs and symptoms can include irritability, depression, small involuntary movements, poor coordination, and trouble learning new information or making decisions. As the disease advanced, these signs become more noticeable. Psychiatric abnormalities include changes in personality such as depression, anxiety, apathy, obsessive-compulsive behaviours, outbursts, addictions, and occasionally psychosis (Dayalu, 2015) . Cognitive decline in many cases ends up as dementia. Although many signs of behavioural changes might be shown at early stages of HD, motor signs are still the best indication for diagnosis.

2.2.1 Bradykinesia in Huntington's disease

The most well-known neurological feature of HD is the movement disorder chorea, an involuntary movement produced by jerk-like contractions of muscles that move randomly from one part of the body to another (Albanese et al., 2013). Chorea is characterised by excessive spontaneous movements that are irregularly timed, non-repetitive and abrupt. Other movement disorders include rigidity, slow or abnormal eye movements, impaired gait, posture and balance, and difficulty with the physical production of speech or swallowing. Progressive motor failure is a major cause of life-ending complications.

Bradykinesia also exists in HD. Although HD is characterised by chorea, a hyperkinetic disorder, studies shown that it can co-exist with bradykinesia. In fact, in Garcia Ruiz et al. (2000), not only HD patients were slower than controls; they were even slower than PD patients of matching ages. A follow-up study of 76 HD patients found that bradykinesia was already evident in early stages and increased linearly with increasing disease stage (van Vugt et al., 2004). On the contrary, other studies challenged the findings and suggest that bradykinesia is not a feature of movement disturbance in HD. Rather, it was suggested that slowness of movement, when observed in HD, is a result of

compensatory adjustments in velocity to improve accuracy in the presence of choleric intrusions (Duval et al., n.d.; Fenney et al., 2008). It raises the question of whether bradykinetic features found in HD share the same characteristics shown by PD patients or healthy people of old age.

2.3 Essential tremor

Tremor is a rhythmic, oscillatory movement produced by alternating or synchronous contractions of antagonist muscles (S. Fahn et al., 2007). It is the most common involuntary movements humans can experience. There are many ways to categorise tremor. Phenomenologically, it can be divided into two, action tremor and rest tremor. Action tremor is recognised when muscles are in use voluntarily. It is the opposite of rest tremor, which is one of the main symptoms of PD. Rest tremor is the type of tremor detected when the muscles are not in action. Action tremor is the main characteristic of Essential Tremor (ET), one of the most common neurological conditions. ET is chronic and progressive, and although once known as merely a motor disorder but psychological complications are now recognised as part of the features of this condition. ET is linked to degeneration of neurons in the cerebellum which is, like the basal ganglia (where movement disorders in PD start), also plays important roles in motor functions. Although patients of these two conditions experience different kinds of movements disorders, they have been mistaken for each other many times. A study disclosed that 30% of patients that were diagnosed as having ET were, in fact, having other conditions where most of them should be diagnosed as PD patients (Jain et al., 2006). The other way, in PD diagnosis, ET was one of the most common causes of misdiagnoses. The relationship between ET and PD always been a great interest in movement disorders research (Berkhout, 2015; Geraghty et al., 1985; Jiménez-Jiménez et al., 2012; Shahed & Jankovic,

2016). This is especially with findings of Lewy bodies in brain matter of ET patients suggesting the possibility that ET is also a neurodegenerative disease. That ET is also a type of disease that progress with the death of certain neurons, a major debate related to the topic of connection between ET and PD is whether ET is a risk-factor of PD (Benito-León, 2014; Louis et al., 2016). Even though the main similarity between PD and ET is the tremor itself, another main symptom of PD, bradykinesia, also exists in ET. This was reported in Duval et al. (2006) and a recent study by Goubault et al.(2017). Duval and Goubault used ET patients recorded movements of executing rapid alternate movements (pronation-supination) to make objective assessment. Prior to Duval, other research also detected a level of bradykinesia in ET patients that was comparable to mild symptoms of PD patients (Montgomery et al., 2000). Then, the idea was challenged by another study by Özekmekçi (2005) that found no significant difference of time taken to push microswitches between ET patients and healthy controls. Interestingly, they found slight prolong in ET patients, but concluded that the delay was due to the patient's effort to control the tremor in executing the task.

With all the controversies about the connections between ET and PD, until recently, no conclusions have been made and we are still far from learning the true nature of the relationship between these two disorders, if any (Algarni & Fasano, 2018). More methods of objective assessments are needed to confirm previous findings.

2.4 Objective assessment of Bradykinesia in movement disorders

Bradykinesia might have the highest potential as a motor progression marker of Parkinson's disease (Maetzler et al., 2009). As it is the only clinical sign that is compulsory for diagnosis, this may also contribute to current high diagnostic inaccuracy of Parkinson's disease (Bajaj et al., 2010). There are many potential objective methods to

assess bradykinesia in movement disorders. Research has previously used patient's digitised performance in drawing the Archimedes spiral, handwriting, rapid alternate movement, finger tapping or doing an everyday task such as walking. Different technologies have been used in clinical settings; for example, wearable sensors to obtain kinematic data such as digitising pads for 2D drawing data, a gyroscope for rapid alternating movements, motion capture cameras and optical devices, as well as electromagnetic sensors for motion in 6 degrees of freedom (Allen et al., 2007; Daneault et al., 2013; Heldman et al., 2014; Heldman et al., 2011; Jobbágy et al., 2005; Rovini et al., 2017; Spasojević et al., 2017; Stamatakis et al., 2013; Yokoe et al., 2009a). In these studies, movement features were extracted from the recorded data. Features such as movement times, speed, rhythm, reaction times and frequencies can be used to describe the structure of the data or make inference by comparing the movement features of involved groups. Popular statistical tests were used for comparison between groups of movement disorders and healthy controls, to look for the difference between the means or standard deviation of variables. Examples of common tests are the t-test and ANOVA for normally distributed data; Wilcoxon rank tests or sign test for data that does not meet statistical assumptions required for parametric tests.

In differentiating PD patients from healthy controls, several features of finger tapping movement have shown a marked difference between the two groups. Example of said features are finger-tapping test score (Jobbágy et al., 2005), maximum opening velocity and total distance of finger-tapping movement (Yokoe et al., 2009b), index for decrementing frequency, index for augmenting frequency, number of hesitations and number of halts (Stamatakis et al., 2013). Using other motor tasks, Daneault et al.(2013) found that maximum velocity of pronation-supination task correlated well with UPDRS clinician's score. Other useful features extracted in diadochokinetic movements studies

that involve hand pronation-supination can be found in Okada & Okada (1983) and Jiménez-Jiménez et al. (2010).

Other than PD patients, similar studies have been conducted on other movement disorders such as Huntington Disease (Garcia Ruiz et al., 2000; Künig & Alba, 2011; Mann et al., 2012; Martínez Pueyo et al., 2016; Müller et al., 2013; Thompson et al., 1988) and essential tremor (Duval et al., 2006; Jiménez-Jiménez et al., 2010; Kwon et al., 2016). From these studies, many important movement features have been extracted and proven as significant features in differentiating the groups. However, statistical analysis is usually focused on understanding the data (descriptive statistics) and making statements about the analysis of in-hand data (inference). Statistical techniques can be used in the prediction of unseen data points, but the performance is usually not optimum for complex, non-linear problems. Hence, the use of machine learning techniques which have been shown to classify unseen data, without prior assumptions about its statistical properties.

2.5 Machine learning in movement disorders

The significant features such as speed and acceleration of movement can be used to train classifiers that able to classify different groups automatically. The classification task is part of what is known as predictive modelling. Predictive modelling is basically the process of developing a mathematical tool or model that generates an accurate prediction by considering relationships between available data. Predictive modelling is commonly divided into two categories depending on type of the predicted data. To predict a continuous value a regression model is used, classification models are used when the response are in form of a categorical variable. There are many studies that have employed machine learning on PD patients' data to evolve classifiers. For example, in Deløcan et al., (2011), several types of Evolutionary Algorithm-Based Neural Networks were used

on a PD dataset taken from repository. In Spadoto et al.(2011), a classification task used an Optimum-Path Forest (OPF) on available PD features of Oxford Parkinson's Disease Detection Dataset. Apart from for classification task to detect PD, machine learning was also used to assess inter-rater agreement of motor task assessment. Martinez-Manzanera et al. (2015) used the popular machine learning method, Support Vector Machine (SVM) to check on inter-rater agreement of Bradykinesia items in MDS-UPDRS. The data in this study was collected using orientation sensors while PD patients performed finger tapping, diadochokinesis (pronation-supination) and toe-tapping. A similar study was conducted by Patel et al. (2009), also using SVM but the data was collected using accelerometers. Naive Bayes and Optimum-Path Forest were used on handwriting and spiral drawing data in Pereira et al.(2015). A review paper by Ahlrichs & Lawo (2013) listed many other studies that employed machine learning methods for automatic detection of PD motor symptoms. However, many of these works focused on classifying movement disorders patients from healthy controls. Whereas, to address the problem of misdiagnosis, the real challenge is to differentiate between groups of different movement disorders that commonly mistaken for each other.

2.6 Conclusions

As discussed in this chapter, bradykinesia is the only physical sign that is obligatory for diagnosis of PD. It is a complex clinical sign that comprises several abnormalities of movement including delayed initiation, reduced amplitude, reduced speed, impaired rhythmicity, and decrementing speed and amplitude. The phenomenology of the components of bradykinesia remain incompletely understood. This means that it can be very difficult to clinically ascertain whether bradykinesia is definitely present or not when a patient presents with only some of the components of bradykinesia, with very subtle

abnormalities or when other abnormal movements such as tremor are also present. Bradykinesia also exits in Huntington disease despite the fact that its main feature is chorea, a hyperkinetic movement. However, there are suggestions that the slowness in HD is not a feature, just an effort to mask other movement problems. The same goes for ET, bradykinesia was detected in a few studies but challenged by others. There is a question whether bradykinesia exists in HD and ET, and if it does, does it have the same characteristics with bradykinesia in PD? At the heart of this problem is uncertainty about the true nature of bradykinesia and its distinguishing characteristics. A better understanding of the characteristics of bradykinesia and how it differs between these groups can be used to inform clinical assessments towards conforming early diagnosis.

Quantitative studies (using statistical analysis of recorded data) have proven it can be used to help objective assessment of bradykinesia and other signs of movement disorders. Statistical techniques are not often used in quantitative studies as the problems are often complex and non-linear, leading to poor performance in describing and making inferences about the data collected. The selection of motor tasks in previous quantitative studies rarely mirror the standard motor tasks used in conventional clinical assessment. Therefore, it is proposed that in supporting the diagnosis of movement disorders and to inform clinical assessment, machine learning methods applied to movement data acquired from standard clinically accepted motor tasks, such as the Movement Disorders Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS), is an optimal approach to providing an effective and practical approach to diagnosing Parkinson's disease and differentiating it from other, similar neurodegenerative conditions.

CHAPTER 3

Machine Learning

Machine Learning is a field that has been growing in popularity for more than 30 years. It has taken inspiration from many aspects of human nature and biology in various ways. The biggest impact of machine learning has included applications in finance, information technologies, decision-making, healthcare, manufacturing and securities, to name a few (Paliouras et al., 2003). Machine Learning is often considered to be a part of artificial intelligence where computer algorithms learn by itself from data and information. It then uses information gathered from the learning process to autonomously identify underlying patterns in a dataset to describe the data and also make prediction based on the given pattern. The capability to produce sensible outputs for new inputs that were unseen during learning (the ability to generalise) is the main advantage of machine learning, especially compared to traditional statistical analysis. Textbooks usually categorise machine learning algorithms based on their learning methods. Marsland (2015) classifies Machine Learning algorithms into four types: supervised learning, unsupervised learning, reinforcement learning and evolutionary learning (Table 3.1). In this thesis, only two types of learning algorithms were applied, supervised learning and evolutionary learning (evolutionary algorithm).

3.1 Supervised learning

The most common type of learning is supervised learning; many real-life applications were based on this type of learning. In any learning algorithm, a set of data is used as the training data. In supervised learning, the training data consists of a set of input data that is mapped to its output; which is the answer that the algorithm should produce.

Table 3.1: Types of machine learning algorithms as defined by Marsland (2015).

Algorithm	Definition
Supervised learning	A training set of examples with the correct responses (targets) is provided and, based on this training set, the algorithm generalises to respond correctly to all possible inputs
Unsupervised learning	Correct responses are not provided, but instead the algorithm tries to identify similarities between the inputs so that inputs that have something in common are categorised together
Reinforcement learning	The algorithm gets told when the answer is wrong but does not get told how to correct it. It must explore and try out different possibilities until it works out how to get the answer right.
Evolutionary learning	Algorithm that is based on idea that biological evolution is a learning process: biological organisms adapt to improve their survival rates and chance of having offspring in their environment.

When an underlying function from inputs to outputs exists, it is referred to as the *target function*. The estimate of the target function output by the learning algorithm is known as the *solution* of the learning problem. In the case of classification this function is sometimes referred to as the *decision function*. The solution is chosen from a set of candidate functions which map from the input space to the output domain (Cristianini & Shawe-Taylor, 2000).

There is an overwhelming range of supervised learning algorithms to choose from. However, until now there is no 'one algorithm fits all' solution. Each algorithm has its

advantages and disadvantages and users need to consider characteristics of the problem in-hand to decide on a suitable algorithm. Two types of very popular supervised learning algorithms, Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) are used in this study to provide a comparison of performance with the algorithm under investigation.

Artificial Neural Networks (ANNs)

ANNs are learning algorithms that base their learning on how the human brain processes information. ANNs learn (or are trained) through experience as in the human neural network, gain their knowledge by identifying the patterns and relationships in the data. ANNs have been used for a variety of tasks including the construction of models for prediction, clustering and classification. Basically, an ANN is a network of simple processing units called neurons. Signals or influences can only pass in one direction along given connection (also called arc or edge). The effect of the signal along a connection may be adjusted by a weight. This means that ANNs are weighted directed graphs. Each node processes the combination of weighted signals presented to it, in a manner that varies according to the type of an ANN. There are many types of ANNs, they can be differentiated from each other by looking at following criteria:

i. connection topology,
ii. basis function,
iii. training method, and
iv. learning algorithm.

The **connection topology** defines how the processing units (nodes) are connected to each other.

The **basis functio**n defines what processing each node carries out on the combination of all its inputs, in order to generate its output value.

The **training method** is concerned with how the ANN learns. ANNs can be used in supervised learning environment where the ANN is provided with training data (input data for which the output is already known), learn from it and try to predict the correct output for inputs not given in the training data. It can also be used in unsupervised learning when the ANN is not provided with outputs, but rather, is left to uncover patterns in the input data without a priori information as to what these patterns may be (Brabazon et al., 2015). The main drawback of ANNs is their black box approach towards problem solving. An ANN does not provide more information on how the learning problem was solved. Symbolic rules may be extracted from trained ANNs, but other algorithms such as genetic programming provides better option in this area, as it is truly white box in its operation providing a discrete mathematical expression describing how data inputs are used in the resulting classification.

Support Vector Machine (SVM)

SVMs are classification algorithms that are based on statistical learning theory. It is a very popular supervised learning methodology in modern machine learning and perhaps the most used one. It can be used for classification, regression and prediction and are particularly suited to binary classification, where there are only two classes to the problem. SVMs aim for good fitting of training data, while avoiding overfitting, so that the solution generalises well to new instances of data (Brabazon et al., 2015). SVMs work on the concept of linearly separating data into their correct classes. The linear boundaries between classes are generally known as hyperplanes: linear or affine subspaces of dimension $n - 1$, where n is the dimension of the space of linearly separable classes.

Hyperplanes can be straight line in case of 2D space (Figure 3.1) or a plane in 3-D space, or a 3-D space in 4-D space. A hyperplane splits the n- dimensional space into two parts, in the positive and negative directions along the nth dimension either side of the hyperplane.

Figure 3.1: Hyperplane: a line that linearly separate two-class dataset (left) compared with a nonlinearly separable two-class dataset (right). Figure from Brabazon et al., (2015)

Figure 3.2 a) shows a simple example of a dataset that can be divided into two classes (labels × and +) which are not linearly separable. However, when mapped into the higher dimensional 'feature space', the classes are now linearly separable (Figure 3.2 b))

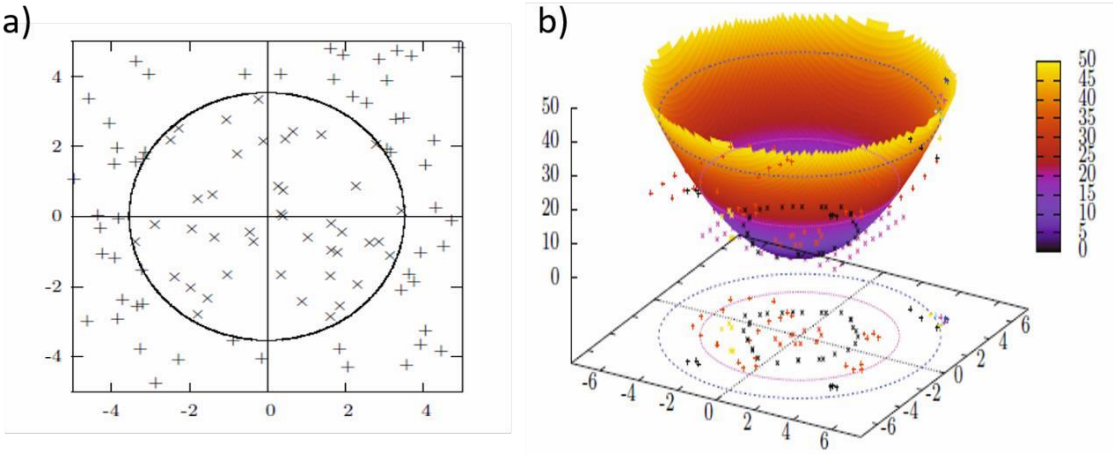

Figure 3.2: a) Two-class dataset that is not linearly separable. b) When the dataset i (a) transformed to higher dimension, it is now separable. Figure from Brabazon et al. (2015)

The mapping of data into a higher dimension is known as kernelling. The kernel function in SVM transforms the messy pattern or input vector space to the higher dimensional feature space by a nonlinear mapping; the classes in the feature space may be linearly separable even if they were not in the original pattern space. In implementing SVMs, choosing an appropriate kernel function and choosing good values for its associated parameters is very important.

SVMs has been implemented in many real-life applications such as text classification tasks, handwritten digit recognition, category assignment, detecting spam and sentiment analysis, among others (Hearst et al.,1998; Ma & Guo, 2014). In the medical field, it has been used extensively for medical image recognition and classification of diseases (Statnikov et al., 2011) . However, SVMs have acknowledge drawbacks (Brabazon et al., 2015):

i. It is difficult to incorporate domain knowledge into an SVM, other than in the data pre-processing or kernel selection steps.
ii. The rationale for the resulting classification decisions can be hard to reverse-engineer, as the support vectors provide limited information to the modeller.
iii. SVMs were originally designed to work with real-valued vectors, so there is no unique way to incorporate noncontinuous data (for example, categorical data) into an SVM.
iv. The SVM methodology also requires that data vectors be scaled, and different methods of scaling can produce different results.

3.2 Evolutionary Algorithms

The Evolutionary Algorithm is a computational intelligence technique inspired by natural evolution and survival of plants and animals. The original basis of the technique is Darwin's theory of natural selection. Biologically inspired computation is one of the fields inspired by Darwin's theory which is considered to be one of the most prominent and important works in the history of human civilization (Serrano & Castillo, 2012) . The notion suggests that all living things on earth today are results of millenniums of adaptations to natural environment. It is based on observation that in an ecosystem, different organisms must share resources. The most skilled organism in securing the resources and reproduction are most likely to have surviving offspring. Whilst organisms that are less skilled may have none or only few descendants left in the future. This was summarised clearly in a chart by Gregory (2009) - see Figure 3.3. The concept is known as 'the fittest survive' or 'the survival of the fittest' (Eiben & Smith, 2003). EAs mimic this part of nature to optimize a solution to a predefined problem.

3.2.1 EAs basic elements and components

Since the 1960s, many algorithms with similar properties have been proposed that resulted in several different types and versions of EAs. The major ones are genetic algorithms, evolutionary programming and genetic programming. However, almost all EAs share a number of common elements (Coello et al., 2007; Yu & Gen, 2010) :

a) Population-based. EAs maintain a group of solutions, called a population (of individuals), to optimize or learn the problem for potential solutions, instead of a single candidate solution at a time.

b) Fitness-oriented. EAs used a selection method biased by fitness. Every individual has its performance evaluation, called its fitness value. The better the fitness of an individual, the more often it is selected and the more some parts of its "genetic material" (parts of its candidate solution) will be passed on to later generations of individuals.

c) Variation-driven. New individuals based on the selection of older ones are generated by randomised processes intended to model mutation and recombination & crossover.

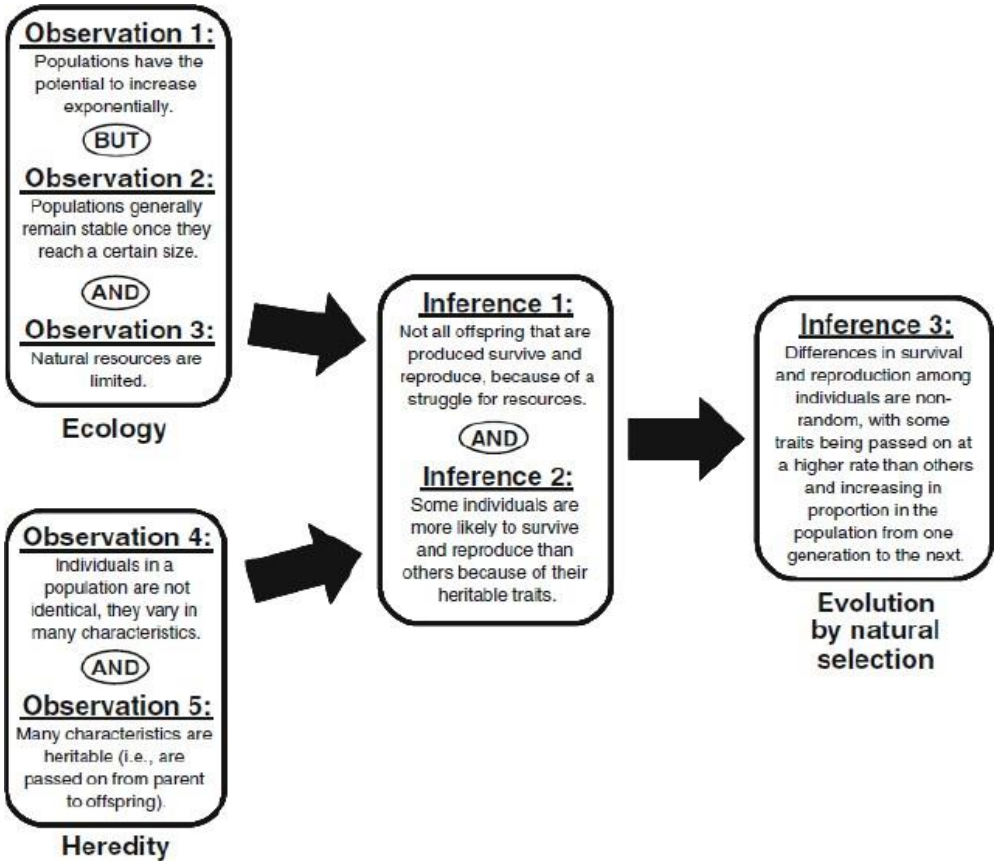

Figure 3.3: The Basis of Natural Selection by Darwin.

Basically, EAs try to find solutions by going through repetitive steps of selecting the best candidate based on its fitness. Initially, given a problem to solve, a set of candidate

solutions is randomly generated. The set is known as a *population* and the candidate as an *individual* to that population. Every individual in the population is assigned, by means of a fitness function, a measure of its goodness with respect to the problem under consideration. Based on this fitness, some of the better candidates are chosen to seed the next generation by applying crossover and/or mutation to them. In essence, crossover swaps some genetic material between two or more individuals, while mutation changes the value of a small part of the genetic material of an individual to a new random value, simulating self-replication of individuals. Executing crossover and mutation leads to a set of new candidates (the offspring) that compete - based on their fitness (and possibly age) - with the old ones for a place in the next generation. This process can be iterated until a candidate with sufficient quality (a solution) is found or a previously determined computational limit is reached. In the form of pseudo code, the process can be written as in Figure 3.4 (figure from Eiben & Smith (2003)). The key components and steps are explained graphically in Coello et al. (2007) (Figure 3.5).

```
BEGIN
  INITIALISE population with random candidate solutions;
  EVALUATE each candidate;
  REPEAT UNTIL ( TERMINATION CONDITION is satisfied ) DO
    1 SELECT parents;
    2 RECOMBINE pairs of parents;
    3 MUTATE the resulting offspring;
    4 EVALUATE new candidates;
    5 SELECT individuals for the next generation;
  OD
END
```

Figure 3.4: The general scheme of EA in pseudocode.

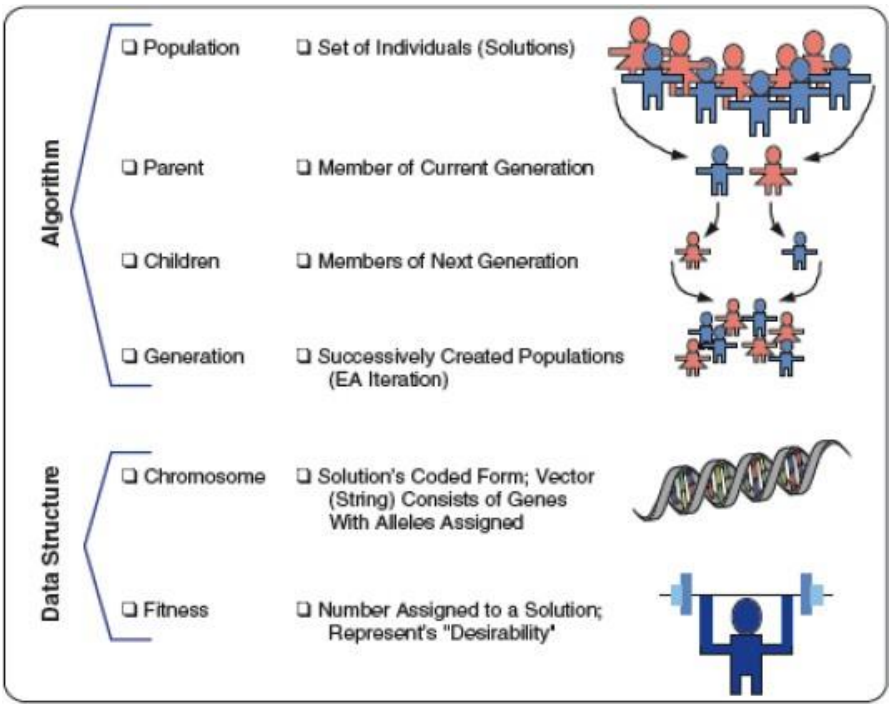

Figure 3.5: Important elements of EA

3.2.2 Applications of evolutionary algorithms

Many activities involve unstructured, real life problems are difficult to model, since they require several unusual factors. Certain engineering problems are complex in nature: job shop scheduling problems, timetabling, traveling salesman or facility layout problems. For all these applications, evolutionary computation provides a near-optimal solution at the end of an optimisation run. Evolutionary algorithms are thus made efficient because they are flexible, and relatively easy to hybridize with domain-dependent heuristics.

Applications of evolutionary computation include the following fields:

- Medicine and healthcare
- Engineering application (including electrical, mechanical, civil, production, aeronautical and robotics).
- Traveling salesman problem.
- Machine intelligence.

- Expert system
- Network design and routing
- Wired and wireless communication networks.

(Sivanandam & Deepa, 2008)

Evolutionary algorithms in Medicine and Healthcare

Since EAs were developed decades ago, they have been used to solve problem in medicine. EAs are considered as an alternative approach to medical quantitative studies (that employ statistical methods) (Smith, 2011). Examples of EA applications in medical and healthcare include:

- Medical imaging: image segmentation, image registration, reconstruction and correction, detection of breast cancer.
- Data mining medical data and patient records
- Clinical expert systems and knowledge-based systems: computer aided detector for breast cancer, decision support system for the diagnosis and classification of heart disease.
- Modelling and simulation of medical processes
- Clinical Diagnosis and Therapy: automated heart disease diagnosis, diagnosis of Parkinson's disease, diagnosis of language impairments in speech pathology, diagnosis of Alzheimer's disease, drug design, medicinal chemistry and predictive toxicology.

(Gandomi et al., 2015; Smith & Cagnoni, 2011)

3.3 Genetic Programming

Genetic Programming (GP) is a well-known branch of EAs. It became more widely adopted after the publication of John Koza's book in 1992 (Koza, 1998). The most important feature of GP is its ability to automatically generate functional programs from a high-level statement of the problem. This addresses one of main challenges of computer science that is to get a computer to do what needs to be done, without telling it how to do it. This idea can be expanded to generate artificial intelligence by computer (Yu & Gen, 2010).

GP may be defined generally as any direct evolution or breeding of computer programs for the purpose of inductive learning. In particular, this definition leaves GP independent of a special type of program representation (Brameier & Banzhaf, 2007). GP is closely related to genetic algorithms (GA) and based on the same Darwin's principal of "survival of the fittest". Only in the case of GP, the initial population consists of computer programs which are generated randomly. However, three important differences exist between GAs and GP (Sivanandam & Deepa, 2008) :

Structure: GP usually evolves tree structures while GAs evolve binary or real number strings.

Active vs Passive: Because GP usually evolves computer programs, the solutions can be executed without post processing i.e. active structures, while GA's typically operate on coded binary strings. i.e. passive structures, which require post-processing.

Variable vs fixed length: In traditional GAs, the length of the binary string is fixed before the solution procedure begins. However, a GP parse tree can vary in length throughout

the run. Although it is recognised that in more advanced GA work, variable length strings are used.

3.3.1 General form of GP

The representation used by genetic programming is in the form of executable computer programs. There are many different forms of computer programs which are used in genetic programming implementations, but most of them use a tree-structured representation. (Other representations commonly used will be introduced later in this section.) Usually there are two types of nodes in representing individuals: Terminals and Functions.

Terminals are inputs to the program; they can be constants or variables. Figure 3.6 shows an example of a tree-structured representation for a genetic programming implementation. In the example a, b, y and integers are terminals.

Functions take inputs and produce outputs and have some possible side-effects. The inputs can be terminals or the output of other functions. In the above example +, /, *,- and sin are functions. The functions may be standard arithmetic operations, standard mathematical functions, logical functions, standard programming functions or domain specific functions.

Genetic Operators of reproduction, crossover and mutation are three basic operators in genetic programming (Koza, 1998).

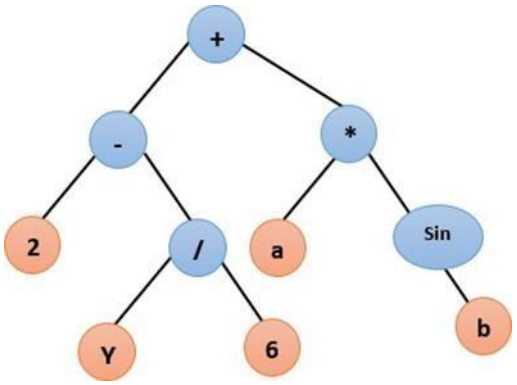

Figure 3.6: An example of tree-structured GP representation.

Reproduction: The reproduction operation for genetic programming is asexual operation where it operates on only one parent and produces only one offspring. The operation of reproduction consists of two steps. First, a single parent is selected from the population according to some selection method based on fitness. Second, the selected individual is copied, without alteration, from the current population into the new population

Crossover: The crossover operation generates a new offspring that consists of parts of its parents. In Figure 3.7(a) there are two parent: A and B. Each parent independently selects a point for crossover operation randomly. Then two offspring will be generated by swapping two selected fragments of parents to produce children Figure 3.7 (b).

Mutation: In genetic programming a point is selected at random within a selected individual. The mutation operation then removes whatever is currently at the selected point and whatever is below the selected point and inserts a randomly generated sub-tree. Another method is to randomly find a function node anywhere on the tree and then replace the function with another selected at random, from the set of functions, where the function takes the same number of arguments.

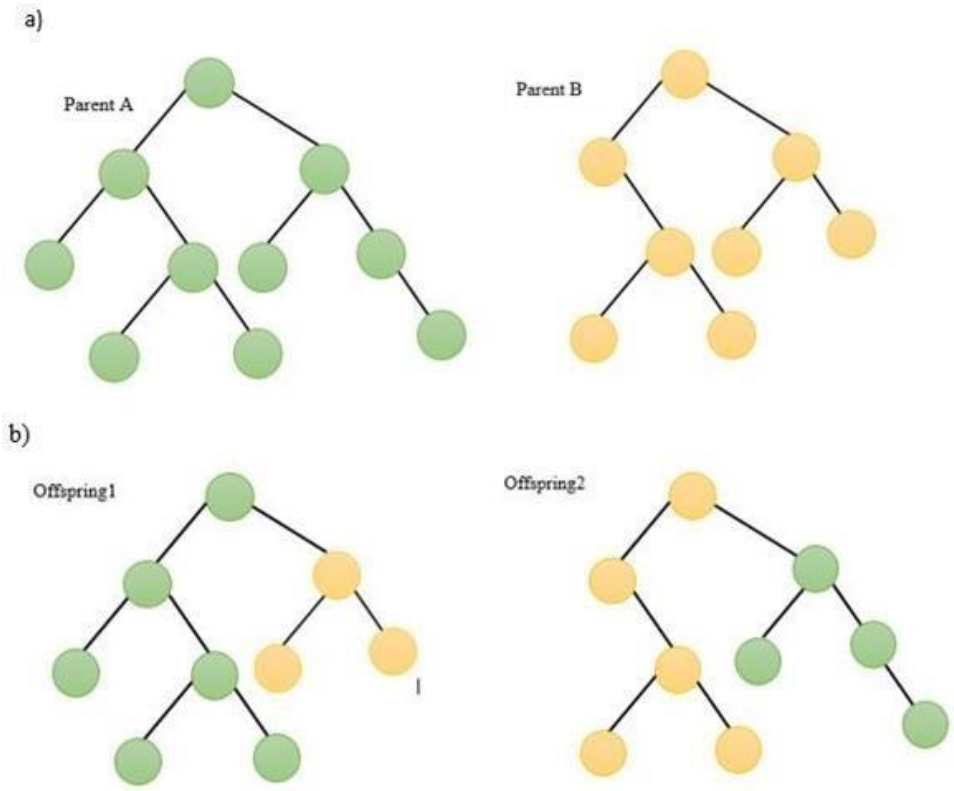

Figure 3.7: Crossover operation.

In general, the process of GP can be executed by the following three steps (Koza, 1998)

(1) Generate an initial population of random compositions of the functions and terminals of the problem (computer programs).

(2) Iteratively perform the following sub steps until the termination criterion has been satisfied:

a) Execute each program in the population and assign it a fitness value according to how well it solves the problem.

b) Create a new population of computer programs by applying the following two primary operations. The operations are applied to computer program(s) in the population chosen with a probability based on fitness.

(i) Copy existing computer programs to the new population.

(ii) Create new computer programs by genetically recombining randomly chosen parts of two existing programs.

(3) The best computer program that appeared in any generation (i.e., the best-so-far individual) is designated as the result of genetic programming. This result may be a solution (or an approximate solution) to the problem. The whole cycle visualised in Figure 3.8 adapted from Langdon & Poli (2002).

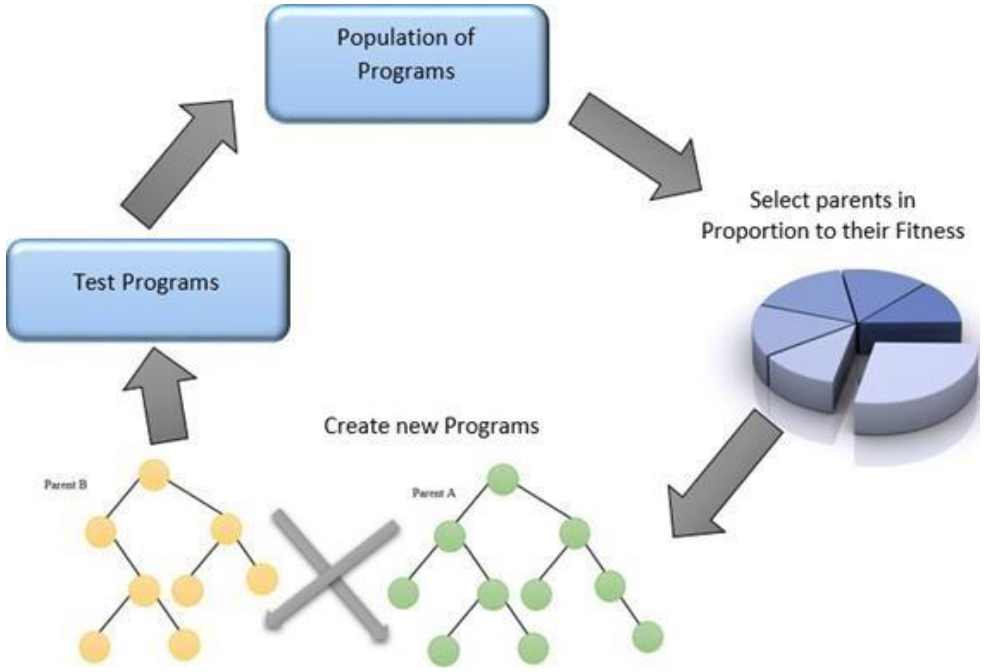

Figure 3.8: Genetic Programming cycle.

3.3.2 Other representation of Genetic Programming

GP is traditionally represented by a tree structure as explained earlier. However, non-tree representations have been suggested and successfully implemented. The other two popular representations of GP are linear and graph representations (Banzhaf et al., 1998).

Linear Genetic Programming (LGP)

LGP is where computer programs in a population are represented as a sequence of instructions from machine language. In linear GP programs are linear sequences of

instructions, as shown in Figure 3.9. The number of instructions can be fixed or varies. It means that all the programs can have the same length or that different individuals can have different size (Poli et al., 2008).

Figure 3.9: Typical LGP representation.

As in machine language instruction execution, in linear GP, instructions read their input(s) from one or more registers or memory locations and store the results of their calculations in a register. Instructions in linear GP all have equivalent roles and communicate only via registers or memory. In linear GP there is no equivalent of the distinction between functions and terminals which is fundamental in tree-based GP. Also, in the absence of loops or branches, the position of the instructions determines the order of their execution.

Originally there are two types of linear GPs: machine code GP, where each instruction is directly executable by the CPU, and interpreted linear GP, where each instruction is executable by some higher-level virtual machine (typically written in an efficient language such as C or C++). When the instructions are actual machine code, then the order of the elements of the representation shown in Figure 3.8 is determined by the particular computer architecture used, and the corresponding data must be packed into bit fields of appropriate sizes. On the contrary, when one is using virtual machine instructions the designer of a GP system has complete freedom as to how the virtual machine will interpret its instructions. If the goal is execution speed, then the evolved code should be machine code for a real computer rather than some higher-level language or virtual machine code (Poli et al., 2008).

Graph-based Genetic Programming

Starting from the middle 1990s, researchers have proposed several extensions of GP that able to evolve graph-like programs. A few examples of variants for this branch of GP are Parallel Distributed GP (PDGP), Parallel Algorithm Discovery and Orchestration (PADO), and Cartesian GP. Riccardo Poli in Poli (1996) proposed an approach called parallel distributed genetic programming (PDGP). Poli stated that PDGP can be considered as a generalization of GP. However, PDGP has more complex representations and evolves finite state automata, neural networks and more. In PDGP programs are represented in as graphs with nodes representing functions and terminals.

In a system called parallel algorithm discovery and orchestration (PADO), a combination of GP and linear discrimination was used to obtain parallel classification programs for signals and images. The actions are drawn from a primitive set including the standard algebraic operations, minimum, maximum, negation, read from indexed memory; write to indexed memory, deterministic and non-deterministic branching instructions, and primitives related to the task of classifying images (Poli et al., 2008). The traditional tree representation of programs is the root of GP, even though many different GP approaches and program representations exist. A general motivation for investigating different representations in evolutionary computation is that for each representation form, as is the case for different learning methods in general, certain problem domains may exist that are more suitable than others (Brameier & Banzhaf, 2007).

3.3.3 Classification applications of GPs

Daily, huge amounts of data are being stored in databases around the world. This creates both an opportunity and a need to discover the knowledge in such databases. If such knowledge discovery activity is successful, discovered knowledge can be used to improve

the decision-making process of an organisation (Freitas, 2010). Research fields such as statistic and machine learning have been used to gain the knowledge from real- world data sets. It is also known as data mining. Classification is probably the most studied data mining task and EA is an evolutionary learning technique that offers a great potential for classification (Espejo et al., 2010). Classification can serve many different purposes, like credit scoring, bankruptcy prediction, medical diagnosis, and so many more.

The advantages of application of GP to classification task include:

- Flexibility, which allows the technique to be adapted to the needs of each problem.
- GP can be employed to construct classifiers using different kinds of representations.
- GP can be useful not only for inducing classifiers, but also for other pre-processing and post processing tasks aimed at the enhancement of classifiers.
- Interpretability, GP can employ more interpretable representation formalisms, like rules and automatic feature selection.

In the classification task each data instance (or database record) belongs to a class, which is indicated by the value of a goal attribute. This attribute can take on a small number of discrete values, each of them corresponding to a class. Each instance consists of two parts, namely a set of predictor attribute values and a goal attribute value. The former are used to predict the value of the latter. The predictor attributes should be relevant for predicting the class (goal attribute value) of a data instance. For example, if the goal attribute indicates whether a patient has or will develop a certain disease, the predictor attributes should contain medical information relevant for this prediction.

In a simplified version of the classification task, first the set of data instances is randomly divided into two subsets, called the training set and the test set. The training set is made entirely available to the search algorithm, so that the algorithm has access to the

values of both predictor attributes and the goal attribute for each data instance. The aim of the algorithm is to discover a relationship between the predictor attributes and the goal attribute using the training set. In order to discover this relationship, the algorithm has access to the values of both predictor attributes and the goal attribute for all instances of the training set. The discovered relationship is then used to predict the class (goal-attribute value) of all the data instances in the test set.Apart from construction of a classifier from a dataset, there are some other related tasks that usually have to be addressed as well such as data pre-processing and model interpretation and enhancement (Espejo et al., 2010).

Data pre-processing. Model extraction techniques are not usually applied to the database in its original form. There are a great variety of pre-processing techniques available in order to prepare the data to take advantage of its maximum potential. Some of the issues to consider are such as the removal of noise or outliers, strategies for handling missing data, feature selection and construction, instance selection, data rebalancing and data projection, and normalisation.

Model interpretation and enhancement. Sometimes the classifier obtained is not readily usable, and some kind of post-processing is necessary. Some of the issues to be addressed at this point may include the elimination of redundant knowledge, application of visualisation techniques, evaluation of the model obtained, translation to a more interpretable form, combination of models, and any other way of improving the model obtained.

3.4 Cartesian Genetic Programming

Cartesian genetic programming (CGP), introduced by Miller and Thomson (Miller & Thomson, 2000), is a form of genetic programming where the candidate solutions are represented as a string of integers of fixed length that is mapped to a directed oriented graph. CGP can efficiently represent common computational structures including mathematical equations, computer programs, neural networks and digital circuits. CGP is Cartesian in the sense that the method considers a grid of nodes that are addressed in a Cartesian coordinate system. Although CGP is relatively new to other GPs, it has shown great potential with a wide range of applications.

> *"Cartesian Genetic Programming (CGP) is now attracting considerable recognition as an evolutionary algorithm that not only delivers high performance, but one that has a representation that is flexible and easy to adapt to a range of applications. "*
>
> (Smith, 2011)

3.4.1 General form of CGP

The graphs in CGP are represented as a two-dimensional grid of computational nodes. The genes that make up the genotype in CGP are integers that represent where a node gets its data, what operations the node performs on the data and where the output data required by the user is to be obtained. When the genotype is decoded, some nodes may be ignored. This happens when node outputs are not used in the calculation of output data. When this happens, the nodes and their genes are referred as 'non-coding'. Programs that result from the decoding of a genotype is called phenotype. The genotype in CGP has a fixed length. However, the size of the phenotype (in terms of the number of computational nodes) can be anything from zero nodes to the number of nodes defined in the genotype.

A phenotype would have zero nodes if all the program outputs were directly connected to program inputs. A phenotype would have the same number of nodes as defined in the genotype when every node in the graph was required. The genotype–phenotypemapping used in CGP is one of its defining characteristics.

The types of computational node functions used in CGP are decided by the user and are listed in a function look-up table. The function gene is the address of the computational node function in the function look-up table. The connection genes determine where the node gets its data from. These genes represent addresses in a data structure (typically an array). Nodes take their inputs in a feed-forward manner from either the output of nodes in a previous column or from a program input (also known as a terminal). The number of connection genes a node has is chosen to be the maximum number of inputs (often called the arity) that any function in the function look-up table has. The program data inputs are given the absolute data addresses 0 to ni - 1 where *ni* is the number of program inputs. The data outputs of nodes in the genotype are given addresses sequentially, column by column, starting from ni to ni +Ln −1, where Ln is the user-determined upper bound of the number of nodes.The general form of a Cartesian genetic program is shown in Figure 3.10 (figure from Miller (2011)). If the problem requires no program outputs, then no integers are added to the end of the genotype. In general, there may be several output genes (Oi) which specify where the program outputs are taken from. Each of these is an address of a node where the program output data is taken from (Miller, 2011).

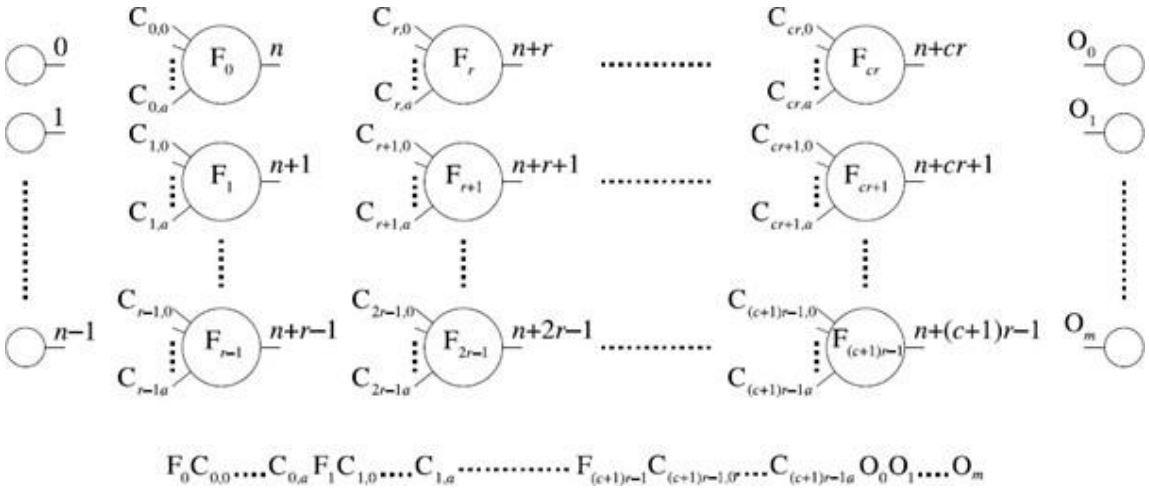

Figure 3.10: General form of CGP.

3.4.2 Encoding

CGP encodes a candidate solution (typically a circuit or a program) using an array consisting of c x r programmable nodes. The c determines the number of columns and r determines the number of rows. These are two out of three parameters that are chosenby the user. The third one is levels back, denoted by l. The product of the first two parameters determine the maximum number of computational nodes allowed: Ln = ncnr. The parameter l controls the connectivity of the graph encoded. Levels-back constrains which columns a node can get its inputs from. If l = 1, a node can get its inputs only from a node in the column on its immediate left or from a primary input. If l = 2, a node can have its inputs connected to the outputs of any nodes in the immediate left two columns of nodes or a primary input. Each node input can be connected either to the output of a node placed in the previous l columns or to one of the program inputs. Because of the complicated evaluation, feedback is not allowed in the standard version of CGP. The main feature of CGP is that all the parameters including the number of programmable nodes is fixed. It means that the array of programmable nodes can be encoded as a string of integers which has the fixed number of items. The main advantage of CGP encoding is that even if the

size of chromosome is fixed, the size of phenotype is variable since some nodes need not to be used.

An example of 2 x 2 nodes CGP network is shown in Figure 3.11 from Smith, (2011). Inputs, I/P 0 and I/P 1 feed the input values to the network, manipulated by function in each node and deliver the result to the output, O/P 0. The reference number of each node is shown in the top right-hand corner of each node. Table 3.2 shown example of a look-up table that listed user defined functions of the nodes in the network.

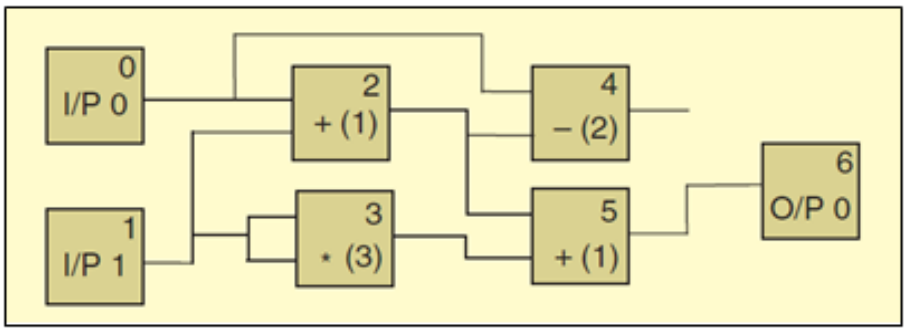

Figure 3.11: Example of a CGP network.

Table 3.2: Example of function set look-up table.

Function Reference	Function
1	/
2	*
3	-
4	+

3.4.3 Evolution of CGP genotypes

CGP uses two common EAs operators; mutation and crossover /recombination. The former is mostly used compare to the latter.

Mutation

A point mutation operator is used in CGP. In a point mutation, an allele at a randomly chosen gene location is changed to another valid random value. If a function gene is chosen for mutation, then a valid value is the address of any function in the function set, whereas if an input gene is chosen for mutation, then a valid value is the address of the output of any previous node in the genotype or of any program input. Also, a valid value for a program output gene is the address of the output of any node in the genotype or the address of a program input. The number of genes in the genotype that can be mutated in a single application is determine by a user defined mutation rate, μr which is normally a percentage of the total number of genes in the genotype. An example of point mutation is shown in Figure 3.12 (Figure from Miller (2011)).

Crossover

Recombination operators have received relatively little attention in CGP. Originally, a one-point crossover operator was used in CGP (similar to the n-point crossover in genetic algorithms) but other studies also suggest improvement to the performance of CGP by using floating-point crossover operator (Clegg et al.,2007).

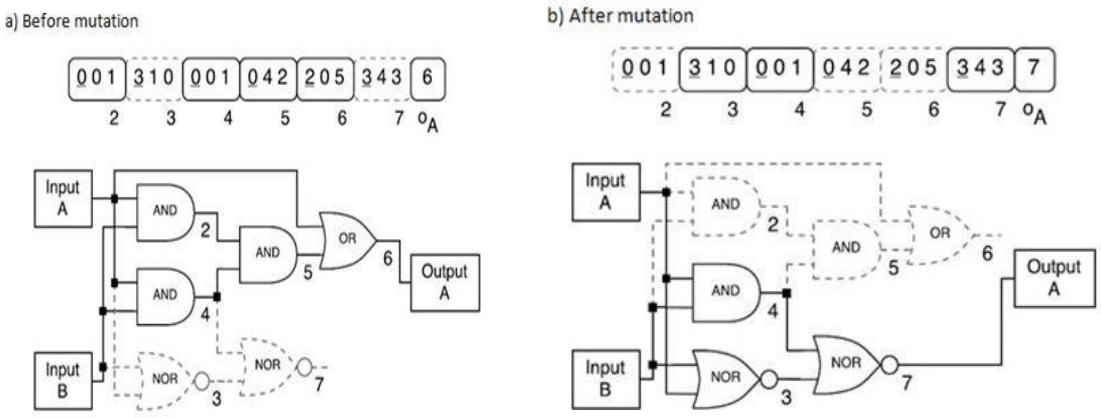

Figure 3.12: Point Mutation process.

3.4.4 Evolutionary strategy

CGP usually uses a variant of a simple algorithm called (1 + 4). Figure 3.13 shown the procedure of 1 + 4 strategy (from (Miller, 2011)). One of the conditions of the algorithm is an offspring is always chosen as a new parent if it is equally as fit or has better fitness than the parent (Figure 3.14). When offspring genotypes in the population have the same fitness as the parent and there is no offspring that is better than the parent, in that case an offspring is chosen as the new parent (Figure from Miller & Thomson (2000)).

1: **for all** i such that $0 \leq i < 5$ **do**
2: Randomly generate individual i
3: **end for**
4: Select the fittest individual, which is promoted as the parent
5: **while** a solution is not found **or** the generation limit is not reached **do**
6: **for all** i such that $0 \leq i < 4$ **do**
7: Mutate the parent to generate offspring i
8: **end for**
9: Generate the fittest individual using the following rules:
10: **if** an offspring genotype has a better or *equal* fitness than the parent **then**
11: Offspring genotype is chosen as fittest
12: **else**
13: The parent chromosome remains the fittest
14: **end if**
15: **end while**

Figure 3.13: : The (1+4) evolutionary strategy.

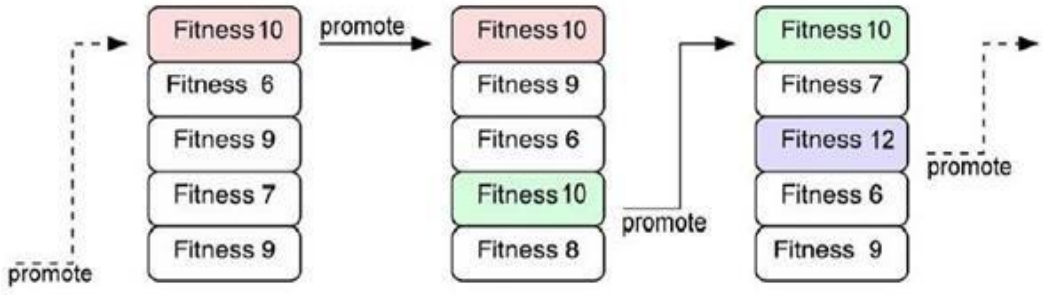

Figure 3.14: Priority to offspring with same fitness as parent.

3.4.5 CGP medical applications

There are many examples where CGP and its variants have been used in classification tasks involving medical data. One of the early examples is a work that classifies mammogram data in detection of breast cancer (Hope et al., 2007). Also, on detection of breast cancer, a CGP variant, Cartesian genetic programming evolved artificial neural networks (CGPANN) have been used in Ahmad et al. (2012) to determine whether a breast cancer is ‘‘malignant’’ or ‘‘benign’’ based on features of breast mass. Other examples are classifications for diagnosis of Alzheimer’s disease (Hazell & Smith, 2008), classification of arrhythmia for Cardiovascular diseases (Ahmad et al., 2013) discriminating normal and cancerous thyroid cell lines (Lones et al., 2010), classifications of PD patients from healthy controls based on movement data (Lones et al., 2014) and many more.

3.5 Summary

Machine learning is not a new founded field, yet it continues to expand with algorithms inspired by new inspirations from biology and nature. Consequently, there are a great number of algorithms to choose from for problem solving. In a classification task, popular algorithms such as Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs) usually perform very well. However, the ANN is a 'black box' that makes understanding of how the problem was solved, difficult to achieve. SVM also does not provide rationale for the resulting classification decisions. This study's main objectives are to differentiate and characterise movement disorders. Although popular algorithms such as SVM and ANN can differentiate very well, they cannot help with characterisation to the same degree. An evolutionary algorithm such as Genetic Programming (GP) is more suitable for this task. In this thesis, the main algorithm used is Cartesian Genetic Programming (CGP). CGP is a type of graph-based genetic programming with increasing applications across many aspects of real life. CGP typically finds good solutions efficiently in few evaluations. Although it uses many generations, it also uses extremely small populations, of typically five individuals, where one individual is the fittest from the previous generation. Other advantages of CGP is that it does not bloat like other forms of GP and sub functions can be easily reused in the Cartesian representation. CGP has been shown to successfully classify and characterised medical data in previous examples, providing further justification for its use in this study.

CHAPTER 4

Methodology

This chapter describes the methodology adopted in obtaining measurements from patients in the respective clinical studies undertaken in medical centres as part of ongoing research in collaboration with University of York .

Participants with different neurological disorders and healthy adults (controls) were recruited by qualified clinicians in two medical centres:

- Leeds Teaching Hospitals NHS Trust, UK
- Monash Medical Centre, Melbourne, Australia

All patients have an established clinical diagnosis according to the relevant diagnostic criteria, established at the respective medical centres. Controls are healthy adults with no expression or history of a medical disorder that could affect movement.

4.1 Data Collection

Data collections were administered in line with the protocols written for the study under investigation but share common clinical settings in comparable testing environments for clinical evaluation. All subjects provided informed written consent and approval was

obtained from Yorkshire and Humber Sheffield Research Ethics Committee (HREC code: 14/YH/0143).

4.1.1 Equipment

The Polhemus Patriot EM tracking sensor system, a motion tracking system is used for data collection. The Polhemus device comprises a system electronics unit (SEU), a magnetic transmitter and two electromagnetic (EM) tracking sensors (Figure 4.1). The transmitter serves as the system's reference frame where the EM sensors record positions and orientations relative to the transmitter in six degrees of freedom with an update rate of 60 Hz per sensor.

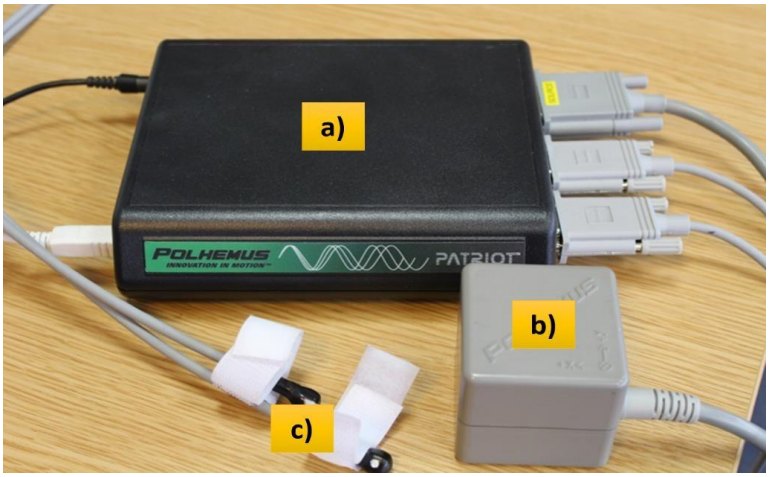

Figure 4.1: Polhemus Patriot system (www.polhemus.com), a motion tracking system used to record movements during the motor assessment. a) SEU, b) magnetic transmitter cube and, c) EM sensors.

The collection of data was administered in a clinical setting as a common practice of motor examination of MDS-UPDRS. The electromagnetic sensors (EM sensors) were attached to the thumb and index finger using Velcro tapes (Figure 4.2). For consistency, the sensors were attached at the same places during the administration of all the three tasks. The Polhemus device is based on varying magnetic fields. Large metal objects can affect the fields, and this adversely affects the results. Therefore, care was taken to

examine the effect of metal objects used in performing the tasks. The position of subjects and their distance from the Polhemus transmitter cube during measurement is also important.

Figure 4.2: EM sensors were attached on nail beds of index finger and thumb.

This is due to the manufacturer recommended maximum distance from the sensors to the transmitter is about 150 cm; beyond this limit, the recordings will be interrupted. To make sure the positions of sensors were always within the working distance, clinicians used a fixed arrangement for the subject and the equipment as shown in Figure 4.3.

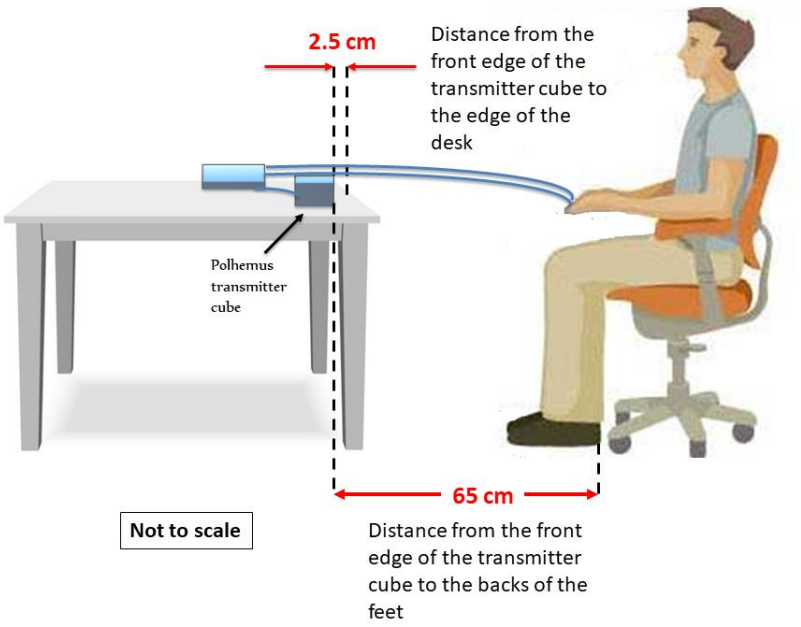

Figure 4.3: Arrangement of the motion tracking system and participant position during movement recordings.

After the sensors were fastened to the thumb and index finger, the participant was requested to sit straight with their back to the chair. Next, they were asked to lift their dominant hand unsupported so that it was around the level of their shoulder. The clinician demonstrates the correct way to perform the task, but the demonstration was not carried on during the execution by the participants. The participants were not allowed any practice movement; the first attempt was recorded as the experimental data. The recording was only repeated if there were technical problems such as the system was not recording, or the output file was not saved automatically to the storage or the sensors were not at correct positions. After the measurement the dominant hand was repeated two times, the sensors then were attached to the non-dominant hand to complete the same measurement. The Patriot system allows for customisation of the interface (Figure 4.4) where the tasks were sequenced along with instructions on how to execute the task. Before the recording started, clinicians can enter demographic and medical information of the participant.

When the recording of movements was completed, the system stores the data file at a designated location with a pre-determined file name that combined unique participant code along with date and time of the assessment. The created interface and automatic storage system ensured clinicians (involved in this study) in different centres to follow the same protocol of the assessment and data storage management. Encrypted data files were hand-delivered or transmitted electronically from the test centres to the University of York to be analysed.

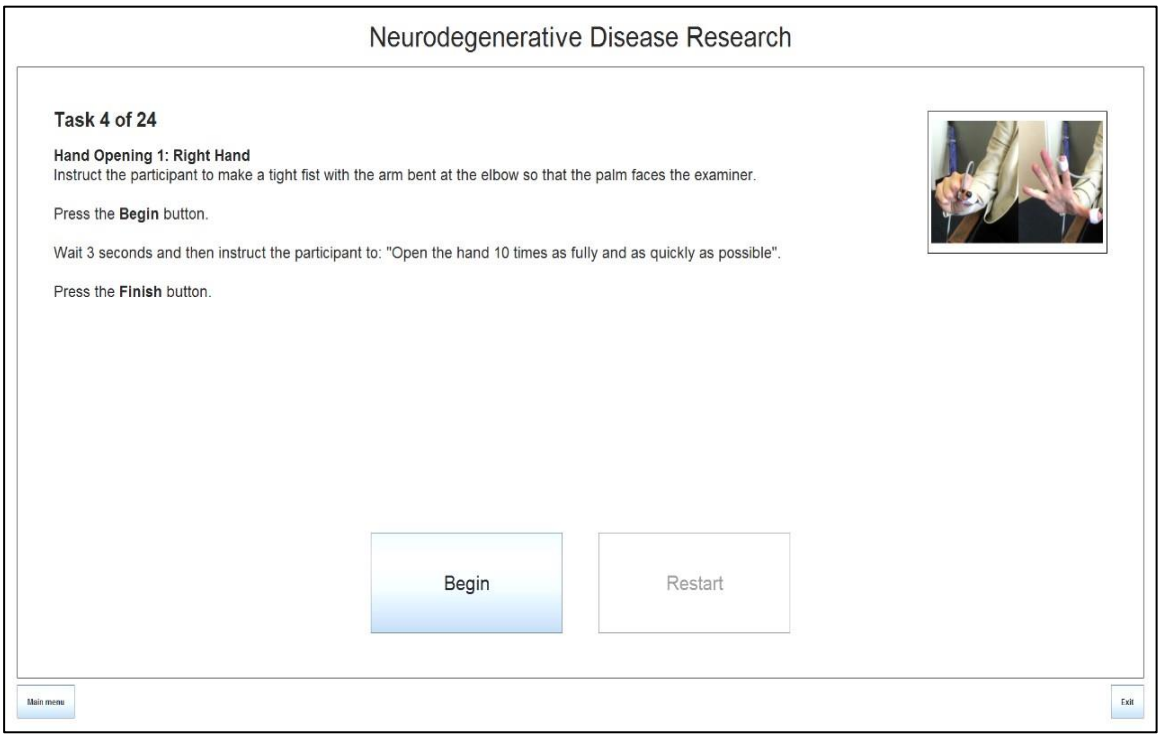

Figure 4.4: An example of the interface created to ensure the same protocol was followed by clinicians from different centres.

4.1.2 Movement tasks

The Movement Disorders Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) was used as the conventional clinical evaluation from which participant measurements were to be captured. The MDS-UPDRS has four sections:

Part I: non-motor experiences of daily living

Part II: motor experiences of daily living

Part III: motor examination, and

Part IV: motor complications

Part I, II and IV are based on the history taking of the subjects by the clinician and patients personal notes (if any). Part III, also referred to as the motor section, is based on the examination by the clinician. This rating scale evaluates the severity of PD symptoms in a 5-points scoring system (0 for no symptom and 4 for a marked severity of the symptom).

For this study, subjects performed the following MDS-UPDRS items

- Finger tapping
- Hand opening-closing (written as hand movements in MDS-UPDRS)
- Pronation/Supination

4.2 Computation of movement features

Based on the literature and direct consultations with clinicians involved in the research, movement features were extracted from the recordings and used as input to the classifiers. The nature of the movement in each task also considered. The features extracted contain important information about the bradykinesia characteristics of every movement task. In order to calculate the features, the data were first pre-processed.

4.2.1 Data pre-processing

The Polhemus device returns the recordings of movements for both sensors in a text file consisting of eight columns (Figure 4.5). The first column is sensor numbers: 01 refers to thumb sensor and 02 is finger sensor. The next three columns give readings for the Cartesian coordinates of X, Y and Z, in that order. Columns four to six are for orientations

in Euler angles: azimuth, elevation and roll respectively. The last column is the time stamp at which the specific measurement was made.

01	21.925	-7.943	0.269	111.180	-0.618	-40.453	1441936584125
02	21.678	-11.362	-3.672	-158.437	73.563	66.148	1441936584125
01	21.936	-7.921	0.283	111.119	-0.568	40.533	1441936584140
02	21.683	-11.321	-3.677	-158.977	73.481	65.596	1441936584140
01	21.937	-7.898	0.302	111.088	-0.454	-40.647	1441936584156
02	21.696	-11.271	-3.692	-159.895	73.431	64.576	1441936584156
01	21.932	-7.887	0.294	111.087	-0.406	40.649	1441936584172
02	21.697	-11.251	-3.699	-160.304	73.368	64.173	1441936584172

Figure 4.5 : Content of the text file returns by Polhemus device contains recordings of the movement in six degrees of freedom.

Digital filtering

In a study of movement where the signal is anatomical coordinate that changes with time, noise might be a problem. The presence of this higher-frequency noise is of considerable importance when considering the problem of trying to calculate velocities and accelerations (Winter, 2009). The kinematics data in this research were pre-processed to remove noise using the Low Pass 5Hz Butterworth filter. The Butterworth filter is the most common filter used in biomechanics data analysis due to its excellent passband response (Christodoulakis et al., 2010). Figure 4.6 shows an example of raw finger tapping data plot against time before and after passing through the filter.

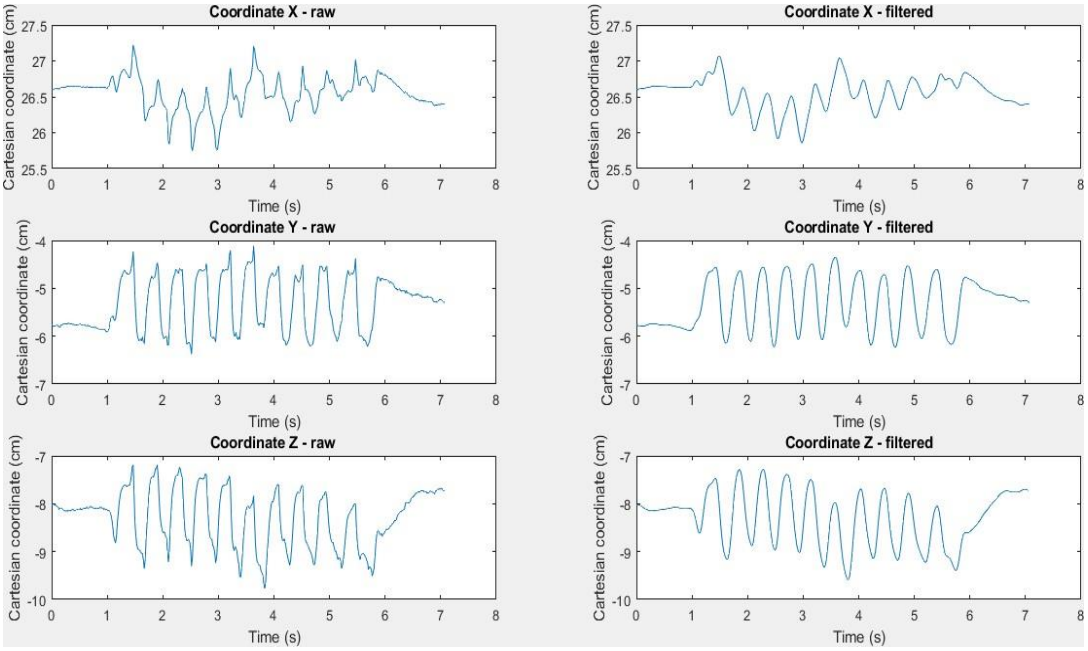

Figure 4.6: The raw data of Cartesian coordinates in the left column effectively smoothed using Butterworth filter as shown in the right column.

4.2.2 Finger Tapping computed features

As mentioned in Chapter 2, finger tapping task is a very popular task in studies that evaluate motor symptoms of PD. Previous studies using finger tapping to assess movement disorders were detailed in Section 2.4. In this study, nineteen kinematic features were selected to represents FT movement. The selection of the features was based on the literature and direct consultation from the clinicians involved in this study. Most of the features were adapted directly from Lacy (2013). Lacy used the same FT protocol and equipment with this study and successfully evolved very high-performance classifiers of PD and healthy controls. The rest of the features were based on findings from Yokoe et al. (2009), Stamatakis et al. (2013), Jobbágyly et al. (2005), Espay et al. (2009) and Teo et al. (2013) Table 4.1 mapped each feature to the research that it was based on.

Table 4.1: Movement features of finger tapping task.

Feature	Previous study
Mean amplitude Mean speed Max opening acceleration Max opening deceleration Max closing acceleration Max closing deceleration Periodicity COV amplitude (amplitude rhythm) COV speed (speed rhythm)	Lacy (2013)
Halts Hesitation Decrementing amplitude Decrementing speed Cycles frequency	Stamatakis et al. (2013)
Amp*freq	Jobbágyly et al. (2005)
Max overall amplitude Max overall speed	Espay et al., (2009)
Max opening speed Max closing speed	Yokoe et al. (2009)

The version of finger tapping in this study was administered as suggested in the motor section of MDS-UPDRS. Participants were guided to continually tap their index finger to the thumb as fast and as widely as possible for ten repetitions. There are two phases of finger tapping task, closing phase and opening phase (Figure 4.7). The sequence started with the closing phase. For the finger tapping task, only positional data of x, y and z are used for the calculation of the features. The separation distance between the finger and the thumb during the finger tapping action was computed by first calculating the difference between the x, y and z coordinate values for the respective sensors (X_d, Y_d and Z_d), and then, calculating the Euclidean distance, or overall positional separation, between index finger and thumb.

a) b)

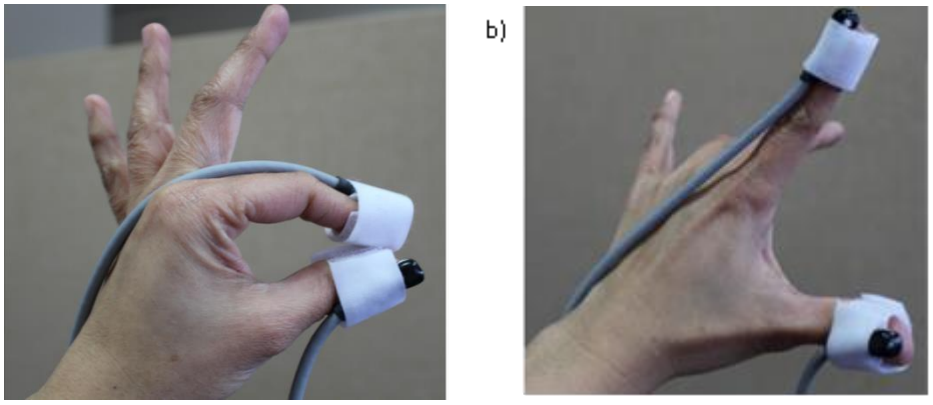

Figure 4.7: a) Finger tapping closed position. b) Finger tapping maximum opening.

The Euclidean distance D, was calculated as:

$$D = \sqrt{X_d^2 + Y_d^2 + Z_d^2}$$

However, different hand sizes of subjects can give the effect of bigger or smaller amplitude that may result in vary speeds and accelerations of tapping sequences dependent on hand size. Additionally, as the sensors were placed on the nails beds, there was a gap between two sensors that depended on the thickness of the subject's thumb and finger. The gap was considered as the minimum separation during measurement. To avoid this human factor affecting the calculation results, the separation data were normalised to scale of 0 to 1 to represent the relative distance between the thumb and finger. To do this, first, the minimum separation value for each subject was subtracted from all its other separation data points. The resulting values were then divided by the maximum separation of the tapping recording. The normalised separation time series data, Dn, then was differentiated to produce the speed time series data (dDn/t) and then differentiated again to give the acceleration time series data (dDn^2/t). Figure 4.8 shows examples of amplitude, speed and acceleration data of a healthy control.

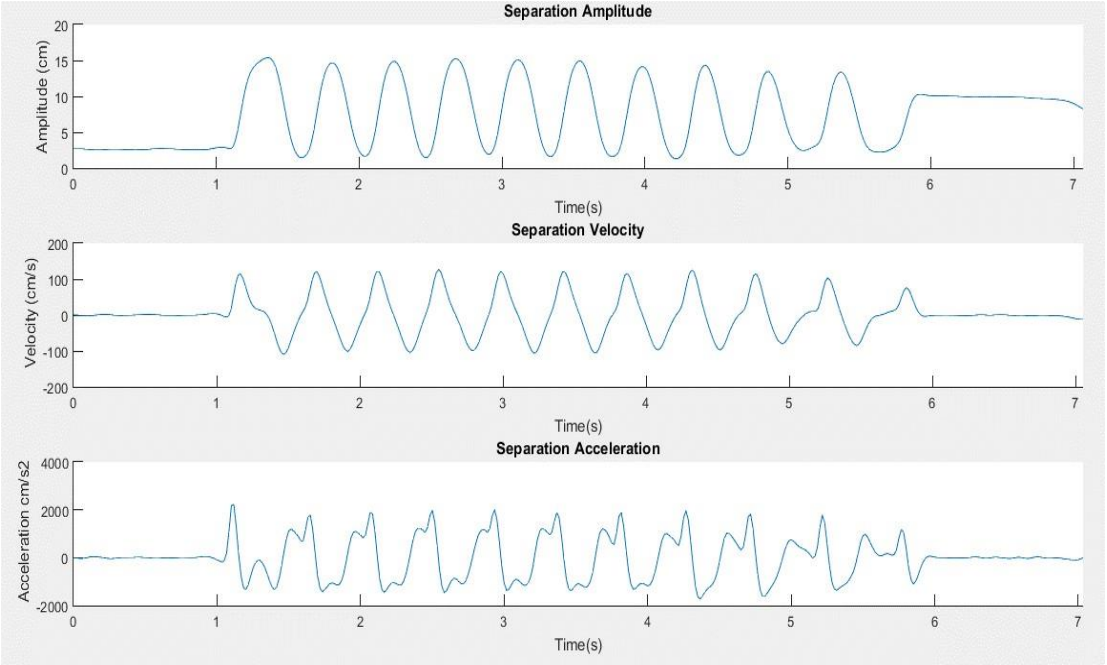

Figure 4.8 : Visuals of time series kinematic FT data. From top is amplitude in centimetres, velocity and acceleration for ten cycles of finger tapping.

Calculation of movement features

There are two approaches for calculating the features; most of the features are based on individual finger tapping cycles, but some features were calculated over the whole taping sequence.

A. Calculation within a single finger tap cycle

A full FT cycle comprises completed closing and opening phases. The closing phase begins once the sensors move towards one another after the point of maximal separation and finishes when the sensors have achieved a minimum separation; the opening phase begins once the fingers are separated, from an initially closed position equating to a minimal distance between the sensors to when they are maximally separated (Figure 4.9). Minimum, maximum and average of the normalised amplitude, speed and acceleration of both cycle phases were computed.

The features were calculated as follows:

Amplitude

The term amplitude here refers to the normalised separation values between sensors during the FT movement, thus, generating values between 0 and 1.

- **Maximum overall amplitude**: The maximum amplitude for each FT cycle was taken, the mean over the whole test then calculated to obtain a single value for each test.
- **Mean overall amplitude**: The mean amplitude of each FT cycle is calculated and then averaged.

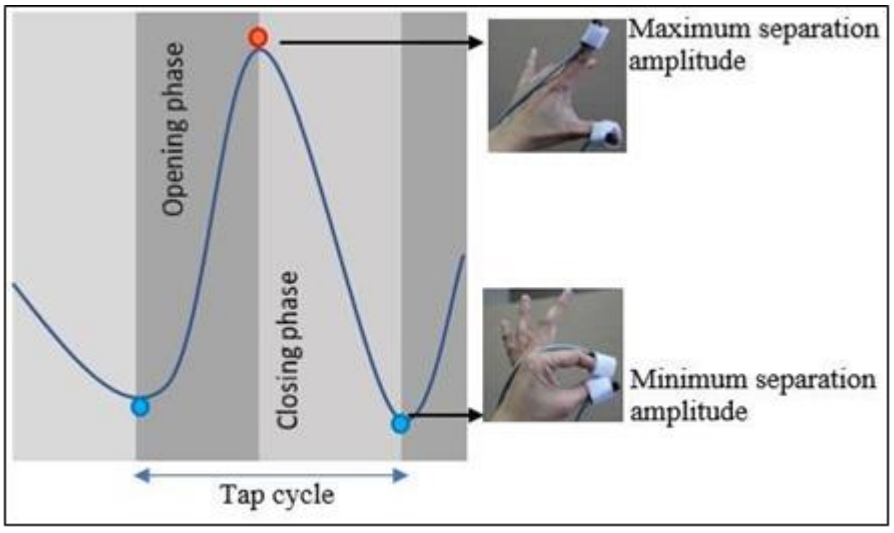

Figure 4.9: Separation data are showing opening and closing phases of a tapping cycle. Maximum amplitude is when the thumb and finger are farthest apart.

Speed

The speed instead of velocity is calculated because the goal is to observe the slowness regardless of the direction of movement. The following speed components were calculated:

- **Maximum overall speed**: The maximum speed reached in every FT cycle is averaged to obtain this value.
- **Maximum opening speed**: The average maximum speed of individual FT cycles during the opening phase
- **Maximum closing speed**: The average maximum speed of individual FT cycles during the closing phase

Acceleration and deceleration

Acceleration and deceleration in cm/s^2 were calculated separately for opening and closing phases of each FT tap. The following components were calculated:

- **Maximum opening acceleration**: The fastest acceleration during the opening phase of each FT tap was taken. The values were then averaged to get a single value for each test.
- **Maximum opening deceleration**: The most negative acceleration during the opening phase of each FT tap was taken. The values were then averaged to get a single value for each test.
- **Maximum closing acceleration**: The fastest acceleration during the closing phase of each FT tap was taken. The values were then averaged to get a single value for each test.
- **Maximum closing deceleration**: The most negative acceleration during the closing phase of each FT tap was taken. The values were then averaged to get a single value for each test.

Periodicity

Measures of amplitude and speed alone may not be sufficient to capture the full movement patterns of subjects. For example, a subject may be faster due to smaller amplitude and vice versa. To express the relationship between these components a variable called periodicity was calculated for each tap cycle as follows:

$$periodicity = maximum\ amplitude \times maximum\ speed$$

The mean was then taken to provide a single value for the whole test.

B. Calculation over tapping sequences

Cycles frequency

Patients often have difficulties in performing the exact number of the cycles as instructed. Therefore, cycles frequency is one of the features selected instead of time taken to finish the task.

$$Cycles\ freq = \frac{number\ of\ cycles\ completed}{time\ taken}$$

Decrementing amplitude and speed

To calculate the decrementing trend, maximum separation amplitude or speed for each tap cycle, each value was linearly regressed against the number of cycles. A negative slope indicates that the overall trend of a movement component measure is decrementing and a zero or positive slope indicates that there is no decrementing pattern**.** Figure 4.10 provides examples of linear regression plots of maximum amplitude to obtain the slope indicating a trend of separation amplitude.

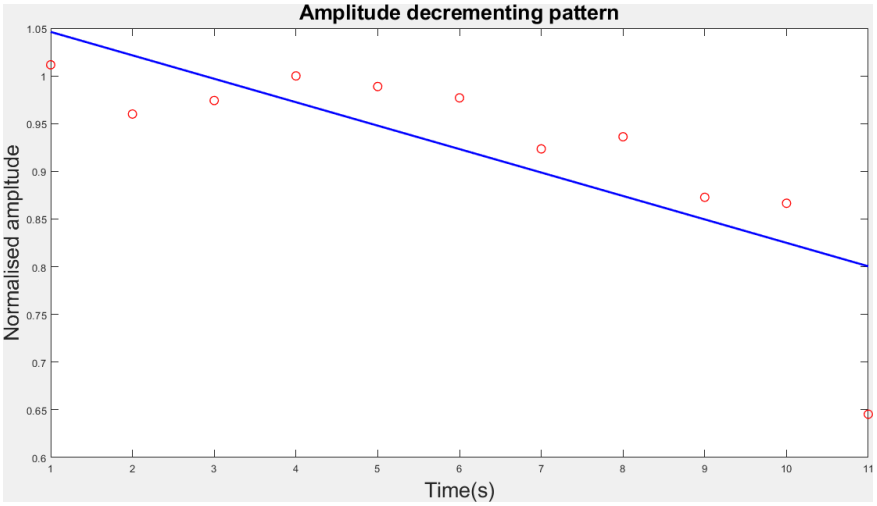

Figure 4.10: Amplitude decrement pattern

Halts

Halts were measured by calculating the percentage of the tap cycle duration spent at 'zero' speed. 'Zero' speed is assumed as when the speed is less than five percent (< 5%) of the maximum speed of the whole test.

$$Halts = \frac{time < 5\% \max speed}{time\ taken\ (test)} \times 100\%$$

Hesitation

When the movement showed smaller peaks between tapping cycle phases (Figure 4.11), it is treated as hesitation. All the hesitations detected in a test total up to a single value.

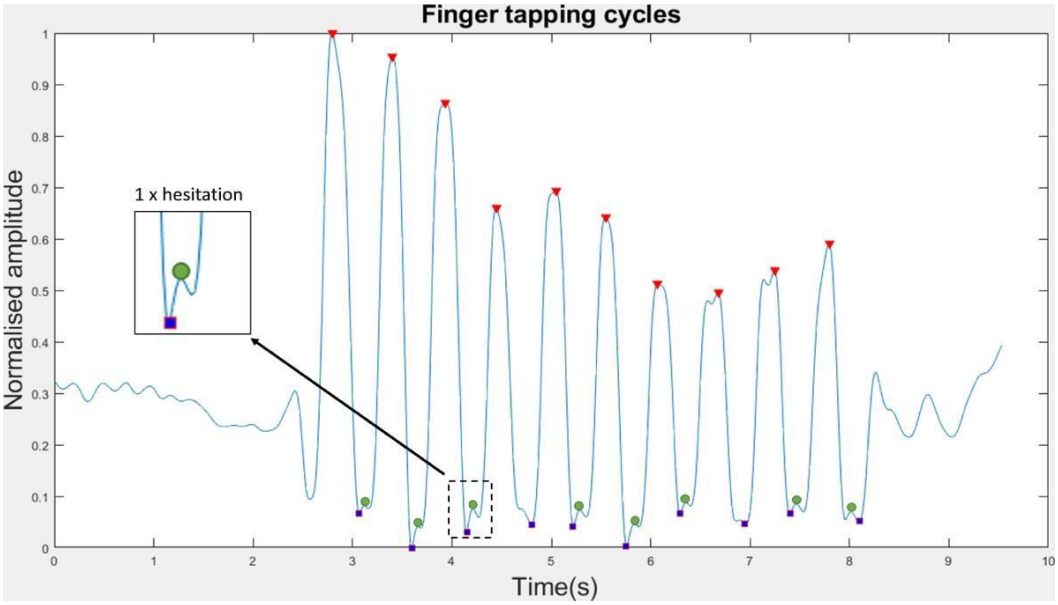

Figure 4.11: Hesitations in tapping cycle

Coefficient of Variation (COV)

To measure rhythm, the Coefficient of Variation (COV) was used. COV reflects how much a movement component measure varies over a defined period. It shows how rhythmic the repetitive movements are**.** High COV values imply less rhythmic movements than small COV values. COV of amplitude and speed were calculated over a period of tapping cycles as following:

$$COV\ amplitude\ = \frac{standard\ deviation\ of\ cycles\ maximum\ amplitude}{mean\ \ of\ cycles\ \ maximum\ amplitude}$$

$$COV\ speed\ = \frac{standard\ deviation\ of\ cycles\ maximum\ speed}{mean\ \ of\ cycles\ \ maximum\ speed}$$

Bigger amplitude with greater frequency during finger tapping means faster finger movement. This is considered to be a better performance. Alternatively, the movement can be executed faster with smaller amplitude. The amplitude × frequency of tapping is suggested in Jobbágy et al. (2005) to characterise the speed. This feature is determined for each tapping cycle and then averaged over the whole test.

All nineteen features extracted from the finger tapping (FT) task were given input numbers to be used as classifier inputs (Table 4.2).

Table 4.2: FT extracted features with classifiers input numbers

Input number	Feature
(0)	Cycles frequency
(1)	Maximum overall amplitude
(2)	Mean amplitude
(3)	Maximum overall speed
(4)	Mean speed
(5)	Maximum opening speed
(6)	Maximum closing speed
(7)	Maximum opening acceleration
(8)	Maximum opening deceleration
(9)	Maximum closing acceleration
(10)	Maximum closing deceleration
(11)	Periodicity
(12)	COV amplitude
(13)	COV speed
(14)	Decrementing amplitude
(15)	Decrementing speed
(16)	Halts
(17)	Hesitation
(18)	Amp*freq

4.2.3 Hand pronation-supination computed features

For the hand pronation-supination task (PS), the MDS-UPDRS requires the participant to extend the arm out in front of their body with the palms face down and then turn the palm up and down alternately 10 times as fast and fully as possible. (Figure 4.12)

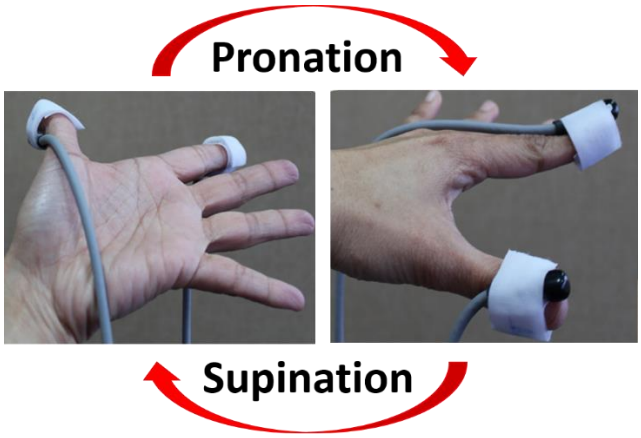

Figure 4.12: Hand pronation-supination task: the phase when the palm is facing down is called pronation. Supination is the phase when subject twists the palm up from pronation phase.

After some experimentation, it was concluded that the most useful data in our pronation-supination recordings came from the movement of the thumb. The PS cycle starts with the pronation phase followed by supination (Figure 4.13). Referring to the thumb sensor recorded movement, when the position of thumb sensor (position A) begin to change at the beginning of the task, it is considered as the start of the pronation phase. When the sensor reaches point B and starts moving towards C, the supination phase starts.

For analysis of movement using only one sensor, the amplitude is defined as the Euclidean distance between thumb sensor and Patriot transmitter.

$$amp\ (t) = \sqrt{x(t)^2 + y(t)^2 + z(t)^2}$$

Speed was calculated by differentiation of each Cartesian coordinate component (x, y, z) over the sampling time to compute the respective velocity components (v_x, v_y, v_z). The total velocity was computed from the sum of its components and its magnitude, the speed:

$$speed(t) = \sqrt{v_x(t)^2 + v_y(t)^2 + v_z(t)^2}$$

Table 4.3: Extracted features of PS.

Input number	Computed features
(0)	Cycles frequency
(1)	Maximum overall amplitude
(2)	Mean amplitude
(3)	Maximum overall speed
(4)	Mean speed
(5)	Maximum pronation speed
(6)	Maximum supination speed
(7)	Maximum pronation acceleration
(8)	Maximum supination deceleration
(9)	Maximum pronation acceleration
(10)	Maximum supination deceleration
(11)	Periodicity
(12)	COV amplitude
(13)	COV speed
(14)	Decrementing amplitude
(15)	Decrementing speed
(16)	Halts
(17)	Hesitation
(18)	Amp*freq

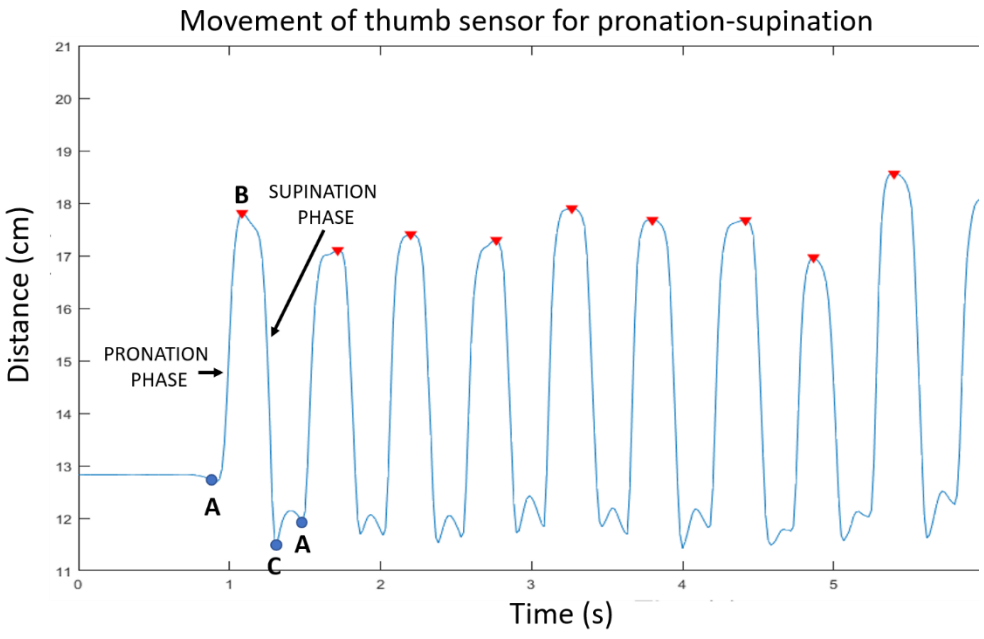

Figure 4.13: Phases of pronation-supination.

Acceleration is obtained by differentiating the speed, using the same sampling time. The same method of calculation as in Section 4.2.2 was used to compute items (0) to (18) in Table 4.3.

4.2.4 Hand opening-closing

Item 3.5 in MDS-UPDRS is hand movements, which is known as hand opening-closing in this study. In this task, participants were first instructed to make a tight fist with the arm bent at the elbow so that the palm faces the examiner (Figure 4.14). Subjects were then asked to open the hand as fully and as quickly as possible. The participant repeated the opening and closing movements ten times.

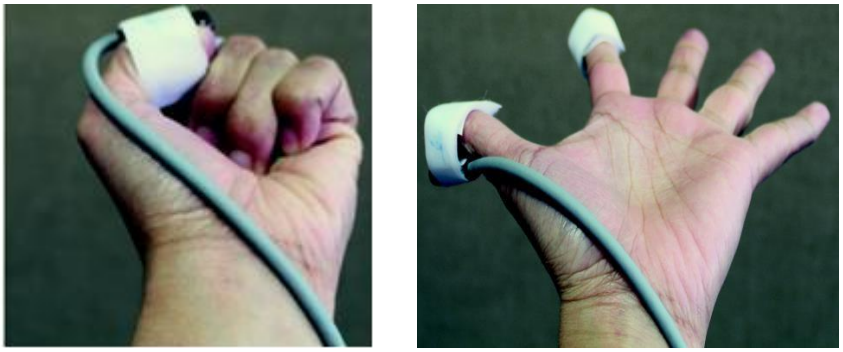

Figure 4.14: The phases of hand opening-closing task. Subject starts with the hand closed by making a fist and proceeds to the opening phase by fully opens the hand..

In the hand opening-closing task, sensors were placed at the same positions as in the finger tapping task. However, unlike finger tapping, which is a simultaneous movement of thumb and fingers, the hand-opening task involves two steps movement. Therefore, the selection of extracted features was considering the measurements of both sensors separately. The cycle of the HO task started from the closing position to maximum amplitude opening and back to closing position. Figure 4.15 shows the task cycle represented by the movement of the finger sensor (FS) amplitude.

The calculation of amplitude for each sensor used the Euclidean distance between the Patriot transmitter and the sensor itself (as in section 4.2.3). The speed and acceleration is then calculated from the amplitude values. . The maximum values for each sensor are the highest from both phases, opening or closing. For example, to get **FS maximum speed**, the maximum speeds of the finger sensor during the opening and closing phases were compared, whichever is higher is taken as the feature.

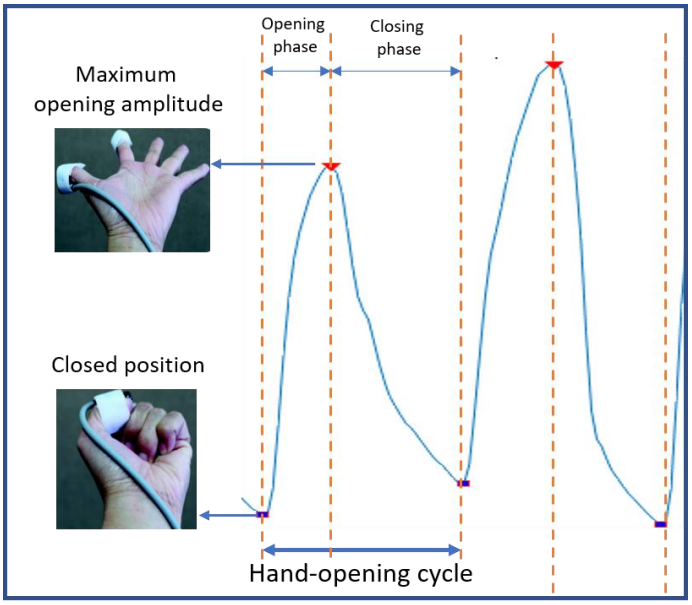

Figure 4.15: The opening and closing phases of the HO task.

The thumb and finger movement data executing the HO task were used to compute the total of seventeen features (Table 4.4).

Table 4.4: Hand-opening extracted features.

Input number	Computed features
(0)	Maximum opening
(1)	COV opening
(2)	HO frequency
(3)	TS maximum speed
(4)	TS average speed
(5)	TS maximum acceleration
(6)	TS maximum deceleration
(7)	TS COV speed
(8)	TS Halts
(9)	FS maximum speed
(10)	FS average speed
(11)	FS maximum acceleration
(12)	FS maximum deceleration
(13)	FS COV speed
(14)	FS Halts

4.3 CGP Classification

The classification used a typical CGP evolutionary strategy which selects one parent from each generation and uses mutation to produce four children. The next generation then comprises the parent and the four children, giving a population of size five - four children plus one parent: (1+4) - ES. The fitness assigned to each classifier is simply the proportion of samples correctly classified. Previous CGP classifiers of FT (refer section 2.5) used the area under a ROC Curve (Fawcett 2006) as fitness function, but in this study, classification accuracy is used for simplicity and directcomparison.

The CGP platform used to evolve the classifiers was developed by Turner & Miller, (2014) called the CGP-Library. The CGP-Library provides a number of structures including which describe the CGP parameters and CGP chromosomes.These structures

are initialised and freed up using provided functions. The following three CGP-Library structures are used to train and test classifiers.

- The *parameters structure* is used to store general CGP parameters including the size of the chromosomes, the evolutionary strategy to use and the mutation rate.
- The *dataSet* structure is used to store input output pairs of data which may be used by the fitness function.
- The *chromosome* structure is used to store the fittest chromosome found after CGP has been applied towards the classification task.

Through experimentation, the CGP parameters values as in Table 4.5 were adopted. The function set comprised ({+, -,×,÷,mean, min, max, mode }).

Table 4.5: CGP parameters adopted to evolve classifiers.

Evolutionary strategies	(1+4) - ES.
Nodes available	15,50
Node arity	2,5
Mutation rate	0.05
Function set	$(\{+, -, \times, \div, mean, min, \max, mode\})$
Generations	10000

The features extracted from the movement tasks were used as the inputs to evolve classifiers. For each movement recording, a set of input values were computed. The input values are floating numbers from calculation explained in Section 4.2. Each set of input was mapped to one output according to the group of the input data. For example, an input set belong to the Control group was given output '0', and '1' for PD group. Figure 4.16 shown an example of CGP network with FT features as inputs. The number in bracket represents a node number; where a node can be an input node or a function node. Each function node is labelled with a textural description of the operation it undertakes such as

'add'. The nodes in the network are numbered sequentially starting at zero with the first of the input nodes. Therefore, the input numbers in the CGP network are tally with the input numbers in Table 4.2, 4.3 and 4.4. The movement features data was fed into the algorithms in the form of dataset files. There were two text files in each

Figure 4.16: Example of a CGP netwrork using FT movement features as input.

training-test cycle; the training file contained data to train a classifier and the test data that is later used to test the predictive accuracy of the evolved classifier. The CGP platform requires a specific format of dataset file. The header of the text file contains information on the number of inputs, outputs and lines of data samples (Figure 4.17). The following lines comprise the inputs followed by the outputs for each sample with all the values comma separated. In every sample line, the input values were arranged according to the sequence of the input numbers of the features extracted.

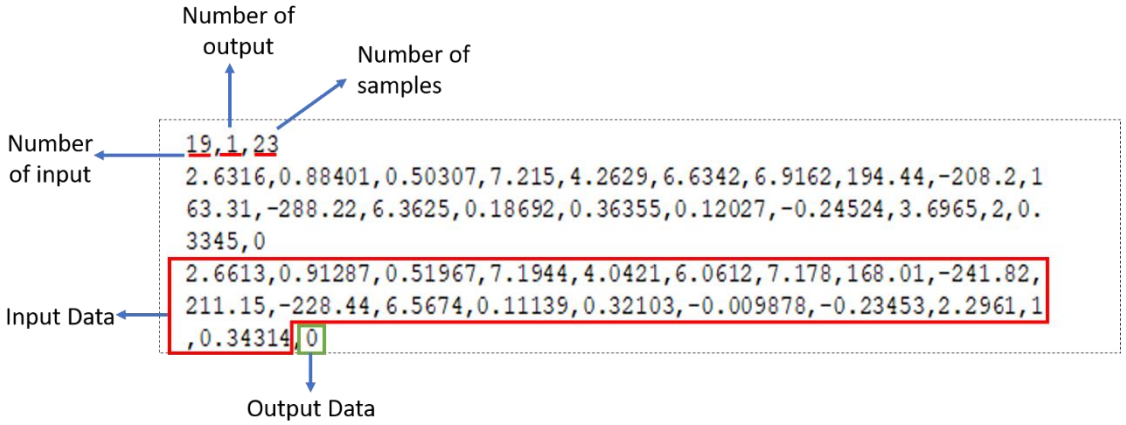

Figure 4.17: The computed features arranged in a text file to be used in the classification task.

4.3.1 Repeated stratified K-fold cross-validation

To compensate for any effect on results caused by small amounts of training and test data, the repeated k-fold cross-validation was used along with the stratification of data.

In the k-fold cross-validation method, the overall data set is divided into separate k groups of the same, or approximately the same size. In each iteration, only a fold is used for evaluation (test set), and the rest of the folds are used for model training (training set). In machine learning, k= 10 is a very popular choice for the number of folds. However, in this study, the number of participants is not big enough to be divided into ten groups.

Thus, k =5 is used, and in some cases, k = 3. Figure 4.18 illustrates the 5-fold cross validation with A as the resulting accuracy which is the average of all iterations.

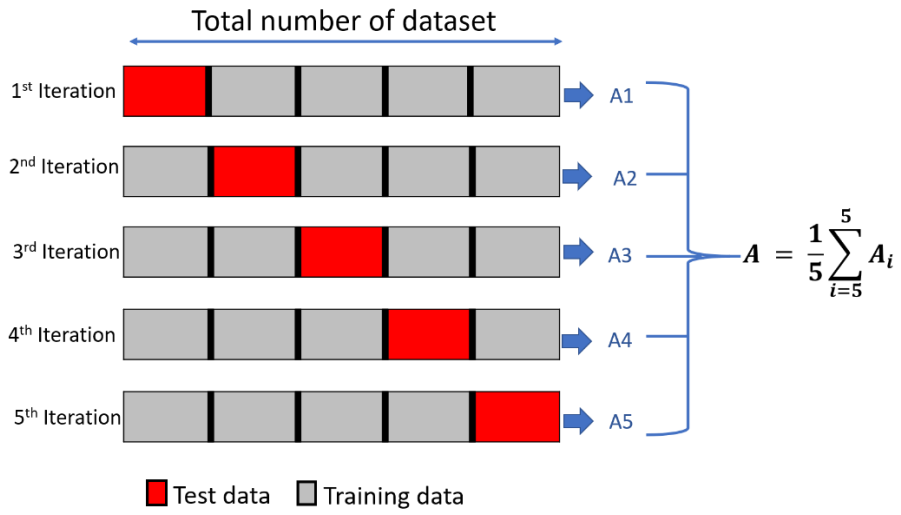

Figure 4.18:5-fold cross-validation technique.

In small sample size classification such as in this study, a class imbalance can easily occur. This can happen in the partition itself, although the overall classes are balanced. For example, consider the healthy controls and PD patients' dataset. There are kinematics data for 22 controls and 20 PD patients which is almost a balanced class distribution. However, when the data set is randomly partitioned, it is possible that the class balancewill be destroyed. The training set may contain more healthy controls than PD patients and the opposite for the test set. To avoid this, the data were stratified. It means that when the data were divided into groups (folds), the program was written to randomly pick a balanced number of subjects from both groups for each fold. The k-fold cross-validation was repeated ten times for improving statistical significance and to increase the number of predictions. The data is reshuffled and re-stratified before each run.

4.3.2 Record-wise and Subject-wise cross-validation.

Participants repeat the movement task twice for each hand. Each participant will have at least four different kinematic records of the same task. There are two ways the data can be partitioned for cross-validation: record-wise or subject-wise. In record-wise validation, all the records were used as different samples and randomly divided into training and test sets. Whilst in a subject-wise validation, the records for the same subject will not be in the training and test set at the same time. Figure 4.19 visualises subject-wise and record-wise cross-validation for clinical predictions.

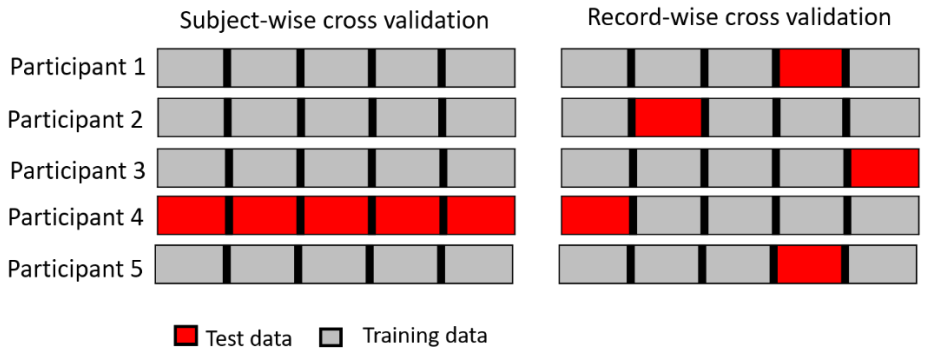

Figure 4.19: Visualisation of subject-wise and record-wise cross-validation used at the stage of selecting kinematic features into training and test set.

Saeb et al. (2017) in their critical article proposed that in the clinical diagnostic application of machine learning, subject-wise CV should be used instead of records-wise CV. Using empirical evidence and simulation of classification on a few public database, they claimed that record-wise CV greatly overvalues the prediction accuracy of the algorithms used. They believed this is because records-wise CV creates a dependence between training and test sets due to shared subjects across train/test sets, so it will produce biased estimates. However, reviewers believed that it is not fair to oversimplify the application of the cross-validation technique and make assumptions based on the results of a few simulations (Little et al., 2017). In this study, both types of CV were used, and comparisons were made between the accuracies of both methods.

4.3.3 Comparisons with other methods of machine learning

In medical classification, there are other methods that are popular and useful for the purpose of classifying classes (refer to Chapter 2). In this study, two other machine algorithms were used as a comparison to evaluate the performance of CGP in the classification task. The algorithms are Support Vector Machine (SVM) and Artificial Neural Network (ANN).

For SVM, the WEKA default algorithm that employed sequential minimal optimization algorithm (SMO) for training a support vector classifier was used. The default settings used polinomial kernel function. Different experiments were done using other kernel functions (Gaussian, linear) without any marked improvement in the accuracies, so the WEKA SMO default settings were used.

The Neural network used was a feed-forward back propagation network with 20 hidden layers and a tan-Sigmoid transfer function. The number of hidden layers was chosen using an iterative method; the number of hidden layers were increased gradually registering the maximum improvement in the results when 20 is considered. A greater number of hidden layers did not improve the results so the hidden layer number was chosen equal to 20. For both algorithms, 10-fold cross-validation was used , and all records treated as different instances (record-wise cross- validation)

4.4 Summary

There are many ways movement data collection can be done. We opted for the low-cost, easy to used equipment for this study. The motion tracker, Polhemus only used two electromagnetic sensors but able to records movement in six degrees of freedom. Good quality of data is very important in machine learning classification task. The data collected were pre-processed and explored to get to know the data. Based on previous

quantitative studies of PD motor tasks, many movement features were extracted. Statistical tests conducted for selection of useful features. The extracted feature were used to evolve classifiers that able to differentiate between groups of movement disorders and healthy controls.

CHAPTER 5

Evaluation of Bradykinesia in Parkinson's Disease and other movement disorders

As discussed in Chapter 2, of the four main signs of Parkinson's disease, bradykinesia is the only mandatory motor sign for a clinical diagnosis of PD (Heldman et al., 2011). The complex nature of bradykinesia makes it difficult for clinicians to be certain of its existence in the early stages of Parkinson's disease. Bradykinesia also exists in Huntington's disease (HD), a genetically inherited neurodegenerative condition that causes movement disorders such as chorea (Garcia Ruiz et al., 2000; Martínez Pueyo et al., 2016; Thompson et al., 1988). On the contrary, there is research suggesting that bradykinesia is not a core feature of movement disturbance in HD (Duval et al., n.d.; Fenney et al., 2008). The same pattern occurs in the case of Essential Tremor. A few papers suggest Bradykinesia exists in ET (Duval et al., 2006; Goubault et al., 2017), but others oppose the idea (Özekmekçi, 2005). The aim of this work is to provide a better understanding of the characteristics of bradykinesia, how they differ between these conditions and may be used to confirm an early diagnosis and inform clinical assessment. In light of the arguments presented in Chapter 3, a type of Evolutionary Algorithm (EA), Cartesian Genetic Programming (CGP), was applied to the movement data of motor tasks

obtained from conventional clinical assessment using the Movement Disorders Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) (Goetz et al., 2008), as described in Chapter 4. The aim being to evolve diagnostic classifiers that can differentiate between the specified movement disorders. Movement data were provided by a team of clinicians and neurologists through a joint research project funded by Centre for Chronic Diseases and Disorders, University of York. Data collection was undertaken in two research centres; Leeds Teaching Hospitals NHS Trust, UK and the Monash Medical Centre, Melbourne, Australia. Participants in the United Kingdom were recruited primarily from Dr Jane Alty's and Dr Stuart Jamieson's consultant caseload. Both Dr Alty and Dr Jamieson are consultant Neurologists at Leeds Teaching Hospital NHS Trust. Patients recruited had an established clinical diagnosis of PD, HD and ET according to diagnostic criteria performed by consultants working at these centres. Healthy controls in this study are healthy adults with no history of a medical disorder that could affect movement i.e. exclusion criteria for controls include tremor, clinically diagnosed parkinsonism, stroke, dopamine receptor antagonists' drugs, significant arthritis of upper limbs, dementia, inability to provide informed consent. In total, 27 healthy controls were recruited. Each participant wore electromagnetic sensors on their thumb and index finger and then performed three MDS-UPDRS motor tasks: finger tapping, hand pronation-supination and hand opening-closing, as described in Chapter 4. The recorded movement data was then used as inputs to train the CGP networks. Full details of the methodology used is described in Chapter 4. This study was designed to answer following questions:

Primary question:

What features of bradykinesia differentiate subjects with PD from subjects with other movement disorders and from healthy age-matched controls?

Secondary questions:

i) What are the essential differences between age-related loss of movement efficiency and bradykinesia in idiopathic PD?
ii) How is the bradykinesia of PD different from other non-PD parkinsonian conditions such as Huntington's disease?
iii) Do subjects with non-parkinsonian tremulous conditions such as Essential tremor also exhibit bradykinesia?

5.1 Bradykinesia in Parkinson's Disease and healthy controls

Using finger tapping (FT) data, Evolutionary Algorithms (EAs) have previously been used to evolve high accuracy classifiers that differentiate Parkinson's disease patients from healthy controls (Lones et al., 2014; Smith & Timmis, 2008). Further investigation into the classifiers evolved was able to characterise movement disorder in PD (Lacy et al., 2013) and inform clinical assessment (Lones et al., 2013). This study is extending these successful previous studies by using other motor tasks to evolve classifiers. Two additional motor tasks are hand pronation-supination (PS) and hand opening-closing (HO). FT data were used as a comparison and a baseline for the study. In total 26 patients with idiopathic Parkinson's disease were recruited along with 27 healthy controls.

5.1.1 CGP classification results (controls vs PD)

As explained in Chapter 4, we used accuracy to evaluate the performance of the CGP classifiers. Accuracies are the prediction of the classifiers trained on an unseen data called the "test set". Final accuracies in this study are the average of 10 runs where each run included 5 or 3 iterations of cross-validation (5-fold cross-validation or 3-fold depending on the number of participants available). The averaged acuracies results are reported in

the form of mean ± standard deviation. Both results for train and test set are present for each task.

For the FT task, the final accuracy of the test set across 10 runs is 82.37%, for the pronation-supination (PS) task, 86.71%, and for the hand opening (HO) task, 71.19%. Classifiers using data from PS task achieved the highest accuracies with the best classifier providing an average of 96.88 %. The best classifier accuracy for the finger tapping task (FT) is slightly lower at 95.12%, and the best classifier accuracy for the hand opening-closing (HO) task, is the lowest with an accuracy below 80%. Overall performance of HO classifiers were poor compared to the other two tasks. Best classifier and average across ten runs accuracies in the case of Controls vs PD of all tasks are summarised in Table 5.1. In Table 5.1, the result reported is in the form: mean ± standard deviation, where the mean and the standard deviation of the accuracies for both values of train and test set are stated. Averaged accuracies are visualised in Figure 5.1 to show the minimum and maximum performance for each task.

Table 5.1: Average and best accuracies of classifiers evolved for all motor tasks (controls vs PD).

Controls vs PD

Task	Accuracy (%)			
	Averaged ten runs		Best classifier	
	Train	Test	Train	Test
Finger Tapping	91.64 ± 1.03	82.37 ± 2.49	93.75	95.12
Hand pronation-supination	92.50 ± 0.65	86.71 ± 3.71	92.65	96.88
Hand opening-closing	82.18± 2.56	71.19± 2.84	84.19	75.87

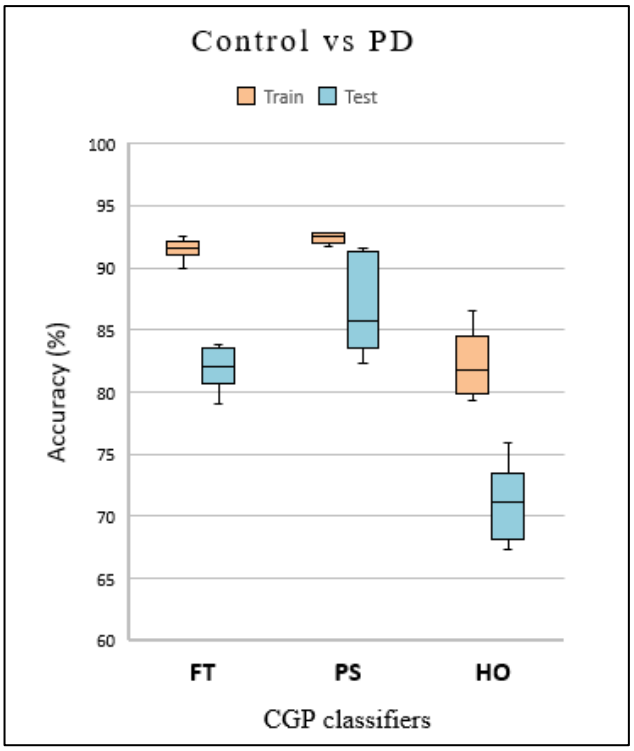

Figure 5.1: Averaged accuracies of CGP classifiers according to type of motor task data.

5.1.2 Discriminating features

The best classifier of each task was selected based on its ability to predict the class of unseen data in the test set. All iterations in the ten runs were considered to find which iteration yielded the best classifier. The best-run accuracies shown in Table 5.1 are an average of 5-fold cross-validation iterations. These high-performance classifiers were considered to identify most discriminating features of the motor task. In order to easily identify discriminative movement features of the best classifiers, the CGP chromosomes of the classifier were visualised using the open source cross-platform Graphviz utility (Ellson et al., 2000). The chromosomes are displayed with the inputs on the left, outputs on the right and the position of the internal nodes optimised by Graphviz. The function nodes are labelled with their functionality and given in bold if active. By looking into the visualised chromosome, the most discriminating features were identified. Only classifiers

with averaged test set accuracy higher than 80% were considered to determine the discriminating features. Therefore, all classifiers using HO data were not considered.

Finger tapping discriminating features

In order to verify which FT movement features make the biggest contribution in classifying the controls and PD groups, five FT classifiers with highest accuracies ranging from 88.24% to 95.12%, labelled FTPD_C1 to C5 were investigated together (Table 5.2).

Table 5.2: Five best FT classifiers differentiating PD patients from healthy controls.

Classifier	Accuracies (%)	
	Train	Test
FTPD_C1	93.75	95.12
FTPD _C2	91.03	93.02
FTPD _C3	87.74	91.95
FTPD _C4	90.68	90.12
FTPD _C5	92.99	88.24

In each classifier, features that were used at least once were ranked based on the number of times they were used. Feature with the highest number of usages were given rank 1, and so on (Table 5.3). If a feature was not used at all, no rank given and indicated by 'X'. For each feature, the rank they get from all five classifiers were summed. By using the combination of the rank, the most discriminating features were determined.

Table 5.3: Ranking of feature usage in five best FT classifiers for class controls vs PD.

Feature	Classifiers FTPD (rank)					Total
	C1	**C2**	**C3**	**C4**	**C5**	
(0) Cycles frequency	3	6	6	4	X	X
(1) Maximum overall amplitude	5	6	6	4	X	X
(2) Mean amplitude	4	5	X	4	1	X
(3) Maximum overall speed	1	X	X	3	4	X
(4) Mean speed	4	6	X	X	4	X
(5) Maximum opening speed	X	2	X	X	4	X
(6) Maximum closing speed	2	6	3	4	X	X
(7) Maximum opening acceleration	1	6	2	4	2	15
(8) Maximum opening deceleration	3	6	2	4	X	X
(9) Maximum closing acceleration	3	1	1	3	X	X
(10) Maximum closing deceleration	1	6	6	1	X	X
(11) Periodicity	4	X	3	3	2	X
(12) COV amplitude	2	6	X	4	4	X
(13) COV speed	2	X	X	X	X	X
(14) Decrementing amplitude	1	4	4	1	3	13
(15) Decrementing speed	3	X	5	3	3	X
(16) Halts	4	3	4	3	3	17
(17) Hesitation	1	X	1	3	X	X
(18) Amp*freq	3	X	X	X	4	X

There are only three features that were used at least once in all the five classifiers (highlighted in Table 5.3). The smaller total of rank values from all five classifiers, the more important the feature is. Combined rank values revealed the most important FT features in classifying healthy controls and PD patient (in order of importance) are as following:

i. (14) Decrementing amplitude.

ii. (7) Maximum opening acceleration.

iii. (16) Halts.

The feature that was least used in all the five classifiers is coefficient of variance (COV) of the speed of movement (13). Feature numbers are from Table 4.2 in Chapter 4.

The CGP network of chromosome of the best FT classifier (FTPD_C1) is shown in Figure 5.2. This classifier achieved 95% accuracy. The most used input by this classifier is input 7, which is averaged maximum acceleration during opening phase of FT (refer to Section 4.1 in Chapter 4). It is used 5 times in finding the solution.

Pronation-supination discriminating features

The top five classifiers using PS data achieved accuracies ranging from 93.75% to 96.88% (Table 5.4). Out of nineteen PS features, only six were used at least once by all the five classifiers.

Table 5.4: The best PS classifiers for class Controls vs PD.

Classifier	Accuracies (%)	
	Train	Test
PSPD_C1	92.65	96.88
PSPD_C2	91.67	94.44
PSPD_C3	91.67	94.44
PSPD_C4	90.91	94.44
PSPD_C5	90.44	93.75

Using the same method of giving ranks based on number of times a feature was used, the PS features that are most useful in classifying healthy controls and PD patients are as follows (ranked most important to least important from the top):

i. (12) COV amplitude
ii. (11) Periodicity
iii. (15) Decrementing speed
iv. (1) Mean amplitude
v. (2) Maximum amplitude
vi. (8) Maximum pronation deceleration

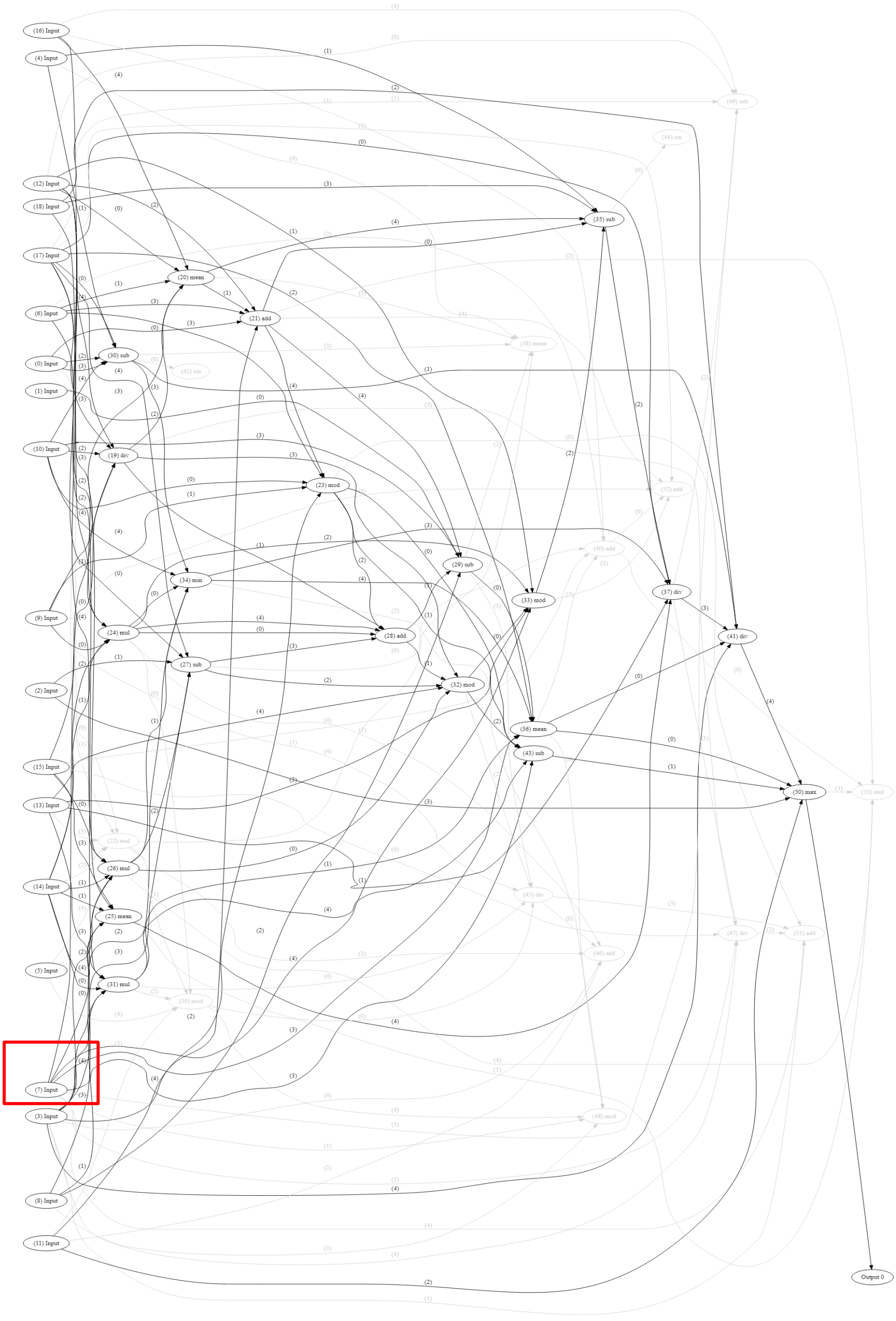

Figure 5.2: Visualised chromosome of the FT best classifier for the case of controls vs PD.

The full report on rank given to each movement feature for all five best PS classifiers is shown in Table 5.5. Feature numbers are from Table 4.3 in Chapter 4.

The best PS classifier for this case of controls vs PD (labelled PSPD C1 in Table 5.4) has a training accuracy of 92.65 % and predicted instances in the test set correctly at 96.88%. The most used features for this classifier are (12) COV amplitude and (2) mean amplitude.

Table 5.5: Rank given to the movement features of top five best PS classifiers based on the number of times the feature is used in evolving the classifiers.

Feature	Classifier PSPD (rank)					Total
	C1	C2	C3	C4	C5	
(0) Cycles frequency	6	X	6	3	4	X
(1) Maximum overall amplitude	1	10	4	3	4	22
(2) Mean amplitude	4	8	4	2	2	20
(3) Maximum overall speed	5	7	8	3	X	X
(4) Mean speed	3	8	2	X	X	X
(5) Maximum pronation speed	3	9	7	2	X	X
(6) Maximum supination speed	4	10	6	3	X	X
(7) Maximum pronation acceleration	7	2	6	X	X	X
(8) Maximum pronation deceleration	7	6	3	3	2	21
(9) Maximum supination acceleration	8	5	8	X	X	X
(10) Maximum supination deceleration	X	10	7	1	2	X
(11) Periodicity	6	3	4	2	1	16
(12) COV amplitude	2	1	3	3	5	14
(13) COV speed	7	6	6	X	X	X
(14) Decrementing amplitude	4	6	1	X	3	X
(15) Decrementing speed	6	5	3	3	1	18
(16) Halts	7	6	6	X	X	X
(17) Hesitation	8	4	7	3	X	X
(18) Amp*freq	8	8	5	X	4	X

5.1.3 Record-wise cross-validation

As explained in Chapter 4, many research studies have employed record-wise cross-validation (cv) in order to provide more data for learning process of machine learning

algorithms (Saeb et al., 2017). This is especially true in clinical studies where real data are difficult to obtain (Forman & Cohen, 2004). We run record-wise cv in classification of controls and PD patients for comparison purposes and to investigate the claim that record-wise cv is not suitable to use in the case of medical data. We retained the same method of stratified k-fold cross-validation (k = 10). However, instead of making sure no files from the same subject appear in both training and test sets of the same fold, files were picked randomly regardless of subjects. Table 5.6 shows accuracies achieved by CGP classifiers from all three tasks using record-wise cross-validation. Using record-wise cv, the best classifiers for FT and PS tasks are able to classify all the instances correctly. HO classifiers at best can classify correctly 88% of the instances.

Table 5.6: CGP accuracies using record-wise stratified 10-fold cross validation.

Controls vs PD

Task	CGP accuracy (record-wise cv)			
	10-fold average		Best classifier	
	train	test	train	test
FT	95.834	93.517	95.96	100
PS	89.679	91.03	90.07	100
HO	81.204	74.444	81.29	88.24

Figure 5.3 showing comparison of accuracies between subject-wise cross-validation and record-wise cross-validation. In two out of three cases, classifiers using record-wise cross-validation have higher accuracies compared to subject-wise classifiers. This might be an indication that record-wise CV is overestimating the accuracy due to dependency of the training and test instances. However, the difference between the accuracies are not so obvious to assume that the learning process was totally bias. In fact, in case of HO classifiers, using subject-wise cross validation yielded better accuracy than using record-wise cross-validation.

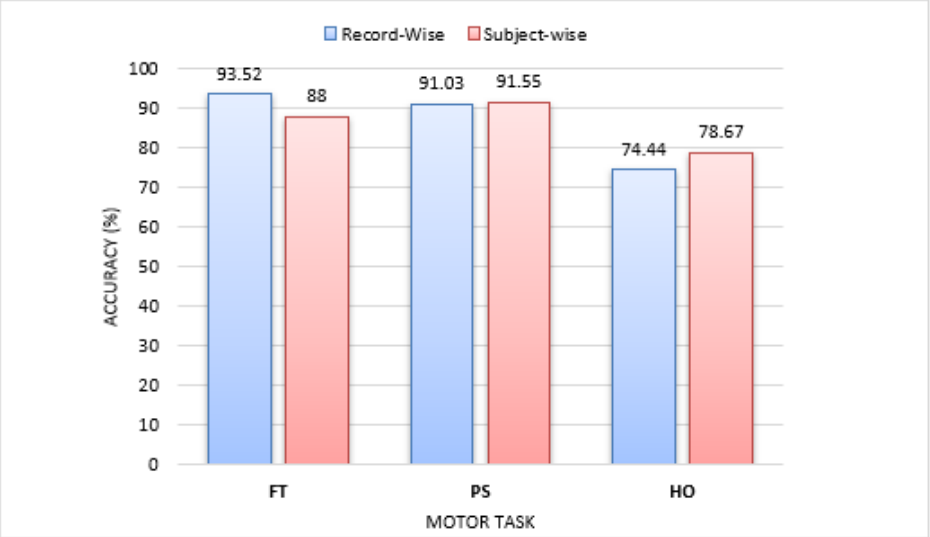

Figure 5.3 : Comparison of CGP classifiers accuracies using subject-wise and record-wise cross-validation

5.1.4 SVM and ANN classifiers

Support vector machines (SVMs) and Artificial Neural Networks (ANNs) are known for their ability to generate good classifiers (refer Section 3.1 in Chapter 3). These two machine learning methods were used to compare to the CGP classifiers.WEKA (Bouckaert et al., 2010), an open source machine learning workbench, was used to run SVM and ANN classification tasks.

For SVM, the WEKA default algorithm that employed sequential minimal optimization algorithm (SMO) for training a support vector classifier was used. 10-fold cross-validation, and all records treated as different instances (record-wise cross-validation) was used. For ANN classifiers, popular feed forward backpropagations were used to learn a multi-layer perceptron to classify instances. Table 5.7 summarise the accuracies of using three types of algorithms to classify controls and PD patients. For fair comparison, CGP classifiers in Table 5.7 were also validated using record-wise cv. For FT classifiers, CGP gave the best performance at 93.51% followed by ANN classifiers with 87.90%, and SVM classifiers have slightly lower accuracies, 84.68%. All classifiers

using PS data showed good accuracies with around 6% difference between three types of algorithms (85% - 91%). HO classifiers showed poor performance, regardless of the type of algorithms used. CGP HO classifiers perform slightly better than both SVM and ANN classifiers. None of the HO classifiers reached 80%. However, the averaged accuracies for HO classifiers are no lower than 70%.

Table 5.7: Performances of SVM and ANN classifiers for controls vs PD compared to CGP classifiers.

Controls vs PD

Task	Accuracy (%)					
	CGP (record-wise cv)		SVM		ANN	
	train	test	train	test	train	test
Finger Tapping	95.83	93.51	87.50	84.68	98.39	87.90
Hand pronation-supination	89.67	91.03	88.10	86.90	98.80	85.1
Hand opening-closing	81.20	74.44	75.58	69.77	76.16	73.26

The comparison between test set accuracies of each algorithm is visualised in following Figure 5.4.

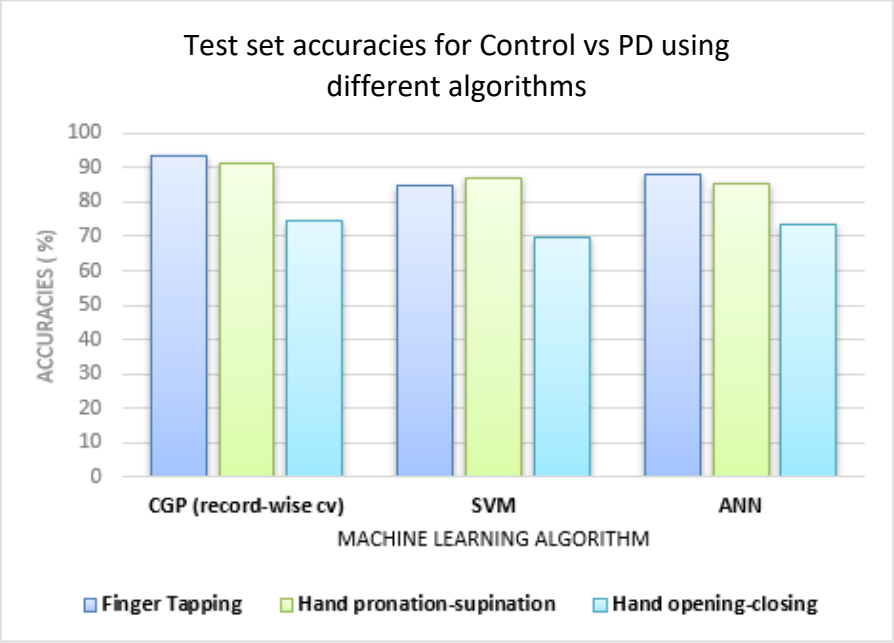

Figure 5.4 : Comparison between performance of different algorithms for Controls vs PD.

5.1.5 Discussion

FT kinematic data have been used many times to distinguish healthy controls and PD patients. Indeed, some studies also used CGP in differentiating these two groups (Lacy et al., 2013; Lones et al., 2013; Lones et al., 2014). However, FT data were investigated again in this study, but for the purpose of validating the method used, as we extend the study to other motor tasks and later other groups of movement disorders. At 80% and above, the results of this study showed accuracies achieved by the FT classifiers are comparable to results from the previous FT (using CGP, SVM and ANN) studies (Lacy et al., 2013; Lones et al., 2013; Lones et al., 2014; Martinez Manzanera et al., 2015; Picardi et al., 2010). The important finding in the FT results is that the most discriminating FT features did demonstrate the method used in this study is applicable. The results report that the most discriminating FT feature is 'decrementing amplitude', an accepted characteristic of Bradykinesia (Fernandez et al., 2014). The second most important FT feature is 'opening acceleration', and has similarity with findings in Yokoe et al. (2009), where they claimed opening velocity as a novel parameter in differentiating PD patients movement from healthy old people. The third most discriminating feature, 'halts' might be an important feature that was known but not deemed important before.

Subsequently, the same bradykinesia features used as input to FT classifiers were applied to hand pronation-supination (PS) data by replacing opening phase with pronation, and closing phase with supination movement. PS classifiers reached accuracies even better than FT classifiers and lead to several important features that might not have been given sufficient attention before. This is especially true because there have not been as many PD objective assessment studies where PS kinematic data is compared to FT. Whilst rhythm and periodicity of amplitude are not important to the FT task, they are the most discriminating features of PS.

Pronation-supination itself is a rhythmic movement; it is a motor task that is commonly used to test the ability of executing diadochokinetic movements. Diadochokinesia is ability to perform rapid, alternating movements (RAM), capability to cease current movement and start the opposite movement instantly after that. The assessment of ability to perform RAM can provide important information about the condition of the nervous system. Unfortunately, a significant amount of information can be lost during the clinical test where qualitative rating scales are used instead of quantitative assessment. The abnormality of diadochokinesia can be seen in the completeness of the sequence, and in the variation of amplitude. Furthermore, two other amplitude characteristics of PS, namely mean and maximum amplitude of both phases, are also deemed as important distinguishing features. The smaller amplitude means subjects did not get to finish the phases (of pronation or supination). The high accuracies of PS classifiers indicate that there is a significant difference between the way healthy old people and PD patients execute RAM such as hand pronation-supination. HO classifiers performance in the controls vs PD case is not very encouraging compared to the other two tasks. However, the accuracies are not so low that the classifications are unuseful. This may be due to the different nature of movement of HO; unlike FT and PS, HO involves two steps movement where thumb and fingers move at different phases. Hence, there are possibilities the same method of data collection and features extraction cannot be applied to HO movement for this purpose.

5.2 Bradykinesia in Parkinson's disease and Huntington's disease

The majority of previous quantitative studies to measure the existence of Bradykinesia in Huntington's disease (HD) focused on slowness of movement (refer section 2.5 in Chapter 2). Bradykinesia itself is not purely slowness, but is a complex disturbance of initiation and execution of actions and the ability to sustain them. Akinesia (failure to initiate movement) and hypokinesia (underactive movement) both relate to bradykinesia, as does the sequence effect—repetitive movements becoming smaller or slower. This part of study observed characteristics of Bradykinesia in Huntington's disease (HD) compared to Parkinson's disease (PD) and healthy old people (controls). As described in Chapter 4, three motor tasks were used, finger tapping (FT), hand opening-closing (HO) and hand opening-closing (PS). Eleven HD patients recruited for this study; however not all of them were able to finish all the three tasks. Table 5.8 lists the number of subjects for each group for different motor tasks.

Table 5.8: Numbers of healthy controls, PD and HD patients performed different motor tasks.

Task	Controls	PD	HD
Finger tapping	27	26	11
Hand pronation-supination	27	26	6
Hand opening-closing	27	26	10

5.2.1 CGP classification results

Classifiers were evolved using the same method as described in Chapter 4. Two cases were considered, healthy controls vs Huntington's disease patients (controls vs HD) and Parkinson's disease vs Huntington's disease (PD vs HD). Table 5.9 presents classification accuracies of controls vs HD for all three tasks in the form: mean ± standard deviation for both values of train and test set . The averaged accuracies are visualised in Figure 5.5 with mean, maximum and minimum accuracy plotted.

Table 5.9: Accuracies for classifiers of controls vs HD.

Controls vs HD

Task	Accuracy (%)			
	Averaged ten runs		Best classifier	
	Train	Test	Train	Test
Finger Tapping	92.36 ± 1.19	83.42 ± 2.11	91.41	91.38
Hand pronation-supination	93.61 ±1.06	85.98 ± 1.85	92.39	95.83
Hand opening-closing	90.63 ± 2.31	76.17 ± 5.60	98.75	77.5

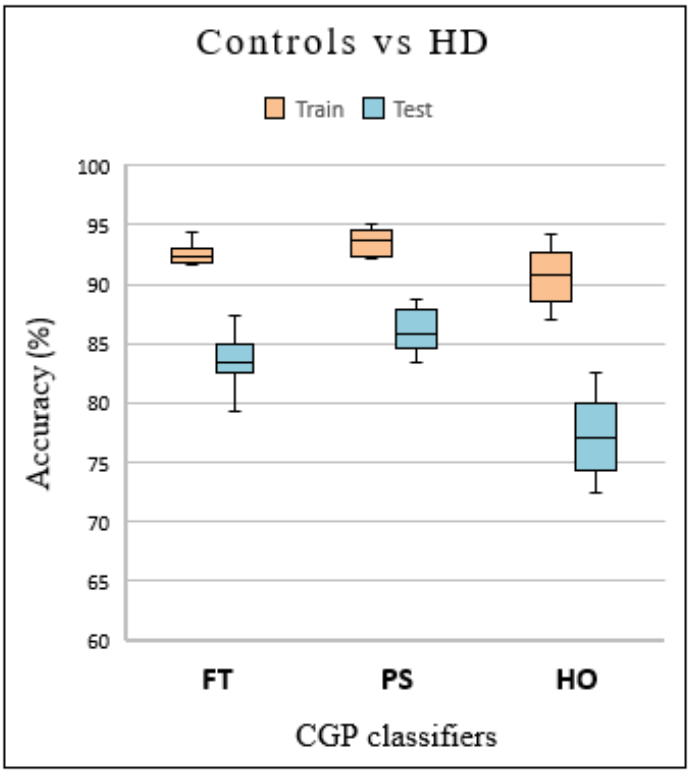

Figure 5.5: CGP Accuracies averaged across ten runs for case controls vs HD.

For the finger tapping task, accuracy of the test set across 10 runs for controls vs HD is 83.42 %. Slightly higher accuracy was obtained for PS classifiers at 85.98%. HO classifiers, as in case of controls vs PD, is still showing poor performance with only 76% accuracy. Overall performance of all classifiers shows that the selected Bradykinesia features are able to differentiate HD from healthy controls. In other words, it can be hypothesised that Bradykinesia does exist in Huntington's disease.

The following results provide information on whether Bradykinesia features in HD have the same characteristic as the features in PD patients' movement. The summary of the classifier's accuracies for PD vs HD is given in Table 5.10 in the form: mean ± standard deviation for both values of train and test set. The averaged accuracies are visualised in Figure 5.6 for comparison of CGP classifiers according to motor tasks.

Table 5.10: Accuracies for CGP classifiers of PD vs HD.

PD vs HD

Task	Accuracy (%)			
	Averaged ten runs		Best classifier	
	Train	Test	Train	Test
Finger Tapping	85.47 ± 1.19	67.97 ± 2.95	83.67	80.95
Hand pronation-supination	95.65 ± 0.85	87.20 ± 1.91	94.19	95.45
Hand opening-closing	93.38 ± 1.19	84.81 ± 3.73	90.28	97.22

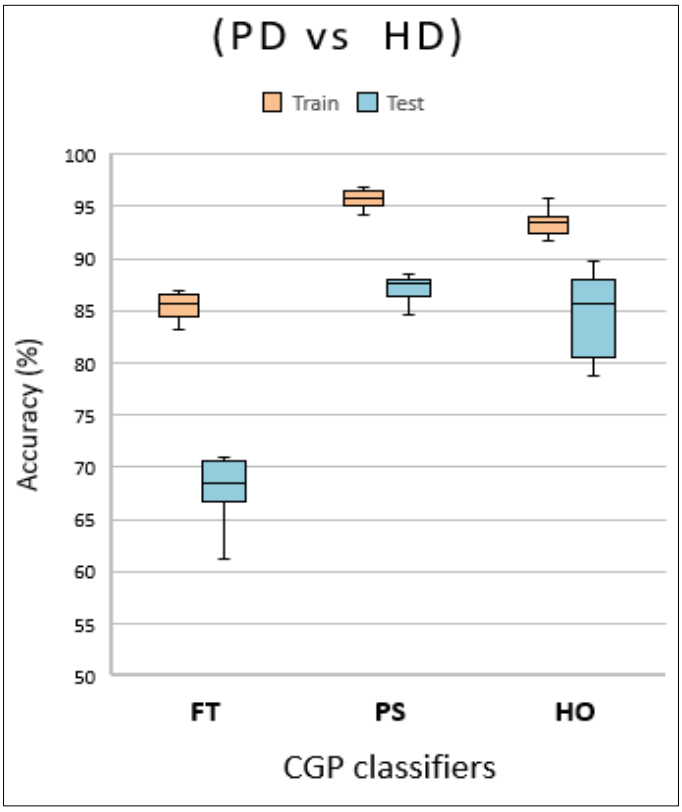

Figure 5.6: CGP Accuracies averaged across ten runs for case PD vs HD.

The best classifiers for PD vs HD case were evolved using kinematic data from the pronation-supination task (PS), with an averaged accuracy of 87.20%. FT classifiers did not show a good performance in differentiating PD from HD. In previous cases, the majority of HO classifiers achieved lower accuracies below 80%, but in the case of HD patients, HO classifiers obtained high accuracies of 85% for the PD vs HD case. This indicates that although HO movement features cannot reliably distinguish healthy controls from PD and HD, it can recognise HD very well when compared with PD patients. It is therefore interesting to determine which HO features evolved the best HO classifiers for class of PD vs HD.

5.2.2 Discriminating features

In the case of PD vs HD, both HO and PS classifiers achieved very good accuracies. However, the number of HD patients is higher for the HO task compared to participants that completed the PS task, giving more confidence in the results. Focusing on the potential of HO data in the classification of PD and HD patients, strong discriminative features have been noted. Table 5.11 shows the number of times every HO movement feature was used in evolving the best HO classifier for PD vs HD. This classifier has accuracy of 97%. CGP network of the chromosome for this classifier is visualised in Figure 5.7 where it can be seen that input number (6) and (3) were used repeatedly in finding the solution, which indicates that features *'thumb maximum deceleration'* and '*thumb maximum speed'* are among most discriminative features in this classification. Other important features are (0) '*maximum opening*', '*COV of speed'* for (7) thumb sensor and (13) finger sensor and (14*)* '*halts'*.

.

Table 5.11: Best HO classifier discriminating features for PD vs HD case.

Feature	Number of times used
Maximum opening	6
COV opening	2
HO frequency	3
TS maximum speed	6
TS average speed	3
TS maximum acceleration	3
TS maximum deceleration	8
TS COV speed	6
TS Halts	1
FS maximum speed	5
FS average speed	5
FS maximum acceleration	5
FS maximum deceleration	3
FS COV speed	6
FS Halts	6

5.2.3 SVM and ANN classifiers

In this section, comparisons are made between accuracies of the CGP classifiers with popular algorithms ANN and SVM. The parameters for ANN and SVM are as explained in Section 5.2.4. Two cases were considered, controls vs HD and PD vs HD. The results for all three types of classifiers are shown in Table 5.12 for controls vs HD. For fair comparison, CGP classifiers in the table used record-wise cv. It can be seen in the table that for the case of controls vs HD, ANN gained high accuracies in all tasks. In fact, using PS kinematic data, ANN get 100% accuracies predicting classes of instances in the test set. Results of SVM are comparable to CGP except for SVM PS classifiers where the accuracy is nearly 100%. Overall, the same patterns are shown by ANN and SVM where the highest accuracies achieved for PS, followed by FT and lowest accuracies for HO. This can be argued to validate the CGP results because the same ranking of accuracies is also shown by CGP.

Table 5.12: Accuracies of classifiers evolved using CGP and two other popular algorithms, SVM and ANN in classification of healthy controls vs Huntington's disease patients.

Controls vs HD

Task	Accuracy (%)					
	CGP (record-wise CV)		SVM		ANN	
	train	test	train	test	train	test
Finger Tapping	92.36	84.87	88.02	86.98	99.48	84.38
Hand pronation-supination	100	100	100	99.04	100	100
Hand opening-closing	88.33	81.67	84.17	77.5	90	80.83

In PD vs HD case, the same pattern of performance is shown by all classifiers where highest accuracies were obtained by PS classifiers, followed by HO and lowest accuracies by FT classifiers. CGP gained comparable results to SVM and ANN for all tasks (Table 5.13). As in previous cases, PS classifiers (for all three algorithms considered) showing the best performance. Poor performance by FT classifiers evolved using CGP are also reflected in ANN and SVM FT classifiers with most of classifiers obtaining accuracies below 70%. As described in Section 5.3.1, surprisingly, CGP evolved HO classifiers with high accuracies for the PD vs HD case. Similar results were gained by ANN HO classifiers. However, SVM HO classifiers do not show good results, with accuracy below 80%. This suggets ANN and CGP were learning characteristics that were missed by SVM.

Table 5.13: Results of classification accuracies using ANN and SVM algorithms compared to accuracies of classifiers evolved using CGP for PD vs HD case.

PD vs HD

Task	Accuracy (%)					
	CGP (record-wise CV)		SVM		ANN	
	train	test	train	test	train	test
Finger Tapping	83.176	78.57	74.29	72.14	95	66.43
Hand pronation-supination	99.602	98.258	100	98.21	100	97.32
Hand opening-closing	88.637	87.255	79.55	77.27	91.67	84.09

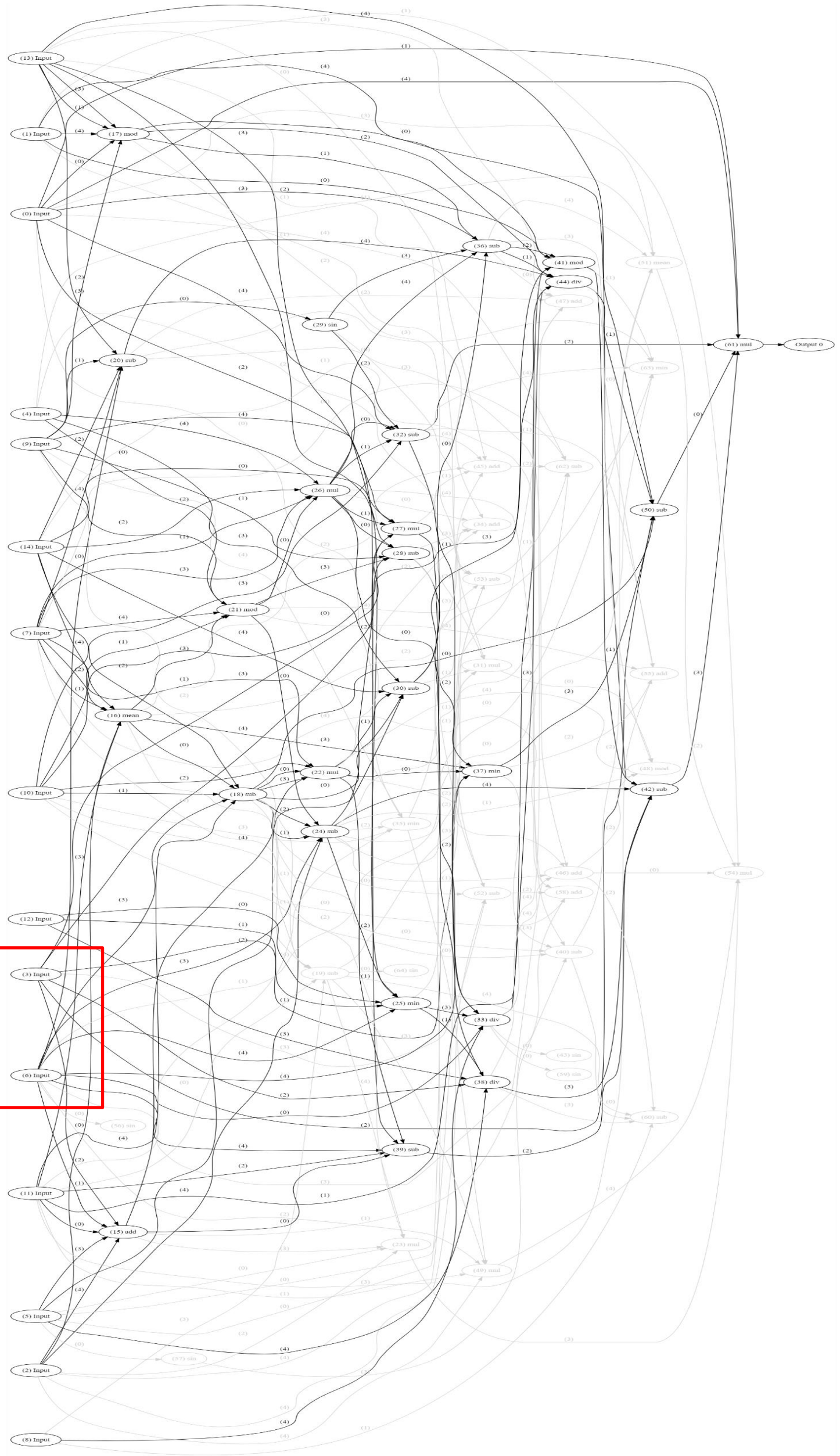

Figure 5.7: Visualised chromosome of the best classifier for PD vs HD class using HO data .

5.2.4 Discussion

FT as a popular task in machine learning classification, which in this work shows expected results of high accuracies in the case of controls vs PD. However, in HD study, compared to HO and PS classifiers, FT classifiers achieved the lowest accuracy when used to differentiate between PD and HD groups. This suggests that selected bradykinesia features in FT are not strong discriminative characteristics of PD and HD patients. PS classifiers continue to consistently achieved high accuracies in HD cases, but the low number of HD patients used in the PS task lessens the confidence in the PS classifiers accuracies.

This is the first time CGP was applied in classifying other movement disorders (apart from PD) using kinematic data of common clinical motor tasks. Both HO and PS achieved very good accuracies for the PD vs HD case. Focusing on the potential of HO data in classification of HD patients, discriminative features have been noted. The best HO classifier (PD vs HD) used the '*thumb sensor (TS) maximum deceleration'* feature eight times in finding the solution. Other important discriminative features are '*maximum opening* (0)', '*COV of speed'* for the thumb sensor (7) and finger sensor (13) and '*halts'* (14). The results suggest that bradykinesia features in healthy old people, PD and HD patients have different characteristics.

5.3 Bradykinesia in Parkinson Disease and Essential Tremor

The objective of this part of the study is to identify the existence of Bradykinesia in Essential tremor (ET). Although ET has been repeatedly mistaken for PD and vice-versa, the symptom perceived as the cause of the confusion is tremor instead of slowness. However, as mentioned at the beginning of this chapter, a few studies found slowness in ET patients compared to healthy old people. If Bradykinesia does exist in ET, then it is important to know if Bradykinesia in ET has the same characteristics as in PD. In pursuit of this, kinematic data of six ET patients was collected for preliminary experiments. Using extracted data of the three motor tasks from the MDS_UPDRS, classification using CGP, SVM and ANN were run for two cases: controls vs ET and PD vs ET.

5.3.1 CGP classification results

Table 5.14 shows classification accuracies of controls vs ET for all three tasks. The averaged accuracies are visualised in Figure 5.8 with mean, maximum and minimum accuracy plotted. It can be seen that the same pattern as seen in controls vs HD had reoccurred where the best performance was shown by PS classifiers with an averaged accuracy of 98.77% followed by FT classifiers with 85.99% accuracy and 78.47% accuracy for HO classifiers.

Table 5.14: Accuracies for CGP classifiers for the case of controls vs ET.

Controls vs ET

Task	Accuracy (%)	
	Averaged ten runs	
	Train	Test
Finger Tapping	95.28 ± 1.01	85.99 ± 1.70
Hand pronation-supination	100.00 ± 0	98.77 ± 1.42
Hand opening-closing	92.09 ± 1.82	78.47 ± 3.19

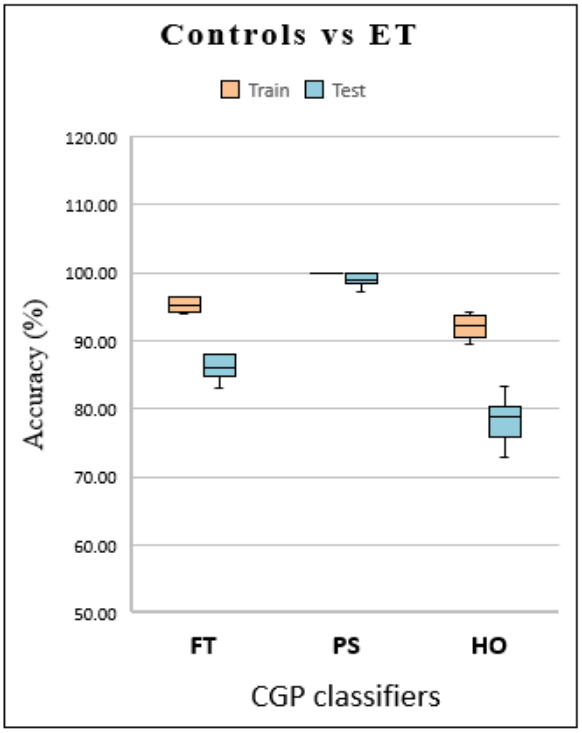

Figure 5.8: CGP Accuracies averaged across ten runs for case controls vs ET.

In the case of PD vs ET, PS classifiers correctly classified 90% of the instances in their correct classes (Table 5.15). In both Table 5.14 and Table 5.15, the result reported is in the form: mean ± standard deviation, where the mean and the standard deviation of the accuracies for both values of train and test set are stated. Although the other two tasks only evolved classifiers with accuracies ranging from 75% to 79%, with a higher number of ET participants better performance of FT classifiers might be expected. The performance of all classifiers across 10 runs are visualised in Figure 5.9.

Table 5.15: Accuracies for CGP classifiers of the case of PD vs ET

PD vs ET

Task	Accuracy (%)	
	Averaged ten runs	
	Train	Test
Finger Tapping	91.43 ± 1.57	78.59 ± 2.81
Hand pronation-supination	97.73 ± 1.44	88.94 ± 2.57
Hand opening-closing	91.89 ± 1.08	75.00 ± 4.09

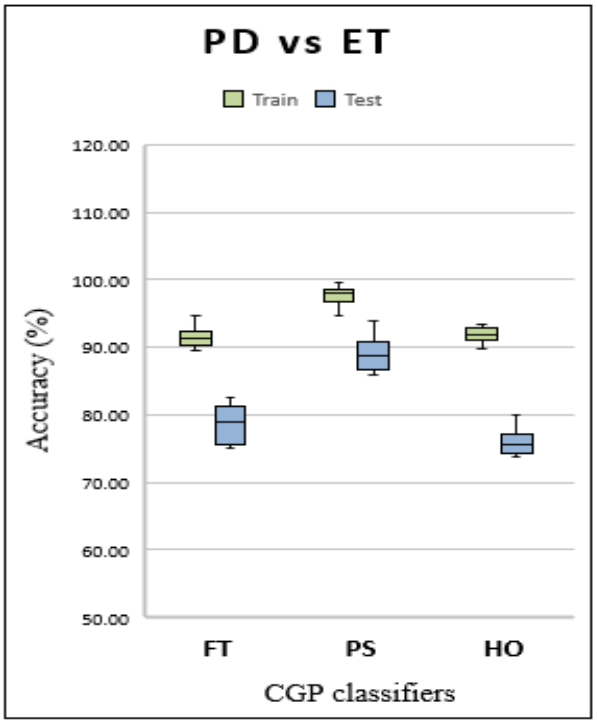

Figure 5.9 : CGP Accuracies averaged across ten runs for case PD vs ET.

5.3.2 SVM and ANN classifiers

In the case of controls vs ET, CGP classifiers performed better than SVM and ANN in all motor tasks. It is argued this is due to CGP's better capabilities of finding connection between instances of the same subject as record-wise cv is used here. CGP PS classifiers classify all instance in the correct classes. Very small differences in performance are achieved by CNN and SVM at 97% and 98% accuracies, respectively. HO classifiers performance is almost uniform between all three algorithms at the border of 80% accuracies.

Table 5.16: Performance of CGP, SVM and ANN in classification of controls vs ET.

Controls vs ET

Task	Accuracy (%)					
	CGP (record-wise CV)		SVM		ANN	
	train	test	train	test	train	test
Finger Tapping	92.59	89.71	87.9	85.06	94.83	86.21
Hand pronation-supination	100	100	98.08	98.08	100	97.12
Hand opening-closing	87.50	81.001	79.81	76.92	86.54	79.81

In the case of PD vs ET, ANN induced the best classifiers in two out of three motor tasks. Only in case of HO classifiers, CGP performed better. In almost all motor tasks, SVM performances are in between of CGP and ANN.

Table 5.17: Performance of CGP, SVM and ANN in classification of PD vs ET.

PD vs ET

Task	Accuracy (%)					
	CGP (record-wise CV)		SVM		ANN	
	train	test	train	test	train	test
Finger Tapping	88.52	83.653	81.15	80.33	99.18	90.16
Hand pronation-supination	97.82	91.213	91.96	91.07	100	96.43
Hand opening-closing	88.41	79.224	80.43	78.26	100	76.09

5.3.3 Discussion

Although the number of ET participants in this study is small, the same pattern of performance between motor tasks (with our studies with higher number of participants; controls vs HD, controls vs HD) indicates that the results of this study can be used to argue the existence of Bradykinesia in ET. High-performance classifiers of three algorithms; CGP, SVM and CGP in differentiating ET from healthy controls supports the argument that Bradykinesia features exist in ET. Bradykinesia is not the main source of confusion between PD and ET, but this study indicates that by using Bradykinesia

features, CGP and other machine learning algorithms managed to perform well in differentiating the two disorders. This suggests that the two disorders have different characteristics in their bradykinetic features.

5.4 Conclusion

Each classification in this study plays an important role in the characterisation of Bradykinesia in PD, HD and ET. Most previous machine learning studies on PD motor symptoms focus on classifying PD patients from healthy controls, as reviews on machine learning classification of PD testify (Ahlrichs & Lawo, 2013). Yet, to address the problem of misdiagnosis, the real challenge is to differentiate between groups of different movement disorders and observe the discriminating features. FT as a popular task in machine learning classification shows expected results of high accuracies in case of controls vs PD. The overall accuracy was lower than shown in previous studies of CGP classifications using finger tapping data (Lacy et al., 2013; Lones et al., 2014), but it is suggested that this is due to smaller numbers of subjects in this study. One of the original intentions of this study was to extend usage of finger tapping (in classification of PD) to other motor tasks. Clearly the results in controls vs PD case show that the PS task can also be used for this purpose. Although HO classifiers accuracies are not as encouraging compared to the other two tasks, the performance is the classification is still quite good. Most HO classifiers achieved accuracies above 70%.

The most interesting cases in this study are regarding HD patients. This is the first time CGP was applied in classifying other movement disorders (apart from PD) using kinematic data in common clinical motor tasks. PS and HO classifiers achieved very good accuracies for the HD cases. The fact that HD patients can be differentiated from healthy control using extracted Bradykinesia features supports the argument that Bradykinesia

does exists in Huntington's disease. In the case of PD vs HD, both PS and HO classifiers show good accuracies, but with higher number of participants in HO tasks giving more confidence to HO classifiers compared to PS classifiers. Hence, the focus of investigating discriminating features (of PD vs HD case) was given to HO classifiers. Several HO movement features were highlighted as strong discriminative features for PD vs HD cases. Among the most important HO features are the speed of thumb movement during the slowing down process and the rhythm of speed of the thumb and fingers during hand opening and closing. Another exciting finding is that classifiers evolved using pronation-supination data not only achieved comparable accuracy with FT classifiers in all cases but are able to classify better in many cases. This implies that features of bradykinesia in PS movement have highly distinguishing characteristics between the classes (controls, PD and HD). Surprisingly, FT classifiers achieved low accuracy for PD vs HD cases. This suggests that selected bradykinesia features in FT are not the most distinctive characteristics of PD and HD patients.

The second study of other movement disorders (than PD) involved patients of Essential Tremor. Only six ET patients were available, but sufficient to provide some insights on Bradykinesia in ET. High-performance PS and ET classifiers of three algorithms, CGP, SVM and CGP, in differentiating ET from healthy controls, suggest that Bradykinesia features exist in ET. Bradykinesia is not the main source of confusion between PD and ET, but this study indicates that by using Bradykinesia features, CGP and other machine learning algorithms managed to perform well in differentiating between the two disorders. It is suggested that the two disorders have different characteristics of bradykinetic features but a higher number of participants can confirm this.

All the CGP results were validated by comparing them to results obtained using widely adopted machine learning algorithms, SVM and ANN. It is the common practice in statistical machine learning to perform classification by random instances divided into

training and test sets without consideration of the instances (records) belonging to the same participant or not (record-wise cv). Therefore, to obtain results that can be compared with previous studies, record-wise cv was used to compare the three algorithms (CGP, ANN and SVM) performances. The results show that the same pattern of performance were demonstrated by SVM and ANN as in CGP classifiers. For example, in the controls vs PD case, the best CGP classifiers were evolved using PS kinematic data, followed by FT classifiers and lowest accuracies obtained by HO classifiers. This ranking of performances also shown by SVM and ANN, validate the CGP results. Overall, CGP accuracies are comparable and sometimes higher than ANN and SVM classifiers.

CHAPTER 6

Differentiation of Organic and Functional Dystonia

In 2013, an international Consensus Committee, consisting of investigators with years of experience in dystonia agreed to proposed revised definition of dystonia as following (Albanese et al., 2013) :

- Dystonia is a movement disorder characterized by sustained or intermittent muscle contractions causing abnormal, often repetitive, movements, postures, or both.
- Dystonic movements are typically patterned, twisting, and may be tremulous.
- Dystonia is often initiated or worsened by voluntary action and associated with overflow muscle activation

Currently, there is no known cure for Dystonia, but effective treatments, coping strategies and support are available. In most cases, dystonia does not shorten a person's lifespan. Historically, Dystonia lacked unifying criteria to put it under a type of disorder; it is not until late 1970's that Dystonia was recognised as its own class of movement disorder. Before that, generalised Dystonia have been managed along the psychiatric line, viewed as not organic disorders (Newby et al., 2017). After Dystonia was recognised as

a class of movement disorder, the psychiatric part of Dystonia still remained, it is known as Psychogenic Dystonia and later renamed to Functional Dystonia (FD). Functional dystonia is a condition where some specific symptoms of dystonia appear, but tests that normally establish the cause of these symptoms are negative. It is part of what is known as Psychogenic Movement Disorders (PMDs). PMDs can look like any organic neurologic condition, but it has a psychiatric cause. A variety of blood, imaging and other tests are usually normal and do not reveal any physical (organic) cause that could explain these motor abnormalities. Functional dystonia is arguably the most challenging functional movement disorder to diagnose and manage. It is especially difficult to differentiate FD from its organic version, Organic Dystonia (OD). Consequently, delays in diagnosis and treatment initiation have a negative impact on outcome, increasing the risk of long-term disability and the resulting financial burden on society. Hence, there is a growing emphasis on the need for 'laboratory supported' criteria, based on objective measurements of movement vectors and/or muscle activity. In order to support this need, this study was designed for objective assessment using evolutionary algorithm by application of classification task between organic and functional dystonia.

6.1 Types of Dystonia

Classification of Dystonia can be made along several dimensions. Among important dimensions or axes include body distribution, age onset, and cause of the dystonia. (Geyer & Bressman, 2006). In details, dystonia experts consensus suggested classification according to recognised using two axes (Albanese et al., 2013):

Axes 1: clinical characteristics

Axes2: aetiology (causes) of the disorder

6.1.1 Axes 1: Clinical characteristics

The clinical characteristics describe the phenomenology of dystonia in a given patient. Descriptors that can be used to specify clinical characteristics:

i. Age at onset

- Infancy dystonia (birth to 2 years)
- Childhood dystonia (3–12 years)
- Adolescence dystonia (13–20 years)
- Early adulthood dystonia (21–40 years)
- Late adulthood dystonia (>40 years)

ii. Body distribution

Body regions involved by dystonia are the upper or lower cranial region, the cervical region, the larynx, the trunk, the upper or lower limbs. These different territories may be involved individually or in different combinations.

Focal. Only one body region is affected. Typical examples of focal forms are blepharospasm, oro-mandibular dystonia, cervical dystonia, laryngeal dystonia, and writer's cramp. Cervical dystonia is considered a form of focal dystonia, although by convention the shoulder can be included as well as the neck.

Segmental. Two or more contiguous body regions are affected. Typical examples of segmental forms are cranial dystonia (blepharospasm with lower facial and jaw or tongue involvement) or bi-brachial dystonia.

Multifocal. Two non-contiguous or more (contiguous or not) body regions are involved.

Generalized. The trunk and at least two other sites are involved. Generalized forms with leg involvement are distinguished from those without leg involvement.

Hemidystonia. More body regions restricted to one body side are involved.

Typical examples of hemidystonia are due to acquired brain lesions in the contralateral hemisphere.

iii. Temporal pattern

Disease course can be either static or progressive. The variability can have four different patterns:

Persistent. Dystonia that persists to approximately the same extent throughout the day.

Action-specific. Dystonia that occurs only during a certain activity or task.

Diurnal fluctuations. Dystonia fluctuates during the day, with recognizable circadian variations in occurrence, severity and phenomenology.

Paroxysmal. Sudden self-limited episodes of dystonia usually induced by a trigger with return to pre-existing neurological state.

iv. Coexistence of other movement disorders

Isolated dystonia. Dystonia is the only motor feature, with the exception of tremor.

Combined dystonia. Dystonia is combined with other movement disorders (such as myoclonus, parkinsonism, etc.).

v. Other neurological manifestations.

The presence or absence of other neurologic or systemic features is a vital component for characterizing dystonia syndromes. Wilson disease is a disorder where dystonia is typically combined with other neurological or psychiatric symptoms and liver disease 30. The broad neurological spectrum evolves over time, with frequent revisions as new information is gained.

Axes 2: aetiology —although there is effort to redefine the categories of Dystonia regarding the causes (Albanese et al., 2013), the still popular etiologic classification divide Dystonia into two broad categories: primary (idiopathic; without no known cause) and secondary (symptomatic) (Albanese & Jankovic, 2012). Primary and secondary Dystonia according to (Geyer & Bressman, 2006) are as following:

Primary Dystonia

In primary dystonia, no abnormality other than dystonia is present, apart from the occasional occurrence of tremor (resembling essential tremor) or myoclonus. Findings such as parkinsonism, seizures, dementia, ataxia, ocular motor abnormalities, weakness, or spasticity suggest that dystonia is secondary. Moreover, when tremor or myoclonus accompanies primary dystonia, these associated movements are less prominent than the dystonia. If these movements predominate, a secondary dystonia is likely. In primary dystonia, there are no structural brain abnormalities on radiographic studies and no inborn errors of metabolism identifiable with conventional investigations. Most primary dystonias are focal or segmental in distribution, with onset in adulthood. Around 10% of patients with primary dystonia have generalised dystonia, usually starting in childhood or adolescence.

Secondary Dystonia

When dystonia is secondary to a hereditary neurological disorder or an exogenous insult, additional neurological abnormalities are likely to be present. An important exception is dystonia resulting from dopamine receptor blocking agents (acute dystonic reaction and tardive dystonia), which usually consists of dystonia only. One subcategory of secondary dystonia comprises the dystonia-plus syndromes. In these inherited disorders, dystoniais accompanied by other neurological abnormalities, but like in the primary dystonias, there is no evidence of brain degeneration. The dystonia-plus syndromes include dopa-responsive dystonia, myoclonus- dystonia, and rapid-onset dystonia-parkinsonism

6.2 Organic and Functional Dystonia

Classification of Dystonia as in Section 6.1 is describing the 'real' or Organic Dystonia (OD). When a patient showing some specific symptoms of dystonia appear, but tests that normally establish the cause of these symptoms are negative, it is what is known as Psychogenic or Functional Dystonia (FD). It is part of what is known as Psychogenic movement disorders (PMDs). Most psychogenic movements are considered involuntary – performed without conscious awareness or effort. Other types of PMDs include Psychogenic tremor (also called functional tremor), Psychogenic parkinsonism and Psychogenic gait disturbances.

6.2.1 Diagnosis of Organic and Functional Dystonia

The diagnosis of dystonia, like that of all neurological disorders, rests most firmly on the history and physical examination. A clinical diagnosis of OD is made through recognition of its core motor features and their distribution and temporal evolution. Non-motor features include disturbed sensory and cognitive processing and psychiatric features

(Albanese et al., 2013). A complex interplay between motor, sensory and limbic centres may give rise to odd and atypical features (e.g. task-specificity, geste antagonist), potentially misleading the inexperienced clinician. The resulting misdiagnosis rate is 25-52% (Stamelou et al., 2012).

On the other hand, a diagnosis of FD rests on assessment of inconsistency and incongruence (in both the history and neurological examination) with organic disease patterns (Espay & Lang, 2015; Ganos et al., 2014). Functional dystonia is arguably the most challenging functional movement disorder to diagnose and manage. Current diagnostic criteria place disproportionate weight on features with poor predictive value: historical features (such as sudden-onset or stress-induced symptoms) and psychiatric comorbidity, which are also prevalent in organic dystonia. Consequent delays in diagnosis and treatment initiation have a negative impact on outcome, increasing the risk of long-term disability and the resulting financial burden on society. Hence there is a growing emphasis on the need for 'laboratory supported' criteria, based on objective measurements of movement vectors and/or muscle activity.

6.3 Classification of organic and functional Dystonia using EA

Research to date has shown that diagnostic criteria for FD have poor inter-rater reliability (Ganos et al., 2014) placing emphasis on historical variables and/or psychiatric comorbidity, which have poor predictive value. The need for 'laboratory supported' criteria for FD, based on electrophysiological signifiers, has been highlighted (Espay & Lang, 2015). Such criteria would permit earlier diagnosis and treatment, reducing the risk of long-term disability and the accompanying health and social care costs (the estimated UK cost for all 'medically unexplained symptoms' being £18 billion) (Hons et al., 2010) In previous quantitative studies, characteristics of Bradykinesia such as variability

(coefficient of variation), rhythm and slowness while performing finger tapping test were found as indications of malingering or psychogenicity (Arnold et al., 2005; Kalogjera-Sackellares & Sackellares, 1999; Rapport et al., 1998). For example, tapping was less rhythmic with higher variability in both malingering and psychogenic disorders. Subjects with suspected malingering performed the finger tapping more slowly than their comparison group counterparts regardless of the type of neurological diagnosis. Objective assessment studies differentiating organic and functional disorders are not common (Criswell et al., 2010). In fact, there are no studies that we know of trying to objectively differentiate FD and OD using kinematic data of any motor task. Based on these facts, this study employed Cartesian Genetic Programming for classification of organic dystonia, functional dystonia and healthy controls using movement data of finger tapping task as the input. The method of movement data collection, data processing and classification steps explained in Chapter 4. However, there is a slight change made for Dystonia study in term of movement data collection method. In PD study (as described in Chapter 4), participants executed finger tapping for ten times as suggested by MDS-UPDRS with two repetitions for each hand. However, in this study, participants tap their fingers for 15 seconds. There are three repetitions for each hand which means, for each participant, six finger-tapping movement recordings were obtained. There are altogether 29 healthy controls, 31 Organic Dystonia and 12 Functional Dystonia patients.

Participants with dystonia were recruited from the current caseloads of the movement disorders consultants at Monash Medical Centre in Melbourne (Australia) and Leeds General Infirmary, Leeds, UK. Assessments took place between September 2015 and February 2018. Ethical approval was obtained from the Monash Health Human Research Ethics Committee (HREC code: 13424B) and the Yorkshire and Humber Sheffield Research Ethics Committee (HREC code: 14/YH/0143).

Patients and controls subjects were chosen following the three inclusion criteria:

i. **Organic dystonia**: Expert diagnosis, according to accepted guidelines, with upper limb or cervical involvement (genetic, idiopathic focal or secondary);
ii. **Functional Dystonia**: Expert diagnosis, documented or clinically established according to the Fahn-Marsden criteria, with upper limb involvement;
iii. **Controls**: Capacity to consent and able to perform assessments. Control subjects were recruited from spouses and friends of the patients who attend clinics at the Monash Medical Centre.

Subjects were removed from the study according the following exclusion criteria:

- Aged under 18 years;
- Lacking capacity to consent;
- Unable to perform movement assessments (e.g. due to cognitive deficit).

6.3.1 CGP classification results

CGP classifiers were evolved using the same method described in Chapter 4 using movement data from the MDS-UPDRS defined finger tapping task. Three cases of pairwise classification were considered:

- Controls vs OD
- Controls vs FD
- OD vs FD

Table 6.1 shows classification accuracies of all the three cases in the form: mean ± standard deviation, where the mean and the standard deviation of the accuracies for both

values of train and test set are stated. The averaged accuracies across ten runs and the accuracies of the best run is visualised in Figure 6.1 with mean, maximum and minimum accuracies plotted.

Table 6.1: CGP classifiers accuracies for all three cases of pairwise classification considered: Con vs OD, Con vs FD and OD vs FD.

Task	Accuracy (%)			
	Averaged ten runs		Best run	
	Train	Test	Train	Test
Finger Tapping	64.30 ± 0.62	60.12 ± 3.19	65.55	64.58
Hand pronation-supination	83.84 ± 0.97	74.33 ± 2.55	85.62	77.00
Hand opening-closing	81.45 ± 1.09	70.09 ± 1.94	83.78	72.49

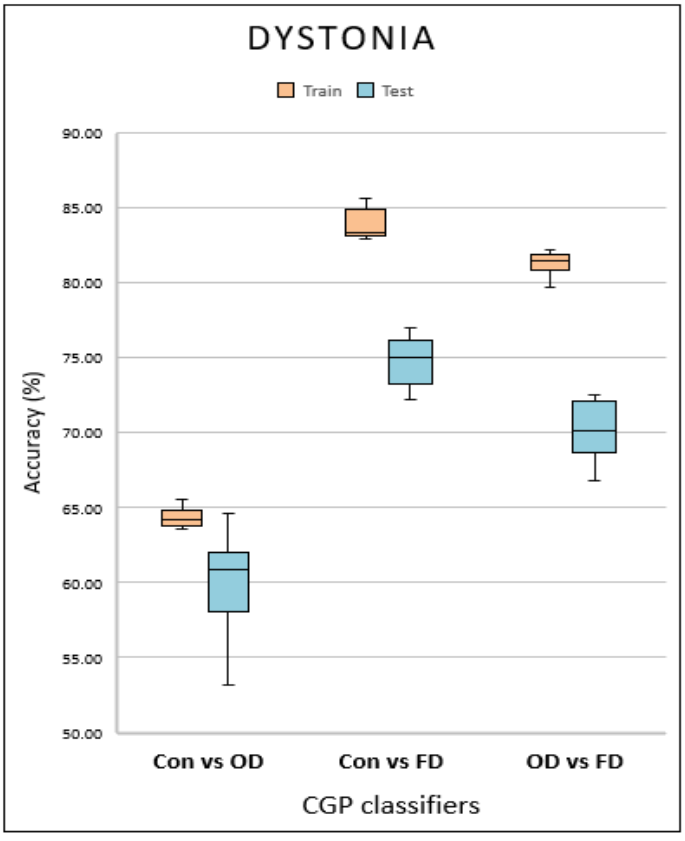

Figure 6.1: CGP accuracies averaged across ten runs for compare according to classification cases considered

None of the CGP classifiers evolved in any of the cases, achieved an accuracy beyond 80%. The classification for the case between controls and OD patients has the worst accuracies, with some of the CGP classifiers with accuracies below 55%. A better performance is shown by classifiers evolved for the case of controls vs FD, achieving results as high as 77% accuracy. Surprisingly, the case that is considered most challenging to diffrentiate, OD vs FD, achieved accuracies ranging from 68% to 77%. The best-run accuracies do not show any vast difference from the averaged values.

6.3.2 Comparison with SVM and ANN

Comparisons were made between accuracies of CGP classifiers and popular machine learning algorithms; SVM and ANN. SVM and ANN parameters used in this study discussed in Section 5.2.4 of Chapter 5.

Table 6.2: Performances of SVM and ANN classifiers for all cases of Dystonia study compared to CGP classifiers

Case	Accuracy (%)					
	CGP (best run)		SVM		ANN	
	train	test	train	test	train	test
Controls vs OD	65.55	64.58	66.10	64.41	89.55	66.95
Controls vs FD	85.62	77.00	77.92	75.00	93.33	68.75
OD vs FD	83.78	72.49	78.00	66.81	84.00	62.92

In all cases, results using CGP and other algorithms are comparable with very small differences between each other. ANN shows signs of overfitting where the accuracies of the training sets were much higher than the those of the test set.

6.4 Discussion

It is a known fact that diagnosis of FD is very difficult. It might be the main reason for poor performance of all machine learning algorithms applied in this study. The fact that all three algorithms, CGP, SVM and ANN, produced comparable results demonstrates that the problem is with the data rather than the method applied.

There are several possible reasons why the data is unreliable. The problem might be with regard to the selected motor task. Although a few previous studies using finger tapping data of Psychogenic Movement Disorders (PMD) shows significant difference between PMD and healthy controls, the studies did not specifically address Functional Dystonia. The movement characteristics of Functional Dystonia and other PMDs might be different. The second reason might be the participants. Dystonia patients in this study consists of various subtypes of Dystonia. Although all Dystonia participants have affected upper limb movements, the symptoms might not be shown in the hand only. In future it is suggested that the use of a combination of motor tasks is considered that involve other affected body parts as well. The third problem might be in number of instances available. The number of participants in the Dystonia study is good compared to PD study. However, due to the high variation of symptoms in Dystonia patients, the study needs more participants for the algorithms to better predict these.

The differentiation of PMD from its organic version is very important to support the clinical diagnosis. The misdiagnosis rate is too high for the costs to be ignored. Therefore, it is suggested that this study be continued but by looking into different arrangements of experimental setup and higher numbers of participants.

CHAPTER 7

Conclusions and future works

In this chapter, a summary of the whole study is presented, stating the findings deemed important and suggestions for future work. Finally, the hypothesis and objectives of the study stated in Chapter 1 are revisited.

7.1 Research summary

The real-world aim of this study is towards improving diagnosis and monitoring of neurological conditions, particularly the conditions that cause movement disorders. Neurological conditions are the biggest caused of loss of healthy life in the world population. The main motivations for this study stem from the multiple impacts of movement disorders to the life of patients, family, society and a country's economy. The symptoms of the condtions cause a decrease in quality of life (QOL) in many ways including loss of independence, negative impact on relationships and work, and having difficulties performing day-to-day activities (Dowding et al., 2006; Soh et al., 2013; World Health Organization, 2006). The disorders affected QOL not only of the patients but the caregivers as well (Carter et al., 2008; Martínez-Martín et al., 2005, 2007). The

degrading quality of life of patients and caregivers is especially true for neurodegenerative disorders such as PD and HD where the burden increase as the disease progress. The economic cost of movement disorders is high, involving public health activities that include formal health care such as personal medical attention, home care and self-care, health promotion and disease prevention (World Health Organization, 2006). The effect is big and global. In 2005, over one million people in Europe were diagnosed with Parkinson's disease, with around 127000 of them living in the United Kingdom. It is projected that the number will double by 2030 (Dorsey et al., 2007). Approximately 2.2% of the US population have essential tremor, smoe 6.38 to 7.63 million people (Louis & Ottman, 2014). Until now, there are no cures for movement disorder conditions; however, accurate and early diagnosis can help lessen the effects through effective management of the disease. Unfortunately, for most movement disorders, diagnosis proves to be difficult and unreliable and requires experts in the field to achieve higher accuracy. This is largely due to the fact that the diagnosis of movement disorders remains clinical. Clinical diagnosis is based on observation of clinical features from examination and history taking and response to medication. Clinical diagnosis is not straightforward, usually complex and may be variable. This is the main reason why misdiagnosis rate in movement disorders are high. In Parkinson's disease, it is estimated that the misdiagnosis rate remains as high as 25% since 1970's (Ali H Rajput & Rajput, 2014). About one in three patients with other tremor conditions were misdiagnosed as having ET, with the most frequent incorrect diagnoses being Parkinson's disease and dystonia (Jain et al., 2006). In additon, due to their sometimes-unusual presentations, patients with movement disorders may be diagnosed as having disorders caused by phycological problems (psychogenic movement disorders, PMD). The misdiagnosed rate of PMD is also high at 25-52% (Pal, 2011).

In effort to overcome difficulties caused by clinical diagnosis, objective assessments have been suggested. Many previous potential objective assessment methods in movement disorders are based on motor symptoms. The studies include using electronic sensors to collect movement data. For example, using patient's performance in drawing an Archimedes spiral, handwriting, rapid alternate movement or doing an everyday task such as walking. From these studies, many important movement features have been extracted and proven as significant features in differentiating the groups healthy old people and movement disorders. However, statistical analysis applied in these studies is usually focused on understanding the data (descriptive statistics) and making statements about the analysis of in-hand data (inference). Statistical analysis is often poor in the prediction of unseen data points. In order to support clinical diagnosis, the prediction part is obviously crucial. The ability to put unseen data into the correct classes, without any prior assumptions about the data can, however, be achieved by classification tasks using machine learning methods.

Considering previous efforts towards objective diagnosis, this study was conducted by employing EA on kinematic data of standard motor tasks. First, movement data collection was done by clinicians and consultant Neurologist in two research centres, Leeds Teaching Hospitals NHS Trust, UK and Monash Medical Centre, Melbourne, Australia. In the first phase, healthy controls and idiopathic PD patients were recruited. A total of 26 controls and 27 PD patients finished three standard motor tasks from MDS-UPDRS, namely finger tapping (FT), hand pronation-supination (PS) and hand opening-closing (HO). The movement data were collected in form of recordings of the positions of electromagnetic sensors worn by participants. Polhemus, a motion tracker system was used for the purpose. In the second phase, 11 HD patients and 6 ET patients were recruited. The last phase involved data collection of Dystonia patients.

The obtained data was then pre-processed using the Digital Signal Processing Toolbox™ in MATLAB®. Simple filtering, a Butterworth filter, was used to filter out noise. Exploration of the data were done to understand the data. Every data file was plotted, checked for missing values and corruption. For example, the maximum distance between two sensors can indicate if the data was corrupted. This can happen if there were strong magnetic sources around the motion tracking system or the basic arrangement of the experiment protocol was not followed. The details of the recording process is explained in Chapter 4.

The next step is to extract useful features that might be able to differentiate movement of participants from different groups. Focus was on features of Bradykinesia. This is due to the importance of Bradykinesia feature in PD as explained in Chapter 2. Previous quantitative studies of PD were used as basis for features selected. For finger tapping and pronation-supination, 19 movement features were extracted. In case of hand-opening task, 15 features were extracted. The same features in PS and FT cannot be used for HO due to the different nature of HO movement. Unlike in FT, where distance of thumb and finger are basis of the extracted features, most of the features in HO considered the movement of the thumb and fingers separately. Basic statistical tests were conducted to find out if there were significant differences between groups of healthy controls and the movement disorders data. The extracted features were then used as inputs to evolve CGP classifiers. CGP's advantages were discussed in Chapter 3. A CGP platform (and library) was used to evolve the classifiers (Turner & Miller, 2014). These classifiers were trained and validated to differentiate groups of healthy elderly people (controls) and movement disorders patients.

The disorders considered are Parkinson's disease, Huntington Disease patients and Essential Tremor. Altogether, there are six cases considered in this study:

- Controls vs PD
- Controls vs HD
- Controls vs ET
- PD vs HD
- PD vs ET
- Controls vs Dystonia
- Organic Dystonia vs Functional Dystonia (OD vs FD).

Based on the accuracies of evolved CGP classifiers, the best classifiers were selected to investigate features that contributed most to the discrimination. Only classifiers from the cases that achieved avaerged accuracies above 80% were considered.

Other machine learning algorithms were also used for comparison and validation. Two popular machine learning algorithms, Support Vector Machines (SVMs) and Artificial Neural networks (ANNs) were chosen. Accuracy is used to measure the performance of classifiers evolved and two forms of cross-validation (cv) were applied. The first approach was to use subject-wise cv where records (repetitions of recordings) from the same participant are not used in both the training and the test sets of the same fold. This is to avoid dependency between instances in training and test set that may cause bias. Stratified repeated k-fold cross-validation was used to assure statistical significance. However, it is quite common practice to disregard the precaution and instances from the same participant picked randomly to be included in both the training and test sets (record-wise cv). Therefore, it was decided to use both approaches to provide a meanful

comparison. Therefore, record-wise cv was used to compare performance of the three algorithms (CGP, SVM and ANN) with previous studies.

7.2 Study findings and future works

This study used one set of experiments with applications to different groups of movement disorders. To the best of knowledge, this is the first time a machine learning study that the three components (motor tasks) from the UPDRS have been measured together on the same patient. This is also the first time that comparisons were made between the performances of machine learning classifiers evolved using the kinematic data of all the three components.

7.2.1 Study findings

To conclude the study, useful findings on the application of EA and other machine learning algorithms on the kinematic data of the motor tasks are presented. Following, are conclusions and observations that can be made based on the results of the classification results and scrutinisation of the chromosomes from the best classifiers evolved.

i) CGP Classifiers using PS kinematic data performed better than FT classifiers in PD study

In the PD study, CGP classifiers had been evolved using kinematic data of three types of motor tasks; finger tapping (FT), hand pronation-supination (PS) and hand opening-closing. These classifiers were trained and validated to differentiate groups of healthy elderly people (controls) and movement disorders patients. The disorders considered were Parkinson's disease, Huntington Disease and Essential Tremor. In all five cases considered for the classification task involving PD and other movement disorders (listed in section 7.1), PS classifiers showed the highest accuracies compared to classifiers

evolved using the other two tasks. This indicates that Rapid Alternating Movements (RAM) such as hand pronation-supination might be a better motor task to differentiate PD from healthy controls and other movement disorders, compared to finger tapping.

ii) HO classifiers differentiate PD from HD better than FT classifiers

FT classifier accuracies are comparable to PS classifiers in most cases except for the case of PD vs HD, where FT classifiers achieved accuracy of 70% at best. HO classifiers showed less convincing performance with the majority of the classifiers' accuracies ranging from 65 % to 78%, in all cases except for the case of PD vs HD, where averaged best run achieved an accuracy of 84%. This indicates that HO is a better task to use in differentiating PD from HD.

iii) Bradykinesia exists in HD and ET

There is research suggesting that bradykinesia is not a core feature of movement disturbance in HD (Duval et al., n.d.; Fenney et al., 2008). It is interesting to note that in this research, the only feature of Bradykinesia considered was slowness, when in fact Bradykinesia comprises other features, as explained in Chapter 2. The same pattern occurred in the case of Essential Tremor. Some researchers suggested Bradykinesia exists in ET (Duval et al., 2006; Goubault et al., 2017), but other research opposes this idea (Özekmekçi, 2005). Looking at the results of this study, classification in the majority of classifiers, be it using CGP, SVM or ANN (for all three motor tasks), achieved at least 80% accuracy in identifying HD from healthy controls. Several high-performance classifiers even achieved 100% accuracies in classifying HD. Since the movement features used to induce the classifiers are Bradykinesia features, it is concluded that Bradykinesia does exists in HD. Excellent results were also achieved in classification of ET and healthy controls. However, because of the small number of ET patients it can't be

confirmed with confidence, but it is a strong indication that bradykinesia exists in ET, but will need further clarification with higher numbers of participants in the future.

iv) Strong discriminative movement features are present in classifications with high accuracies

For classifiers that showed good performance, with averaged accuracies above 80%, the best discriminative movement features were investigated. For every case considered, the five best classifiers were investigated, by rank-ordering the movement features most influential in finding the solution, and identifying these as discriminating features. In case of controls vs PD, two of the discriminating features are established characteristics in identifying PD (from previous objective assessment studies). This gives more confidence in the findings. Discriminative features for every case considered were summarised in Chapter 5**.** The ability of CGP to inform how the solution was achieved is very important and is especially useful in future clinical studies.

v) Classification using CGP is in all cases comparable with, and in certain cases better than, ANN and SVM

Accuracies of CGP classifiers in all cases were compared to classifiers induced by SVM and ANN. SVM and ANN results are used as a baseline to validate the results obtained using CGP. In all cases, accuracies of CGP classifiers are comparable to those obtained using SVM and ANN. In the case of controls vs PD, and controls vs ET, CGP classifiers even achieved higher accuracies than SVM and ANN**.**

vi) Record-wise cross-validation yielded better accuracies than subject-wise cross-validation in most of the cases

In record-wise validation, all the records were used as different samples and randomly divided into the training and test set datasets. Whilst in subject-wise validation, the records for the same subject will not be in the training and test set in the same iteration.

We were taking advantage of having real medical data to test the effect of using different style of cross-validation. The results showed that record-wise cross-validation yielded better accuracies than subject-wise cross-validation in most cases. This might be a sign of bias and overestimating as claimed in (Saeb et al., 2016); however further clarification is needed. All SVM and ANN classifiers induced in this study were achieved using record-wise cross-validation; a practice that is common. For example, leave one out cross-validation is a type of record-wise cross-validation.

vii) Organic and Functional Dystonia cannot be clearly differentiated using the same FT movement features in PD study

The movement features of FT have been used to differentiate between organic and functional Dystonia. The classification results were not very encouraging with averaged accuracies below 80%. Early statistically significance tests done on the at the beginning of the experiments shown promising results. However, as we include more participants, the performance decreased. There are few factors that might be the reasons of the lackof good performance by the classifiers. The first is, the classification between OD and FD - itself a very difficult clinical case. No study known has been successful in objectively differentiating between these two types of Dystonia. The second reason might be the motor task selected to evolve the classifiers is not suitable to measure the differences between the two groups. Finally, it may not be possible to differentiate between OD and FD by movement features alone as noe difference is actually expressed. Further work needs to be done as suggested in the next section (Section 7.2.2).

7.2.2 Future work

This study has two main limitations:

- The electromagnetic sensors were put at the same places during all motor tasks although the nature of movement for each task might be different.
- The small number of participants for the HD and ET studies.

In future work, new experiments should be considered for movement data collection. The electromagnetic sensors should be placed at suitable locations according to the task. For example, in case of hand pronation-supination, a sensor placed at wrist and another at thumb or finger might lead to better results. This way, the angels of movement can be measured. In case of Dystonia study, other motor tasks should be considered as well apart from finger tapping.

Although classifications of HD cases yielded very good accuracies, a higher number of participants will give more confidence in the accuracies and discriminating features achieved. As for ET cases, the classification in this study can be regarded as pilot study or preliminary experimentation to a future study with more participants. Another suggestion is to use a smaller set of features as CGP classifiers inputs. CGP is an algorithm than can predict solutions without any prior assumption and able to process any type and distribution of data. Also, this structure should be taken advantage of by using raw data points as inputs instead of using a set of features. This way, new movement features other than stated in literature might emerge as discriminating features.

7.3 Final conclusions

Having summarised, this study in Section 7.1 and highlighted the useful findings in Section 7.2, we can now revisit the hypothesis defined in section 1.6:

"Evolutionary Algorithms offer a means of differentiating and characterising a range of movement disorders using digitised kinematic data from common conventional clinical tasks."

In order to test the hypothesis, three study objectives were set:

- Apply EA predictive modelling to classify Parkinson's disease patients from healthy controls, essential tremor and Huntington's disease patients using available datasets.
- Observe characteristics of different movement disorders performing the same motor tasks by analysing the EA classifiers evolved.
- Use EAs to discriminate patients of functional (psychogenic) dystonia from healthy controls, and its organic counterpart, primary dystonia.

In Chapter 5, we concluded that CGP successfully differentiated PD from healthy controls, PD and ET. Although in order to generalise the results and inform clinical assessment, further clarifications are needed by increasing the number of participants. We also highlighted that one of the CGP advantages is the ability to investigate the solution. In this case, the study on the chromosomes of the best classifiers exposed the most discriminating features of that differentiated each group.

In Chapter 6, it was presented that the performance of Dystonia classification is not so encouraging. Nonetheless, it shows the positive trends where the almost all the

classifiers evolved achieved accuracies above 65%. With further work as suggested in Section 7.2, we believed better results can be obtained.

In light of experimental results presented in Chapter 5 and 6, a case can be made that Evolutionary Algorithms in form of Cartesian Genetic Programming offer a means of differentiating and characterising a range of movement disorders specifically Parkinson's disease, Huntington's disease and Essential Tremor using digitised kinematic data from common conventional clinical tasks of MDS-UPDRS, namely finger tapping, hand pronation-supination and hand opening-closing. It is hoped that this study can be used as the foundation to further efforts of using EAs in supporting clinical diagnosis of movement disorders. This study can be extended in many ways such as application to other types of movement disorders, or using kinematic data of other motor tasks, or using components from other assessment scales.

REFERENCES

Adler, C. H., Beach, T. G., Hentz, J. G., Shill, H. a, Caviness, J. N., Driver-Dunckley, E., Dugger, B. N. (2014). Low clinical diagnostic accuracy of early vs advanced Parkinson disease: clinicopathologic study. *Neurology*, *83*(5), 406–412.

Ahlrichs, C., & Lawo, M. (2013). Parkinson's Disease Motor Symptoms in Machine Learning: A Review. *Health Informatics - An International Journal*, *2*(4), 1–18.

Ahmad, A. M., Khan, G. M., & Mahmud, S. A. (2013). Classification of Arrhythmia Types Using Cartesian Genetic Programming Evolved Artificial Neural Networks, 282–291.

Ahmad, A. M., Khan, G. M., Mahmud, S. A., & Miller, J. F. (2012). Breast cancer detection using cartesian genetic programming evolved artificial neural networks. *Proceedings of the Fourteenth International Conference on Genetic and Evolutionary Computation Conference - GECCO '12*, 1031.

Albanese, A., Bhatia, K., Bressman, S. B., Delong, M. R., Fahn, S., Fung, V. S. C., Teller, J. K. (2013). Phenomenology and classification of dystonia: A consensus update. *Movement Disorders*, *28*(7), 863–873.

Albanese, A., & Jankovic, J. (2012). *Hyperkinetic Movement Disorders : Differetial diagnosis and treatment*. Oxford, UK: Blackwell Publishing.

Algarni, M., & Fasano, A. (2018). Parkinsonism and Related Disorders The overlap between Essential tremor and Parkinson disease, *46*, 101–104.

Allen, D. P., Playfer, J. R., Aly, N. M., Duffey, P., Heald, A., Smith, S. L., & Halliday, D. M. (2007). On the use of low-cost computer peripherals for the assessment of motor dysfunction in Parkinson's disease - Quantification of bradykinesia using target tracking tasks. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *15*(2), 286–294.

Arnold, G., Boone, K. B., Lu, P., Dean, A., Wen, J., Nitch, S., & McPherson, S. (2005). Sensitivity and specificity of finger tapping test scores for the detection of suspect effort. *Clinical Neuropsychologist*, *19*(1), 105–120.

Bajaj, N. P. S., Gontu, V., Birchall, J., Patterson, J., Grosset, D. G., & Lees, A. J. (2010). Accuracy of clinical diagnosis in tremulous parkinsonian patients: a blinded video study. *Journal of Neurology, Neurosurgery, and Psychiatry*, *81*(11), 1223–1228.

Banzhaf, W., Nordin, P., Keller, R. E., & Francone, F. D. (1998). Genetic programming: an introduction (Vol. 1). San Francisco: Morgan Kaufmann.

Bates, G., Harper, P. S., & Jones, L. (2002). *Huntington's Disease* (Third Edit). New York: Oxford University Press.

Benito-León, J. (2014). Essential tremor: a neurodegenerative disease? *Tremor and Other Hyperkinetic Movements (New York, N.Y.)*, *4*, 252.

Berkhout, J. (2015). Differences in brain connectivity between Essential tremor and Parkinson ' s disease : an EEG study.

Bouckaert, R. R., Frank, E., Hall, M. A., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2010). WEKA — Experiences with a Java Open-Source Project. *Journal OfMachine Learning Research*, *11*, 2533–2541.

Brabazon, A., O'Neill, M., & McGarraghy, S. (2015). Natural Computing Algorithms (Natural Computing Series).

Brameier, M. F., & Banzhaf, W. (2007). *Linear Genetic Programming*. *Igarss 2014*. New York,USA.

Carter, J. H., Stewart, B. J., Lyons, K. S., & Archbold, P. G. (2008). Do motor and nonmotor symptoms in PD patients predict caregiver strain and depression? *Movement Disorders*, *23*(9), 1211–1216.

Chaudhuri, K. R., & Ondo, W. G. (2009). *Handbook of Movement Disorders*. London: Current Medicine Group.

Christodoulakis, G., Busawon, K., Caplan, N., & Stewart, S. (2010). On the filtering and smoothing of biomechanical data. *Communication Systems Networks and Digital Signal Processing (CSNDSP), 2010 7th International Symposium On*, 512–516.

Clegg, J., Walker, J. A., & Miller, J. F. (2007, July). A new crossover technique for cartesian genetic programming. In Proceedings of the 9th annual conference on Genetic and evolutionary computation (pp. 1580-1587). ACM Press.

Coello, C., Lamont, G., & Veldhuisen, D. Van. (2007). Evolutionary algorithms for solving multi-objective problems.

Cristianini, N., & Shawe-Taylor, J. (2000). An Introduction to Support Vector Machines and other Kernel-Based Learning Methods. Cambridge: Cambridge Press.

Criswell, S., Sterling, C., Swisher, L., Evanoff, B., & Racette, B. A. (2010). Sensitivity and specificity of the finger tapping task for the detection of psychogenic movement disorders. *Parkinsonism and Related Disorders*, *16*(3), 197–201.

Daneault, J.-F., Carignan, B., Sadikot, A. F., & Duval, C. (2013). Are quantitative and clinical measures of bradykinesia related in advanced Parkinson's disease? *Journal of Neuroscience Methods*, *219*(2), 220–223.

Dayalu, P. (2015). Huntington Disease. *Neurologic Clinics of NA*, *33*(1), 101–114.

Deløcan, Y., Özyilmaz, L., & Yildirim, T. (2011). Evolutionary Algorithms Based RBF Neural Networks For Parkinson ' s Disease Diagnosis. *Electrical and Electronics Engineering (ELECO), 2011 7th International Conference On*, (1), 311–315.

Dorsey, E., Constantinescu, R., Thompson, J., Biglan, K., Holloway, R., & Kieburtz, K. (2007). Projected number of people with Parkinson disease in the most populous natios, 2005 through 2030. *Neurology*, *68*, 384–386.

Dowding, C. H., Shenton, C. L., & Salek, S. S. (2006). A review of the health-related quality of life and economic impact of Parkinson's disease. *Drugs & Aging*, *23*(9), 693–721.

Duval, C., Fenney, A., & Jog, M. S. (n.d.). The Dynamic Relationship Between Voluntary and Involuntary Motor Behaviours in Patients with Basal Ganglia Disorders, 521–534.

Duval, C., Sadikot, A. F., & Panisset, M. (2006). Bradykinesia in patients with essential tremor. *Brain Research*, *1115*(1), 213–216.

Egerton, T., Williams, D. R., & Iansek, R. (2012). Comparison of gait in progressive supranuclear palsy, Parkinson's disease and healthy older adults. *BMC Neurology*, *12*(1), 116.

Eiben, A. E., & Smith, J. E. (2003). Introduction to Evolutionary Computing. New York.

Ellson, J., Gansner, E., Koutsofios, L., North, S., & Woodhull, G. (2000). Graphviz - Open Source Graph Drawing Tools, 3–4.

EPDA. (2012). THE EUROPEAN PARKINSON'S DISEASE STANDARDS OF CARE CONSENSUS STATEMENT.

Espay, A. J., & Lang, A. E. (2015). Phenotype-Specific Diagnosis of Functional (Psychogenic) Movement Phenotype-Specific Diagnosis of Functional (Psychogenic) Movement Disorders, (April).

Espejo, P. G., Ventura, S., & Herrera, F. (2010). A Survey on the Application of Genetic Programming to Classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *40*(2), 121–144.

Evans, S. J. W., Douglas, I., Rawlins, M. D., Wexler, N. S., Tabrizi, S. J., & Smeeth, L. (2013). Prevalence of adult Huntington ' s disease in the UK based on diagnoses recorded in general practice records, 1156–1160.

Fahn, S. (2011). Classification of movement disorders. *Movement Disorders*, *26*(6), 947–957.

Fahn, S., Jancovic, J., & Hallett, M. (2011). *Principles and Practice of Movement Disorders* (2nd ed.). Philadelphia: Elsevier.

Fahn, S., Jankovic, J., Hallett, M., & Jenner, P. (2007). *Principles and practice of movement disorders*. *Motor Control*.

Fenney, A., Jog, M. S., & Duval, C. (2008). Bradykinesia is not a "systematic" feature of adult-onset Huntington's disease; implications for basal ganglia pathophysiology. *Brain Research*, *1193*, 67–75.

Fernandez, H. H., Machado, A. G., & Pandya, M. (2014). *A Practical Approach to Movement Disorders: Diagnosis and Management.* Demos Medical Publishing.

Fernandez, H. H., Machado, A. G., & Pandya, M. (2015). A paractical approach to movement disorders: Diagnosis and Mnagement. Demos Medical (Second). New York.

Forman, G., & Cohen, I. (2004). Learning from Little: Comparison of Classifiers Given Little Training. *Knowledge Discovery in Databases: PKDD 2004*, *19*(September), 161–172.

Freitas, A. A. (2010). A Review of Evolutionary Algorithms for Data Mining. In O. Maimon & L. Rokach (Eds.), *Data Mining and Knowledge Discovery Handbook* (2nd ed.). Springer US.

Gandomi, A. H., Alavi, A. H., & Ryan, C. (2015). *Handbook of Genetic Programming Applications*. Springer.

Ganos, C., Edwards, M. J., & Bhatia, K. P. (2014). The Phenomenology of Functional (Psychogenic) Dystonia Phenotypic Characteristics of FD, (January), 36–44.

Garcia Ruiz, P. J., Gomez Tortosa, E., Sanchez Bernados, V., Rojo, A., Fontán, A., & Garcia de Yebenes, J. (2000). Bradykinesia in Huntington's disease. *Clinical Neuropharmacology*, *23*(1), 50–52.

Geraghty, J. J., Jankovic, J., & Zetusky, W. J. (1985). Association between essential tremor and Parkinson's disease. *Annals of Neurology*, *17*, 329–333.

Geyer, H. L., & Bressman, S. B. (2006). The diagnosis of dystonia, *5*(September).

Gibb, W. R., & Lees, a J. (1988). The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *51*(6), 745–752.

Goetz, C. G., Chmura, T. A., & Lanska, D. J. (2001). History of movement disorders as a neurological specialty: Part 14 of the MDS-sponsored history of movement disorders exhibit, Barcelona, June 2000. *Movement Disorders*, *16*(5), 954–959.

Goetz, C. G., & Stebbins, G. T. (2004). Assuring Interrater Reliability for the UPDRS Motor Section : Utility of the UPDRS Teaching Tape †, *19*(12), 10–13.

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., Zweig, R. M. (2008). Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disorders*, *23*(15), 2129–2170.

Goubault, E., Nguyen, H. P., Ayachi, F. S., Bogard, S., & Duval, C. (2017). Do Bradykinesia and Tremor Interfere in Voluntary Movement of Essential Tremor Patients ? Preliminary Findings. *Tremor and Other Hyperkinetic Movements*, 1–7.

Gregory, T. R. (2009). Understanding Natural Selection: Essential Concepts and Common Misconceptions. *Evolution: Education and Outreach*, *2*(2), 156–175.

Hazell, A., & Smith, S. L. (2008). Towards an Objective Assessment of Alzheimer ' s Disease : The Application of a Novel Evolutionary Algorithm in the Analysis of Figure Copying Tasks, 2073–2079.

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, *21*(9), 1263–1284.

He, H., & Ma, Y. (2013). Imbalanced Learning: Foundations, Algorithm, and Applications. New Jersey: John Wiley & Sons.

Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. IEEE Intelligent Systems and their applications, 13(4), 18-28.

Heldman, D. a., Espay, A. J., LeWitt, P. a., & Giuffrida, J. P. (2014). Clinician versus machine: Reliability and responsiveness of motor endpoints in Parkinson's disease. *Parkinsonism and Related Disorders*, *20*(6), 590–595.

Heldman, D. A., Giuffrida, J. P., Chen, R., Payne, M., Mazzella, F., Duker, A. P., … Espay, A. J. (2011). The modified bradykinesia rating scale for Parkinson's disease: Reliability and comparison with kinematic measures. *Movement Disorders*, *26*(10), 1859–1863.

Hilten, J. J. Van, Zwan, A. D. Van Der, Zwinderman, A. H., & Roos, R. A. C. (1994). Rating Impairment and Disability in Parkinson ' s Disease : Evaluation of the Unified Parkinson ' s Disease Rating Scale, *9*(1), 84–88.

Hons, S. L. B., Frcgp, A. C., Hague, J., Bs, M. B., Parsonage, M., & Hons, B. A. (2010). The cost of somatisation among the working-age population in England for the year 2008 – 2009, 71–84.

Hope, D. C., Munday, E., & Smith, S. L. (2007). Evolutionary Algorithms in the Classification of Mammograms. *2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing*, (CIISP), 258–265.

Hornykiewicz, O. (2006). The discovery of dopamine deficiency in the parkinsonian brain. *Journal of Neural Transmission. Supplementum*, (70), 9–15.

Hughes, A. J., Daniel, S. E., Ben-Shlomo, Y., & Lees, A. J. (2002). The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. *Brain : A Journal of Neurology*, *125*(Pt 4), 861–870.

Hughes, A. J., Daniel, S. E., Kilford, L., & Lees, A. J. (1992). Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *Journal of Neurology, Neurosurgery & Psychiatry*, *55*(3), 181–184.

Hughes, A. J., Daniel, S. E., & Lees, A. J. (2001). Improved accuracy of clinical diagnosis of Lewy body Parkinson's disease. *Neurology*, *57*, 1497–1499.

Jain, S., Lo, S. E., & Louis, E. D. (2006). Common Misdiagnosis of a Common Neurological Disorder. *Archives of Neurology*, *63*(8), 1100.

Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry*, *79*(4), 368–376.

Jiménez-Jiménez, F. J., Alonso-Navarro, H., García-Martín, E., & Agúndez, J. A. G. (2012). The relationship between Parkinson's disease and essential tremor: review of clinical, epidemiologic, genetic, neuroimaging and neuropathological data, and data on the presence of cardinal signs of parkinsonism in essential tremor. *Tremor and Other Hyperkinetic Movements (New York, N.Y.)*, *2*(5), 799–804.

Jiménez-Jiménez, F. J., Rubio, L., Alonso-Navarro, H., Calleja, M., Pilo-De-La-Fuente, B., Plaza-Nieto, J. F., … Agúndez, J. A. G. (2010). Impairment of rapid repetitive finger movements and visual reaction time in patients with essential tremor. *European Journal of Neurology*, *17*(1), 152–159.

Jobbágy, A., Harcos, P., Karoly, R., & Fazekas, G. (2005). Analysis of finger-tapping movement. *Journal of Neuroscience Methods*, *141*(1), 29–39.

Jobbágy, Á., Harcos, P., Karoly, R., & Fazekas, G. (2005). Analysis of finger-tapping movement, *141*, 29–39.

Joutsa, J., Gardberg, M., Röyttä, M., & Kaasinen, V. (2014). Diagnostic accuracy of parkinsonism syndromes by general neurologists. *Parkinsonism & Related Disorders*, *20*(8), 840–844.

Kalogjera-Sackellares, D., & Sackellares, J. C. (1999). Intellectual and neuropsychological features of patients with psychogenic pseudoseizures. *Psychiatry Research*, *86*(1), 73–84.

Keränen, T., Kaakkola, S., Sotaniemi, K., Laulumaa, V., Haapaniemi, T., Jolma, T., … Takala, A. (2003). Economic burden and quality of life impairment increase with severity of PD. *Parkinsonism and Related Disorders*, *9*(3), 163–168.

Klein, C. (2005). Movement disorders: Classifications. *J Inherit. Metab Dis*, *28*(3), 425–439.

Kowal, S. L., Dall, T. M., Chakrabarti, R., Storm, M. V., & Jain, A. (2013). The current and projected economic burden of Parkinson's disease in the United States. *Movement Disorders*, *28*(3), 311–318.

Koza, Jo. R. (1998). Genetic Programming: On the Programming of Computers by Means of Natural Selection (Sixth prin). London: The MIT Press.

Künig, G., & Alba, A. B. (2011). Bradykinesia in early Huntington ' s disease, 277–278.
Kwon, K. Y., Lee, H. M., Lee, S. M., Kang, S. H., & Koh, S. B. (2016). Comparison of motor and non-motor features between essential tremor and tremor dominant Parkinson's disease. *Journal of the Neurological Sciences*, *361*, 34–38.

Lacy, S. E. (2015). Modeling Movement Disorders in Parkinson ' s Disease using Computational Intelligence.

Lacy, S. E., Lones, M. a, & Smith, S. L. (2013). Characterisation of Movement Disorder in Parkinson ' s Disease using Evolutionary Algorithms Categories and Subject Descriptors, (August 2009), 1479–1485.

Langdon, W., & Poli, R. (2002). *Foundations of Genetic Programming*. New York: Springer-Verlag Berlin Heidelberg.

Little, M. A., Varoquaux, G., Saeb, S., Lonini, L., Jayaraman, A., Mohr, D. C., … Triangle, A. (2017). Using and understanding cross-validation strategies . Perspectives on Saeb et al ., (March), 1–6.

Lones, M. A., Alty, J. E., Duggan-Carter, P., Turner, A. J., Jamieson, D. R. S., & Smith, S. L. (2014). Classification and Characterisation of Movement Patterns During Levodopa Therapy for Parkinson's Disease. *Proceedings of the 2014 Conference on Genetic and Evolutionary Computation - GECCO '14*, 1321–1327.

Lones, M. A., Alty, J. E., Lacy, S. E., Jamieson, D. R. S., Possin, K. L., Schuff, N., & Smith, S. L. (2013). Evolving Classifiers to Inform Clinical Assessment of Parkinson ' s Disease, 76–82.

Lones, M. A., Member, S., Smith, S. L., Alty, J. E., Lacy, S. E., Member, G. S., … Tyrrell, A. M. (2014). Evolving Classifiers to Recognize the Movement Characteristics of Parkinson ' s Disease Patients, *18*(4), 559–576.

Lones, M. A., Smith, S. L., Alty, J. E., Lacy, S. E., Possin, K. L., Jamieson, D. R. S., & Tyrrell, A. M. (2014). Evolving classifiers to recognize the movement characteristics of parkinson's disease patients. *IEEE Transactions on Evolutionary Computation*, *18*(4), 559–576.

Lones, M. a., Smith, S. L., Harris, A. T., High, A. S., Fisher, S. E., Smith, D. A., & Kirkham, J. (2010). Discriminating normal and cancerous thyroid cell lines using implicit context representation Cartesian genetic programming. *IEEE Congress on Evolutionary Computation*, 1–6.

Louis, E. D., Benito-Leon, J., & Faust, P. L. (2016). Essential tremor is a risk factor for Parkinson's disease. *Parkinsonism and Related Disorders*, *24*, 143–144.

Louis, E. D., & Ottman, R. (2014). How many people in the USA have essential tremor? Deriving a population estimate based on epidemiological data. *Tremor and Other Hyperkinetic Movements (New York, N.Y.)*, *4*, 259.

Ma, Y., & Guo, G. (Eds.). (2014). Support Vector Machines Applications. Springer Science & Business Media.

Maetzler, W., Liepelt, I., & Berg, D. (2009). Progression of Parkinson's disease in the clinical phase: potential markers. *The Lancet Neurology*, *8*(12), 1158–1171.

Mann, R. K., Edwards, R., Zhou, J., Fenney, A., Jog, M., & Duval, C. (2012). Comparing movement patterns associated with Huntington's chorea and Parkinson's dyskinesia. *Experimental Brain Research*, *218*(4), 639–654.

Marsland, S. (2015). *MACHINE LEARNING An Algorithmic Perspective* (Second Edi). New York: CRC Press.

Martinez-Manzanera, O., Roosma, E., Beudel, M., Borgemeester, R. W. K., Laar, T. Van, & Maurits, N. M. (2015). A method for automatic , objective and continuous scoring of bradykinesia. *IEEE International Conference on Body Sensor Networks (BSN)*.

Martínez-Martín, P., Benito-León, J., Alonso, F., Catalán, M. J., Pondal, M., Zamarbide, I., de Pedro, J. (2005). Quality of life of caregivers in Parkinson's disease. *Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *14*(2), 463–472.

Martínez-Martín, P., Forjaz, M. J., Frades-Payo, B., Rusiñol, A. B., Fernández-García, J. M., Benito-León, J., … Catalán, M. J. (2007). Caregiver burden in Parkinson's disease. *Movement Disorders*, *22*(7), 924–931.

Martinez-Martin, P., Gil-Nagel, A., Morlan Gracia, L., Balseiro Gomez, J., Martinez-Sarries, J., Bermejo, F., & Group, T. C. M. (1994). Unified Parkinson ' s Disease Rating Scale Characteristics and Structure. *Movement Disorders*, *9*(1), 76–83.

Martinez-Martin, P., Jeukens-Visser, M., Lyons, K. E., Rodriguez-Blazquez, C., Selai, C., Siderowf, A., Schrag, A. (2011). Health-related quality-of-life scales in Parkinson's disease: Critique and recommendations. *Movement Disorders*, *26*(13), 2371–2380.

Martinez Manzanera, O., Roosma, E., Beudel, M., Borgemeester, R., van Laar, T., & Maurits, N. (2015). A method for automatic and objective scoring of bradykinesia using orientation sensors and classification algorithms. *IEEE Transactions on Biomedical Engineering*, *9294*(c), 1–1.

Martínez Pueyo, A., García-Ruiz, P. J., Feliz, C. E., Garcia Caldentey, J., Del Val, J., & Herranz, A. (2016). Reaction time and rhythm of movement in Huntington's disease. *Journal of the Neurological Sciences*, *362*, 115–117.

Miller, J. F. (2011). *Cartesian Genetic Programming*. Springer-Verlag Berlin Heidelberg.
Miller, J. F., & Thomson, P. (2000). Cartesian genetic programming. Genetic Programming, Proceedings of the Third European Conference on Genetic Programming (EuroGP2000)., 1802, 121–132.

Montgomery, E. B., Baker, K. B., Lyons, K., & Koller, W. C. (2000). Motor initiation and execution in essential tremor and Parkinson's disease. *Movement Disorders*, *15*(3), 511–515.

Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease. (2003). State of the Art Review The Unified Parkinson ' s Disease Rating Scale (UPDRS): Status and Recommendations. *Movement Disorders*, *18*(7), 738–750.

Müller, T., Saft, C., Andrich, J., & Harati, A. (2013). Diadochokinetic movements differ between patients with Huntington's disease and controls. *NeuroRehabilitation*, *33*(4), 649–655.

Naghavi, M., Wang, H., Lozano, R., Davis, A., Liang, X., Zhou, M., … Temesgen, A. M. (2015). Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, *385*(9963), 117–171.

Napierala, K., & Stefanowski, J. (2016). Types of minority class examples and their influence on learning classifiers from imbalanced data. *Journal of Intelligent Information Systems*, *46*(3), 563–597.

National Collaborating Centre for Chronic Conditions. (2006). Parkinson's disease: national clinical guideline for diagnosis and management in primary and secondary care. *Royal College of Physicians*.

Newby, R. E., Thorpe, D. E., Kempster, P. A., & Alty, J. E. (2017). A History of Dystonia: Ancient to Modern. *Movement Disorders Clinical Practice*, *4*(4), 478–485.

Okada, M., & Okada, M. (1983). A method for quantification of alternate pronation and supination of forearms. *Computers and Biomedical Research*, *16*(1), 59–78.

Özekmekçi, S. (2005). Assessment of movement time in patients with essential tremor, 964–967.

Pal, P. K. (2011). Electrophysiologic Evaluation of Psychogenic Movement Disorders, 21–32.

Paliouras, G., Karkaletsis, V., & Spyropoulos, C. D. (Eds.). (2003). Machine learning and its applications: advanced lectures (Vol. 2049). Springer.

Patel, S., Lorincz, K., Hughes, R., Huggins, N., Growdon, J., Standaert, D., … Bonato, P. (2009). Monitoring Motor Fluctuations in Patients With Parkinson's Disease Using Wearable Sensors. *IEEE Transactions on Information Technology in Biomedicine*, *13*(6), 864–873.

Pereira, C. R., Pereira, D. R., Silva, F. A. D., Hook, C., Weber, S. A. T., Pereira, L. A. M., & Papa, J. P. (2015). A step towards the automated diagnosis of parkinson's disease: Analyzing handwriting movements. *Proceedings - IEEE Symposium on Computer-Based Medical Systems*, *2015–July*, 171–176.

Perlmutter, J. S. (2010). NIH Public Access, 1–16.

Petrie, A., & Sabin, C. (2000). *Medical Statistics at a Glance*. London: Blackwell Science.

Picardi, C., Cosgrove, J., Smith, S. L., Jamieson, D. R. S., & Alty, J. E. (2010). *Objective Assessment of Cognitive Impairment in Parkinson's Disease Using Evolutionary Algorithm*. (C. Di Chio, A. Brabazon, G. A. Di Caro, M. Ebner, M. Farooq, A. Fink, N. Urquhart, Eds.), *Florida dental journal* (Vol. 6025). Berlin, Heidelberg: Springer Berlin Heidelberg.

Poli, R. (1996). Parallel distributed genetic programming. Technical Report CSRP-96-15, School of Computer Science, The University of Birmingham.

Poli, R., Langdon, W. B., & McPhee, N. F. (2008). *A Field Guide to Genetic Programing*. Published via http://lulu.com and freely available at http://www.gp-field-guide.org.uk.

Popescu, A., Lippa, C., Lee, V. M.-Y., & Trojanowski, J. Q. (2004). Lewy Bodies in the Amygdala. *Arch Neurol.*, *61*, 1915–1919.

Quinn, N. P. (2005). How to diagnose multiple system atrophy. *Movement Disorders : Official Journal of the Movement Disorder Society*, *20 Suppl 1*(Box 13), S5–S10.

Rajput, A. H., & Rajput, A. (2014). Accuracy of Parkinson disease diagnosis unchanged in 2 decades. *Neurology*, *83*(5), 386–387.

Rajput, A. H., Rozdilsky, B., & Rajput, A. (1991). Accuracy of clinical diagnosis in parkinsonism--a prospective study. *Canadian Journal of Neurological Sciences.*, *18*(3), 275–278.

Rapport, L. J., Farchione, T. J., Coleman, R. D., & Axelrod, B. N. (1998). Effects of coaching on malingered motor function profiles. *Journal of Clinical Experimental Neuropsychology Official Journal of the International Neuropsychological Society*, *20*(1), 89–97.

Respondek, G., Roeber, S., Kretzschmar, H., Troakes, C., Al-Sarraj, S., Gelpi, E., … Höglinger, G. U. (2013). Accuracy of the national institute for neurological disorders and stroke/society for progressive supranuclear palsy and neuroprotection and natural history in Parkinson plus syndromes criteria for the diagnosis of progressive supranuclear palsy. *Movement Disorders*, *28*(4), 504–509.

Rovini, E., Maremmani, C., & Cavallo, F. (2017). How wearable sensors can support parkinson's disease diagnosis and treatment: A systematic review. *Frontiers in Neuroscience*, *11*(OCT).

Saeb, S., Lonini, L., Jayaraman, A., Mohr, D. C., & Kording, K. P. (2016). Voodoo Machine Learning for Clinical Predictions, 1–11.

Saeb, S., Lonini, L., Jayaraman, A., Mohr, D. C., & Kording, K. P. (2017). The need to approximate the use-case in clinical machine learning, (October 2016), 1–9.

Serrano, J. I., & del Castillo, M. D. (2012). On the origin of the evolutionary computation species influences of Darwin's theories on computer science. *Artificial Intelligence Review*, *38*(1), 41–54.

Shahed, J., & Jankovic, J. (2016). Exploring the relationship between essential tremor and Parkinson's disease Joohi. *Parkinsonism and Related Disorders*, *22*, S162–S165.

Sharma, N. (2008). *Parkinson's Disease (Biographies of Disease)*. Wesport, USA: Greenwood.

Sivanandam, S. N., & Deepa, S. N. (2008). *Introduction to Genetic Algorithms*. Springer Berlin Heidelberg.

Smith, S. (2011). Cartesian Genetic Programming and its Application to Medical Diagnosis. *IEEE Computational Intelligence Magazine*, *6*(4), 56–67.

Smith, S. L., & Cagnoni, S. (2011). Genetic and evolutionary computation: medical applications. John Wiley & Sons.

Smith, S. L., & Timmis, J. (2008). An immune network inspired evolutionary algorithm for the diagnosis of Parkinson's disease. *Bio Systems*, *94*(1–2), 34–46.

Soh, S.-E., McGinley, J. L., Watts, J. J., Iansek, R., Murphy, A. T., Menz, H. B., … Morris, M. E. (2013). Determinants of health-related quality of life in people with Parkinson's disease: a path analysis. *Quality of Life Research*, *22*(7), 1543–1553.

Spadoto, A., Guido, R. C., Carnevali, F. L., Pagnin, F., Falc, A. X., & Papa, P. (2011). Improving Parkinson ' s Disease Identification Through Evolutionary-Based Feature Selection, 7857–7860.

Spasojević, S., Ilić, T. V., Stojković, I., Potkonjak, V., Rodić, A., & Santos-Victor, J. (2017). Quantitative assessment of the arm/hand movements in Parkinson's disease using a wireless armband device. *Frontiers in Neurology*, *8*(AUG).

Stamatakis, J., Ambroise, J., Crémers, J., Sharei, H., Delvaux, V., Macq, B., & Garraux, G. (2013). Finger tapping clinimetric score prediction in Parkinson's disease using low-cost accelerometers. *Computational Intelligence and Neuroscience*, *2013*, 717853.

Stamelou, M., Edwards, M. J., Hallett, M., & Bhatia, K. P. (2012). The non-motor syndrome of primary dystonia: clinical and pathophysiological implications. *BRAIN A JOURNAL OF NEUROLOGY*, (2011), 1668–1681.

Statnikov, A., Aliferis, C. F., Hardin, D. P., & Guyon, I. (2011). A Gentle Introduction to Support Vector Machines in Biomedicine: Volume 1: Theory and Methods. World Scientific Publishing Company.

Sulzer, D., Chaudhuri, K. R., & Fahn, S. (2015). Introducing a new journal on Parkinson's disease. *Npj Parkinson's Disease*, *1*, 15006.

Thompson, P. D., al., et, & Marsden, C. D. (1988). The coexistence of brady kinesia and chorea in Huntington's disease and its implications for theories of basal ganglia control of movement, *111*, 223–244.

Turner, A. J., & Miller, J. F. (2014). Introducing a cross platform open source Cartesian Genetic Programming library. *Genetic Programming and Evolvable Machines*.

Van Vugt, J. P. P., Piet, K. K. E., Vink, L. J., Siesling, S., Zwinderman, A. H., Middelkoop, H. A. M., & Roos, R. A. C. (2004). Objective assessment of motor slowness in Huntington's disease: Clinical correlates and 2-year follow-up. *Movement Disorders*, *19*(3), 285–297.

Weiss, G. M., & Provost, F. J. (2003). Learning When Training Data are Costly: The Effect of ClassDistribution on Tree Induction. *J. Artif. Intell. Res. (JAIR)*, *19*, 315–354.

Wenning, G. K., Kiechl, S., Seppi, K., Müller, J., Högl, B., Saletu, M., … Poewe, W. (2000). Prevalence of movement disorders in men and women aged 50 – 89 years (Bruneck Study cohort): a population-based, 815–820.

Winter, D. A. (2009). Biomechanics and motor control of human movement, Fourth Edition. John Wiley & Sons.

World Health Organization. (2006). *Neurological Disorders: Public Health Challenges*. Geneva: World Health Organization.

Yokoe, M., Okuno, R., Hamasaki, T., Kurachi, Y., Akazawa, K., & Sakoda, S. (2009a). Opening velocity, a novel parameter, for finger tapping test in patients with Parkinson's disease. *Parkinsonism & Related Disorders*, *15*(6), 440–444.

Yokoe, M., Okuno, R., Hamasaki, T., Kurachi, Y., Akazawa, K., & Sakoda, S. (2009b). Opening velocity, a novel parameter, for finger tapping test in patients with Parkinson's disease. *Parkinsonism and Related Disorders*, *15*(6), 440–444.

Yu, X., & Gen, M. (2010). *Introduction to Evolutionary Algorithms*. London: Springer-Verlag London.