TALLINN UNIVERSITY OF TECHNOLOGY
Faculty of Information Technology
Department of Computer Science

Modelling Parkinson's Disease with Gait Analysis Approach

Master's thesis

Anna Krajuškina
176838IAPM

Supervisors
Sven Nõmm, PhD
Aaro Toomela, PhD, MD
Kadri Medijainen, MSc

Tallinn 2019

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia Teaduskond
Tarkvarateaduse instituut

Parkinsoni Tõve Modellerimine Kõnnaku Analüüsi Meetodil

Magistritöö

Anna Krajuškina
176838IAPM

Juhendajad
Sven Nõmm, PhD
Aaro Toomela, PhD, MD
Kadri Medijainen, MSc

Tallinn 2019

Declaration

I declare that this thesis is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Anna Krajuškina

May 7, 2019

.....

(Signature)

Abstract

In spite of the recent advances in the areas of motion capture and motion analysis, medical community remains sceptical about applying computer aided systems to support modelling and diagnostics of Parkinson's disease. The main goal of this work is to analyse the dataset collected during previous studies and propose an extended set of parameters that will have a higher discriminative power in Parkinson's disease classification than conventional methods. The dataset initially collected for this work [1] includes recordings of a Up-and-Go gait mobility tests in a form of Comma Separated Values (CSV) files. The data was collected using the Kinect sensor that is able to track three dimensional coordinates of joints of a human skeleton.

In this thesis, motion mass based approach is adopted to describe gait movements of the patients. In a clinical setting, the set of measurable parameters is limited by the time, step lengths and in some cases angles between the limbs. This study compliments this set of parameters with motion mass parameters. These parameters describe amount and smoothness of the motion and are computed for each step. The feature selection process turns a lot of attention to interpretability of the features and their combination. Then classifiers are trained, treating the values of motion mass parameters as features. The baseline to which the classifiers are compared is a classifier that has duration of the test as an only feature - the only measurable parameter commonly used with Up-and-Go test.

In particular, in scope of this thesis is analysis of the existing dataset, implementation of step extraction and calculation of parameters. After the feature extraction is performed, another step is feature selection with stress on medical interpretability of the parameters used. The implementation also includes the model training and reporting layer.

It is demonstrated that the time durations related to the tests does not possess highest discrimination power. Using Motion Mass parameters it was possible to outperform the time-only classifier. In addition, the study showed the importance of the upper-body

features on diagnosing Parkinson's disease.

The thesis is in English and contains 65 pages of text, 6 chapters, 13 figures, 9 tables.

Annotatsioon

Liikumise jäädvustamise ja analüüsi valdkond on hiljuti näinud mitmeid edusamme. Sellele vaatamata on meditsiiniline kogukond jäänud skeptiliseks arvutipõhiste süsteemide rakendamisel Parkinsoni tõve modelleerimise ja diagnostika toetamiseks. Antud töö põhieesmärk on analüüsida eelmiste uuringute käigus kogutud andmed ja pakkuda välja laiendatud parameetrite kogum, millel on kõrgem diskrimineeriv võimsus Parkinsoni tõve klassifitseerimisel kui tavapärastel meetoditel. [1] käigus jäädvustatud andmekogu sisaldab Up-and-Go kõnnaku liikuvuse testide salvestusi CSV failides. Andmed jäädvustati Kinecti sensorite abil, mis on võimeline jälgima inimese skeleti liigeste kolmemõõtmelisi koordinaate.

Patsientide kõnnak kirjeldatakse kasutades liikumismassi parameetreid (ing. motion mass parameters), mis kirjeldavad liikumise suurust ja sujuvust. Kliinilises keskkonnas mõõdetakse testile kuluv aeg, sammude pikkust ja mõnel juhul ka jäsemetevaheliste nurkade suuruseid. Antud uurimistöö käigus täiendati eelnevate tööde andmeid lisades iga sammu kohta liikumismassi parameetrid.

Parimate parameetrite ja nende kombinatsioonide valimisel pöörati tähelepanu sellele, et need oleksid ka meditsiiniliselt tõlgendatavad. Seejärel ehitatakse masinõppega klassifikaatorid, kasutades sisendina liikumismassi parameetrite väärtusi. Saadud klassifikaatorid võrreldakse baasklassifikaatoriga, mille ainsaks sisendparameetriks oli katse kestus - seda kasutatakse tavaliselt Up-and-Go testiga.

Käesolevas töö raames teostati olemasoleva andmestiku analüüs, salvestusest sammude eraldamine ja parameetrite arvutamine. Parameetrite arvutamisele järgnes parimate parameetrite valik, rõhutades kasutatud parameetrite meditsiinilist tõlgendatavust. Töö hõlmab ka mudelite väljaõpet ja aruandluse genereereerimist.

Tulemused näitavad, et Up-and-Go katse ajaline kestus ei ole kõrgeima diskrimineerimisvõimega. Liikumismassi parameetrite kasutamisel oli võimalik ületada ainult kestusel põhineva klassifikaatori tootlust. Lisaks näitas uurimus ülakeha omaduste täht-

sust Parkinsoni tõve diagnoosimisel.

Lõputöö on kirjutatud Inglise keeles ning sisaldab teksti 65 leheküljel, 6 peatükki, 13 joonist, 9 tabelit.

Contents

List of Figures

List of Tables

List of Acronyms

PD Parkinson's Disease

TUG Timed Up-and-Go test

SDK Software Development Kit

MM Motion Mass parameters

C Controls

MRI Magnetic Resonance Imaging

ML Machine learning

IMU Inertial Measurement Unit

DT Decision Tree

LogR Logistic Regression

kNN k-nearest neighbours

SVM Support Vector Machines

HTML Hypertext Markup Language

CSV Comma Separated Values

1. Introduction

1-2 % of the population older that 65 years suffer from Parkinson's Disease (PD) [2]. It is a neurodegenerative disease, which affects human motor functions, resulting in unintentional movements and joints rigidity [3]. Characteristic symptoms include a consistent decline in the gait stability, increasing tremor and muscle rigidity [2]. There are also some non-motor symptoms that often occur such as sleep abnormalities, depression and others [2, 4, 5]. The disease is thought to be caused by dopamine deficiency due to death of dopaminergic neurons [6].

Those symptoms vary from person to person depending on lifestyle and severity of the PD. The wide variety of combinations and the similarity of Parkinsonian diseases [4] make the diagnosis of this disease extremely challenging, especially on the early stages, and requires a skilled and an experienced therapist. Although imaging technologies such as Magnetic Resonance Imaging (MRI) cannot capture dopamine deficiency [2], MRI scans are often used to exclude other possible diseases with similar symptoms. There are a number of tests used in medicine to support the medical staff, but without the right supporting tools, the results are biased and depend heavily on the level of the expertise of the therapist.

Despite recent advances in understanding of Parkinson's disease, underlying mechanisms of neurodegeneration in PD are unknown [7]. At the present moment, there is no known cure against PD [7]. However, early diagnosis and monitoring of PD is crucial to prolong the normal way of life of the patients. The dopamine replacement therapy is able to control the symptoms for a number of years and is associated with near-normal life expectancy [2].

The current thesis is organized as follows. Introduction gives an overview of research done in area of classification of PD and describes the motivation for this work stating the main research questions and baselines. It describes experimental setting and presents

formal problem statements. Chapter 2 gives an implementation details of the approach. It brings the description of the dataset and goes over the main parts of the classification flow describing reasoning behind each decision. Analysis and the interpretation of the outcomes of the second part are presented in Section 3. Some immediate conclusions are drawn in this section. Section 4 discusses the outcomes of the thesis, problems and possibilities for practical applications. It also describes ideas for future work. Last chapter sums up the work and points out main deductions.

1.1 Background

PD severely affects human motion causing body tremor, muscle rigidity and gait instability. This result in variability in movement performance efficiency, amount and smoothness of the movements. These abnormalities can be detected using gait analysis.

Gait analysis is a study of human motion focusing on measuring body movement and muscle activities. Usually some parameters are measured in order to quantify the results and better detect motor abnormalities or wrong technique in case it is used for sport evaluation.

There are numerous gait mobility tests that can be used to asses the balance of older or injured people and detect gait abnormalities that can occur due to different physical disorders. They can provide a useful insight into the state of an individual and help plan the necessary rehabilitation or choose the right strategy for care. Some of the examples are Sit-to-Stand test, Go test, 4-Meter-Walk test, Up-and-Go test [8].

Up-and-Go test is one of the gait mobility tests. It starts with a person sitting on a chair, then, after the command is given by a therapist, the person stands up from a chair, walks three meters, turns back, walks back to the chair, turns back around and sits back to the chair (Figure 1.1). The timed version of the test is also referred to as Timed Up-and-Go test.

[9] showed that it is possible to distinguish between elderly people without PD and with PD by the results of Timed Up-and-Go test (TUG) test. It was also shown in this study that the measurements reflected the change in performance for patients before and after levodopa usage (a drug used to treat symptoms of PD). That makes TUG a suitable tool for tracking the progression of the illness and estimate the effectiveness of therapy. [10] also showed that the results differ for healthy individuals and patients. Time of per-

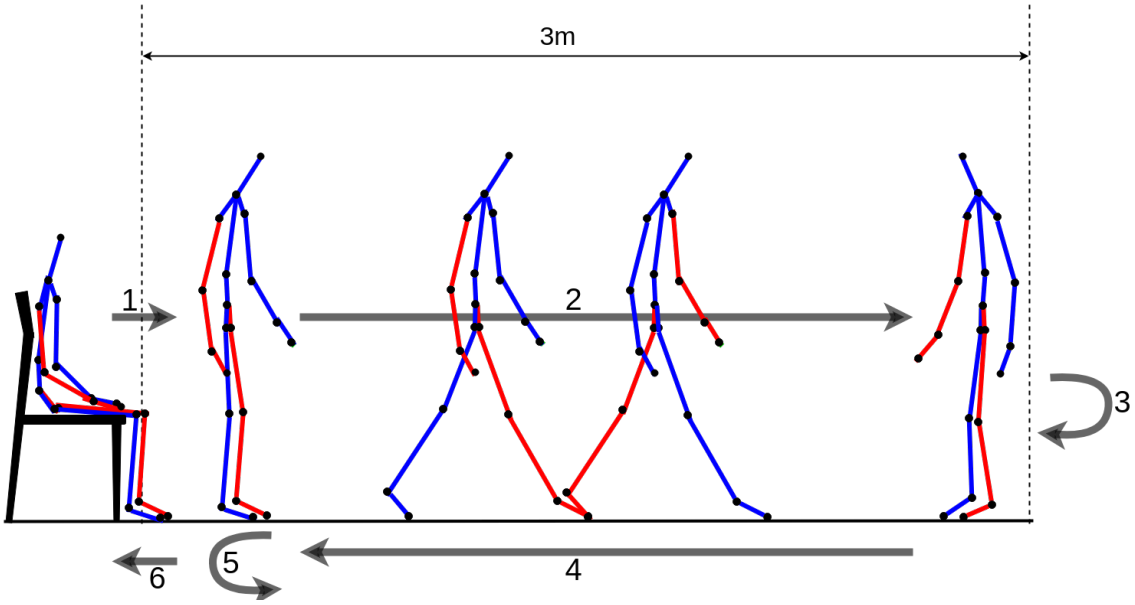


Figure 1.1: TUG test consists of six phases. Firstly, a person stands up from a chair (phase 1). Then, they walk three meters (phase 2), turn around (phase 3), walk back to the chair (phase 4). When reaching the chair, the person turns back around (phase 5) and sits back to the chair (phase 6).

forming the test was proved to be a reliable measure for determining H & Y score (the severity level of the disease) [10]. In current medical practice, the only measurable parameter used to distinguish between healthy individuals and Parkinson's disease patients is the duration of the experiment.

A novel approach to detect PD uses Kinect, a motion sensing device that consists of a color camera and an infrared depth camera. To generate the dataset used in this thesis, Kinect 2012 was used. It was able to track skeletons of two people, detecting 20 points in a human body [11] (Figure 1.2). The newer version of Kinect released in 2013 was improved to be able to track up to six skeletons simultaneously. The number of joints detected by the sensor increased from 20 to 25 and the skeleton became more anatomically correct [12]. A newer version of Kinect sensor could be used instead of the older one in the same set-up. Kinect was proven to be sensitive enough to capture the differences between the PD patients and healthy elderly people [13].

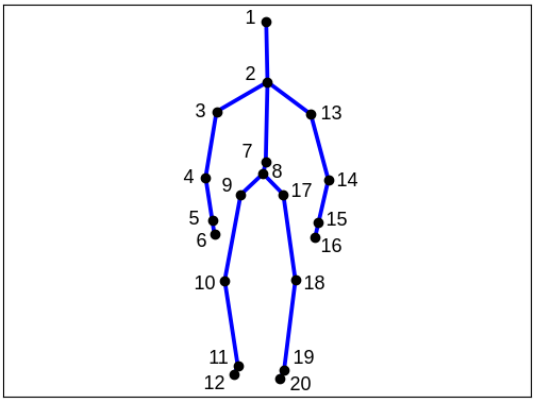


Figure 1.2: 1 - Head, 2 - Spine Shoulder, 3 - Right Shoulder, 4 - Right Elbow, 5 - Right Wrist, 6 - Right Hand, 7 - Spine Mid, 8 - Spine Base, 9 - Right Hip, 10 - Right Knee, 11 - Right Ankle, 12 - Right Foot, 13 - Left Shoulder, 14 - Left Elbow, 15 - Left Wrist, 16 - Left Hand, 17 - Left Hip, 18 - Left Knee, 19 - Left Ankle, 20 - Left Foot

1.2 Problem statement and contribution of the thesis

Formally, research problem may be stated as follows. Design a method to distinguish between PD patients and controls based on the data that can be extracted from Up-and-Go test. This problem naturally splits into three following sub problems.

1. Precise the setting and capture the gait during the Up-and-Go test.
2. Choose the set of features possessing the highest discriminative power.
3. Choose and train machine learning technique(s) able to differentiate between healthy elderly people and PD patients.

In this work, the data previously recorded with Kinect were used [1]. The significance of this setting in combination with Motion Mass parameters (MM) to differentiating healthy individuals from PD patients was also proved by [1]. Compared to the approach described in their publication, the main novelty of the current thesis is calculating the same parameters on steps extracted from moving forward motion (phase 2, Figure 1.1) of the experiment as opposed to calculating the parameters for the whole phase without splitting it down to smaller periods. Furthermore, other phases of the experiment are also analysed in the framework of this work. The main focus lays in feature selection, selection of appropriate machine learning techniques and developing a framework for training

classifiers. All the methods used should be acceptable from medical point of view. Combination of features and classifiers should have a clear meaning and interpretation from therapists' perspective. Analysis relies on extraction of MM parameters for all of the phases of the experiment and step extraction algorithm developed during the Bachelor's thesis[14].

The expected outcome of this thesis is...

- extended set of parameters to use with Up-and-Go test;
- PD classification models;
- a framework for automated reporting and model training.

1.3 Motivation

Healthcare is one of the numerous areas where application of Machine learning (ML) techniques is rapidly evolving to complement the work of the professionals. These applications have definitely not reached their limit. There are still many unexplored ways in which we can benefit from them. Naturally, machine learning or any other current technology will not be able to replace a human doctor in the foreseeable future. However, they can become useful tools to support medical specialists, to provide a way of minimizing the influence of human biases to the outcome of their work and to drive their attention to the deviations from the norm that can be an indicator of a disease.

However, there is a certain barrier that needs to be overcome to make such techniques an integral part of the medical evaluation. There are several reasons to that. Firstly, while people are more open to changes and new technologies in less substantial parts of their lives, they are more conservative in questions that relate to their health. Any new technology has to prove its efficiency and usefulness to the public to become widely acceptable. Another issue mentioned frequently in the last decade is interpretability of the results with respect to medical knowledge. More and more attention is drawn towards explaining the decision making process of Machine Learning models. So called "black-box" models such as neural networks usually provide a better accuracy but cannot be used in medicine because of the lack of traceability of the decision.

As a result, up to a very recent time medical community has remained sceptical about applicability of the computer assisted systems for diagnostics and modelling of

neurological disorders. Simple systems are considered too imprecise to capture the minor changes in movements and, therefore, unsuitable for medical environment. For the more advanced systems the main obstacle is lengthy and complicated set-up procedure which makes their practical application nearly impossible.

To make the transition to a computer-supported PD diagnostics procedure smoother and to only introduce the minimal necessary changes to a widely used routine, this thesis focuses on analysis of the Up-and-Go test with the essential expansion of the set of parameters used for estimating the condition of a patient. This work analyses the performance of the setting and the parameters applied to Parkinson's disease. In addition, the fact that this test is used for detecting other diseases such as Alzheimer's disease or multiple sclerosis [15] makes it potentially extendable to capturing other disorders.

1.4 Methodology and validation

In this thesis, the dataset acquired in [1] is analysed as a proof-of-concept for applicability of ML methods to PD diagnostics. This work solves the feature extraction problem from the Kinect recordings, followed by feature selection, engineering and choosing suitable machine learning techniques.

The dataset consists of recordings for 23 Parkinson's disease patients and 20 healthy controls of matching sex and age. It was captured using Kinect v1, which provided coordinates of a set of points in the subjects' bodies changing in time. The experiments were performed within a week before and after the medical intervention, with up to two recordings per week. The data from before and after the treatment are used separately in all of the following steps except where noted otherwise. Each recording contains data from several test trials (usually three).

First of all, the recordings are split into TUG test trials (experiments). Each experiment is then divided into phases as marked in the recording's file. Next, the forward walk (phase 2, Figure 1.1) is split further into steps. MM parameters are calculated for each step and for the whole phases. Those features are then averaged across the week for corresponding periods (step or the whole phase). Statistical hypothesis test is used to evaluate the applicability of the setting to differentiate PDs and controls. Student's t-test or Welch's t-test is performed depending on whether sample sizes are equal or not in order to determine p-values.

Feature selection process starts with calculating Fisher scores. This is a supervised feature selection method that allows to estimate the goodness of a given feature to differentiate between the result classes. This is done independently for first and second week. Features with the higher Fisher scores are chosen as candidates to build the classifiers. A lot of attention is drawn towards the medical interpretation of the results, meaning that symmetry of the joints used for classification is also a major factor in the feature selection process. It is important to include both sides of the body because it is then possible to take into account the differences between right and left side. Asymmetry in motor features are a major factor in PD diagnostics and sometimes the lack of it can indicate an alternative disability [4]. In order to avoid the dependence of the results from the way the dataset is split, evaluation of the outcome is performed with k-fold cross validation technique. The data points are split into k independent sets. At each iteration one set is left out as a testing data and the model is trained using all other sets. Accuracy is then calculated on the left-out set.

Average accuracy across the folds is calculated. And it is used as a main measure for goodness of a classifier. Apart from the accuracy, precision, recall and F1 score are investigated in the same way across different folds. The goodness of the outcome was determined in comparison with a corresponding metric of time only classifier.

To see how the classifiers would generalise and perform on the previously unseen and potentially different data, the following procedure was followed for best classifiers reported for first and second week. The classifier is trained with the most promising set of features on the data from the whole week and then tested on the other week.

1.5 Related Work

The presence of the motor symptoms of PD is considered cardinal for making PD diagnosis [5]. The changes in gross motor functioning influence the gait and can, therefore, be detected with gait analysis approach. The method that is commonly used to gather the data necessary for the assessment is marker based [16]. Special light-reflecting markers are places onto the body of a patient to track the movements of body segments. The number of markers used can vary. It is fifteen to sixteen to capture the lower-body movement and more than 30 to capture the kinematics of the whole body [16]. The drawback of such system is that it is intrusive, hard-to-use and the placement of the markers greatly influences the results. It is also time-consuming to place all the markers onto a body of

each patient.

There are also a lot of systems that use wearable sensors to collect the data for diagnosing Parkinson's disease. For example, [17] used a setting consisting of 8 Inertial Measurement Unit (IMU) sensors. They tried different combinations of range-of-motion and spatio-temporal parameters with different ML techniques. Using data from all 8 sensors and weighted Majority of votes classifier combining 6 simpler methods they achieved accuracy of 96%. Reducing the amount of sensors also reduced the accuracy. In this setting the knee-range of motion were shown to perform the best among the groups of features they investigated. This approach is hard-to-use for Parkinson's disease patients because of its lengthy setup and wearable placement process.

An alternative approach uses force sensitive resistors placed in the shoes. For example, [18], [19] and [20] used sets from the Physionet database, obtained with this technique, to extract parameters for the PD classification problem. Their findings suggest the importance of stride and stance length in gait analysis. The accuracies achieved are generally around 90 %.

A lot of effort has been made to simplify the data gathering process. Kinect and similar motion capture devices have a number of advantages compared to motion sensors used in the studies below: they are non-intrusive, easy to set-up and they eliminate the risk of inconsistent sensor placement. As a result, they have been extensively used not only in entertainment but also in different areas of medicine [21], [22]. For example work of [23] propose a non-intrusive gait analysis system based on Kinect sensor. Lack of precision for capturing lower limbs motion is claimed to be one of the major drawbacks in such systems.

However, [24] have shown that Kinect v2 sensors are sensitive enough and the data recorded with the device is sufficient to distinguish between Parkinson's disease patients and healthy individuals. In addition, [13] has compared the two versions of Kinect and explored the applicability of the technology in clinical setup. Naturally, the second version was shown to have higher percentage of gait parameters suitable to distinguish healthy people and PD patients.

[25] have also investigated the performance of Kinect sensors and their applicability to medical assessment. As Kinect sensor is discontinued, they also give an overview of the alternative hardware and software that can be used instead. There are several similar devices that can be used such as ASUS Xtion and LIPSedge DL. Furthermore, [25] elaborate the software alternatives that can be used instead of Kinect Software Development

Kit (SDK). For example, OpenPose project is able to identify and track anatomical points using just a 2D camera images. Using several 2D cameras the system is capable of 3D tracking as well.

In May 2018 Microsoft announced the development of Azure Kinect that will be combining the new improved depth sensor with the Azure AI services. The company also promises the continuation of support for Kinect SDK through forums and suggests Intel's RealSense depth cameras as an alternative for the hardware [26].

PD also affects fine motor skills and therefore usage of fine motor test for patients is under active research. For instance, [27] digitalised and proposed a set of parameters to support visual analysis of the clock drawing test used in diagnosing PD. Another test suitable for detecting changes in fine motor skills specific to PD patients called the Poppelreuter's Test is converted to a digital representation and described by [28].

Some people suffering from PD may experience voice and speech disturbances. Several studies [29–39] have made attempts to detect dysphonia (a weakness in voice production) and distinguish the voice data of PD patients and healthy people. The number of research papers devoted to this approach can be partly explained by the availability of publicly available datasets containing voice data [40]. Although, the accuracy usually exceed the 90 % threshold of classification accuracy, changes in vocal cords are exhibited in only some of the cases making this features not generally applicable.

Unlike the voice-oriented approaches [29–39] where ML models are used to estimate the probability of PD, the features used in the present research have clear medical meaning and therefore provide an easy way to interpret achieved results. In contrast to the system proposed by [17], the data gathering system used to collect the data for this work is non-intrusive, simple and is an extension to the widely used TUG test. Kinect-based approach used in this thesis gives a bigger range of possible features to use for classification enabling to analyse how different joints contribute to the outcome of the assessment. This data is missing in the research that concentrate their attention on legs and feet [18–20]. The problem of discontinuation of Kinect sensors can be solved with other similar devices and the software can also be replaced by open-source solutions.

2. Implementation

This chapter presents the tools and algorithms used for this work. It describes the dataset and the data collection setting. The calculation of MM parameters is explained and choice of classifiers is justified.

2.1 Tools

The feature extraction scripts were developed in Python language in PyCharm IDE environment. Data were processed and transformed using numpy, pandas and scipy libraries. Statistical tests were performed using scipy library stats module. For calculation of Fisher scores an open-source library scikit-feature was used [41]. Machine learning models were trained using scikit learn library. Visualisations were created with matplotlib and graphviz.

2.2 Dataset

For the purpose of this experiment the recordings of 20 Controls (C) and 23 PD patients of matching age and sex were used. The number of observations may vary for different weeks due to filtering out the inconsistent phase marking. So the second week's data uses data from 20 controls and 20 PD patients, while for the first week is 18 and 20 correspondingly.

For the purpose of data collection the classical setting of TUG test was extended with Kinect camera and a laptop. The position of the Kinect camera was dictated by the limitations of the sensor - Kinect is able to provide skeletal tracking of people that are at distance from 0.5 to 4 meters from the sensor, whereas the tracking is most accurate in range of 1.5 to 3.5 meters [42, 43]. The experimental set-up and camera placement is

depicted in Figure 2.1.

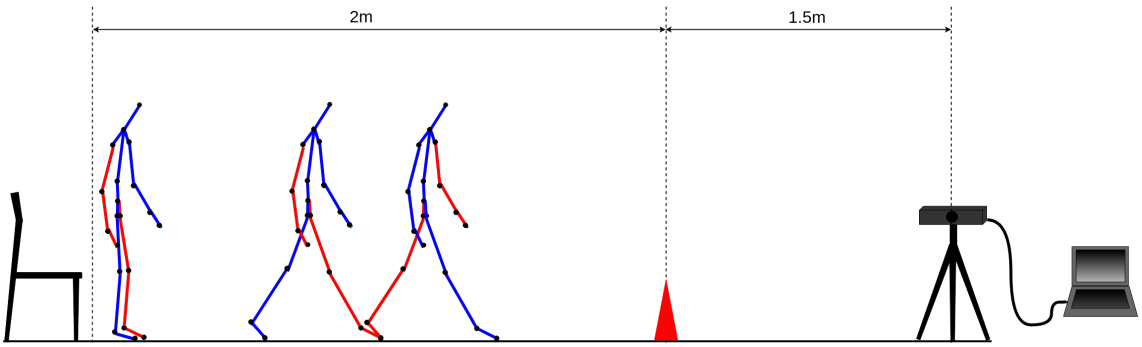


Figure 2.1: Experimental setting. The distances in TUG are chosen according to the limitations of the Kinect sensor. The sensor is located 3.5 meters away from chair and 1.5 meters away from marker where the patients should turn around. Marker is depicted by a triangle.

The data are stored in CSV format. Each recording contains up to three repetitions of the experiment. Each trial of the experiment contains all the phases: stand up, walk forward, turn around, walk back, turn back around, sit down. Each row of the recording contains a timestamp, and three coordinates for each joint detected by Kinect. If a joint is not recognised by the sensor or is occluded, there is an indication that the joint is untracked.

During the collection of this dataset the phases were detected manually by therapists who also gave the appropriate commands to the patients. However, since the collection of the dataset, the data acquisition process was simplified by Bernstein in his Bachelor's thesis [44]. That system is able to detect phases autonomously as well as give commands for action (such as "GO" for example).

2.3 Motion Mass parameters

The notion of Motion Mass (MM) parameters was firstly introduced by [22] to measure the smoothness and amount of motion. It was then proved in [1] that using the setting employed in this work in combination with MM parameters is promising approach for distinguishing PD patients from healthy people.

Splitting the walking forward phase into smaller intervals - steps - were shown to give better results for PD classification problem [45]. It was proven to improve the separability of the data for the two classes [14, 45]. This approach is also adopted in this work.

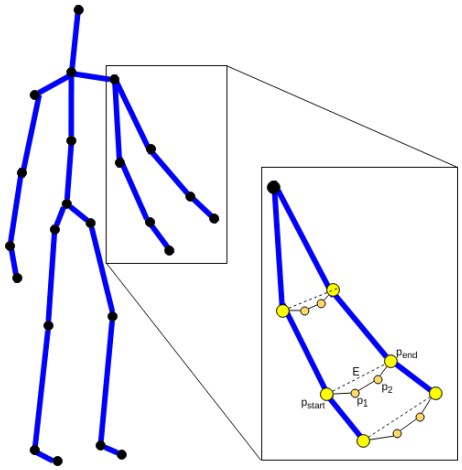


Figure 2.2: Depiction of the movement of a hand performed in time t. The p_{start} indicates the starting position of a joint. The p_{end} shows the position of a joint after the movement. The distance between those two points is E (illustrated by a dotted line). p_1 and p_2 represent intermediate points captured by a system. Tm is sum of the distances between the points p_{start}, p_1, p_2, p_{end}.

MM features can be computed for any point (in our case joint) and for any time interval. In this work, the main focus is drawn towards steps, so the starting time is the beginning of each step and the end is the time when the step ends. Intermediate points are snapshots of coordinates at each point in time that Kinect was able to capture.

For the set of joints $J = \{j_1, \dots j_n\}$ Motion Mass is defined as a vector $M = \{E_j, Tm_j, Vm_j, Am_j, Jm_j, E_j/Am_j, E_j/Tm_j, t\}, j \in J$. Trajectory mass Tm_j, velocity mass Vm_j, acceleration mass Am_j, jerk mass Jm_j, and Euclidian distance E_j are defined as follows:

$$Vm_j = \sum_{t=1}^{n} v_{j_t} \tag{2.1a}$$

$$Am_j = \sum_{t=1}^{n} a_{j_t} \tag{2.1b}$$

$$Jm_j = \sum_{t=1}^{n} j_{j_t} \tag{2.1c}$$

$$Tm_j = \sum_{t=1}^{n} d_{j_t} \tag{2.1d}$$

$$E_j = d \tag{2.1e}$$

where v_{j_t}, a_{j_t} and j_{j_t} are correspondingly velocity, acceleration and jerk calculated in each point. d stands for the distance between the staring and ending position of a joint and d_{j_t} distance between the intermediate points.

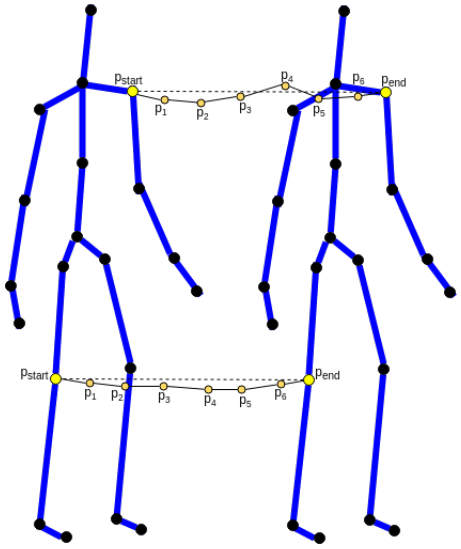


Figure 2.3: MM enable the analysis of upper-body joints as well as lower body joints. It is also possible to measure the movements of a combination of joints, for example right knee and left shoulder.

Other two features are computed as Euclidean distance divided by corresponding parameter.

E_j is a distance between starting and ending positions of a joint. Tm_j is the length of a trajectory the joint travelled in time t, that is sum of distances between all intermediate points (Figure 2.2). Vm_j, Am_j and Jm_j are the sum of corresponding parameters calculated in each point detected by the Kinect. The features are calculated for all joints that Kinect sensor is able to detect. In addition, it is possible to sum the features to investigate the movement of several joints together (Figure 2.3).

2.4 Classification

The main goal of this work is to distinguish between healthy people and those who suffer from PD. This is a two-class classification problem so the output is 0 for controls and 1 for PD patients. In this work several different ML techniques were investigated: Decision Tree (DT), Logistic Regression (LogR), kNN (k=5). The choice of those algorithms is justified by the simplicity and high interpretability of the models and the ability to give reasonable results considering the amount of data and the amount of features that we have. Moreover, the Support Vector Machines (SVM) model is used for comparison as well, but no hyperparameter tuning is performed.

To estimate the goodness of a classifier the cross validation technique was used. Dataset is split into 4 equal parts and in each iteration a classifier is trained on three parts and tested on the one that is left out. Average accuracy, precision, recall and F1 score are then reported and compared to the baseline. The baseline of the performance of a classifier is average accuracy of the classifier that uses the time taken by the subject to complete the experiment as an only feature. This was chosen as a baseline because it is the only parameter used by therapists for the evaluation. In addition, to see if the models generalize well to different data, the best models for each week are trained on data from one week and then tested on the other week's data producing an accuracy and confusion matrix.

2.5 Reporting

As a side result of the thesis, an automatic model training and reporting system was developed. It trains a set of different classifiers and creates a Hypertext Markup Language (HTML) and pdf reports. Each report contains a detailed information about each of the folds of k-fold cross validation such as testing accuracy, precision, recall, F1 score. In addition, it shows the averaged statistics for all of the four classifiers. It also contains pictures of the data, visualisations of decision tree classifiers and depiction of which joints are investigated for each particular case. An example report can be found in Appendix A.5.

3. Analysis

Figure 3.1: General workflow

The main workflow can be divided into three major parts: feature extraction, feature selection and models' training as shown in Figure 3.1 (full diagram in Appendix A.4). This chapter goes through the flow of the work step by step, presenting the results of each phase and stating the implications those results have on the next step.

3.1 Feature Extraction

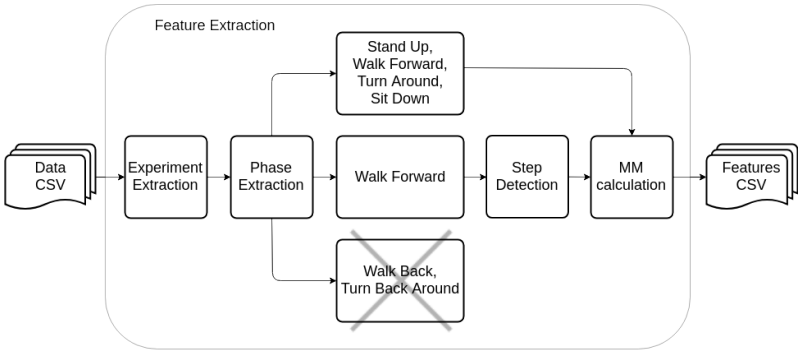


Figure 3.2: Feature extraction workflow

The main procedure of feature extraction produces 180 parameters (9 MM parameters for each of 20 joints, see Figure 1.2). It can be performed for arbitrary time intervals and then aggregated by taking average. The process adapted in this work is depicted in Figure 3.2.

This work mostly focuses on analysis of steps, thus, the main set of features contains

MM parameters that are calculated for each step and then averaged across the steps for each week. The results are one set of features per week and per person.

However, to investigate the possibility of combining the step-based features with features of other phases, the same set of parameters is also calculated for each phase of Up-and-Go test without splitting down to smaller periods. The walking back motion as well as turning back around were excluded because Kinect cannot reliably detect the human skeleton from the back.

3.2 Statistical Analysis

To prove the applicability of the set-up for classification task statistical hypothesis test is performed. The null-hypothesis is that the data is normally distributed and there are no significant difference between the two groups. The alternative hypothesis states that the groups are different. The results of the t-test are presented in Tables 3.1, 3.2. Welch's t-test was performed instead of the classical Student's t-test for the first week because it tends to give more reliable results with samples of unequal sizes. The second week's p-values were calculated with Student's t-test. The results show that the setting is sensitive to capture the differences between the PD and controls with significance level 0.01. Full results can be found in Appendix A.1.

Table 3.1: p-values characterizing ability of MM parameters to distinguish the two groups for the first week.

Joint	parameter	p-value	t-statistic
HipRight	Vm	0.000114	4.331885
HipLeft	Vm	0.000118	4.317425
HipRight	Tm	0.000591	3.768417
ShoulderLeft	Vm	0.000649	3.734449
HipLeft	Tm	0.000704	3.706690
ShoulderLeft	Am	0.001359	3.514941
KneeRight	Vm	0.001364	3.519113
ShoulderRight	Vm	0.001912	3.366591
KneeLeft	Vm	0.002594	3.242248
KneeRight	Tm	0.003656	3.141354
ShoulderLeft	Tm	0.004071	3.068720
HipRight	E	0.004306	3.047474
ShoulderRight	E	0.004533	3.028032
KneeLeft	E	0.004563	3.025357
KneeLeft	Tm	0.006445	2.897070
ShoulderRight	Tm	0.007292	2.855015
KneeRight	E	0.007616	2.827512
HipLeft	E	0.007685	2.824928
ShoulderLeft	E	0.010102	2.716117

Table 3.2: p-values characterizing ability of MM parameters to distinguish the two groups for the second week.

Parameter	Joint	p-value	t-statistic
ShoulderRight	Tm	0.000009	5.122567
ElbowRight	Tm	0.000011	5.066750
ShoulderRight	Vm	0.000011	5.051259
ElbowRight	Vm	0.000011	5.049883
KneeRight	E	0.000014	4.986571
ShoulderRight	E	0.000016	4.944651
SpineShoulder	Tm	0.000016	4.936453
HipRight	Tm	0.000017	4.917614
HipRight	Vm	0.000018	4.904116
HipRight	E	0.000021	4.858270
HipLeft	Tm	0.000049	4.577930
HipLeft	Vm	0.000053	4.551003
HipLeft	E	0.000062	4.502647
ShoulderLeft	Tm	0.000117	4.292722
ShoulderLeft	Vm	0.000120	4.284801
KneeRight	Tm	0.000134	4.248813
KneeLeft	E	0.000148	4.216791
KneeRight	Vm	0.000174	4.162549
ShoulderLeft	E	0.000187	4.138537

3.3 Feature Selection

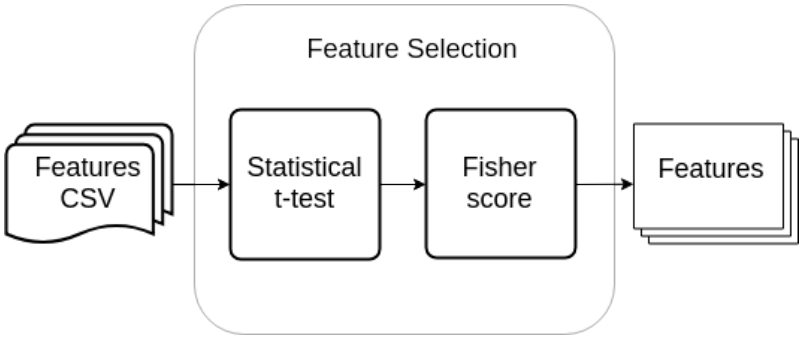


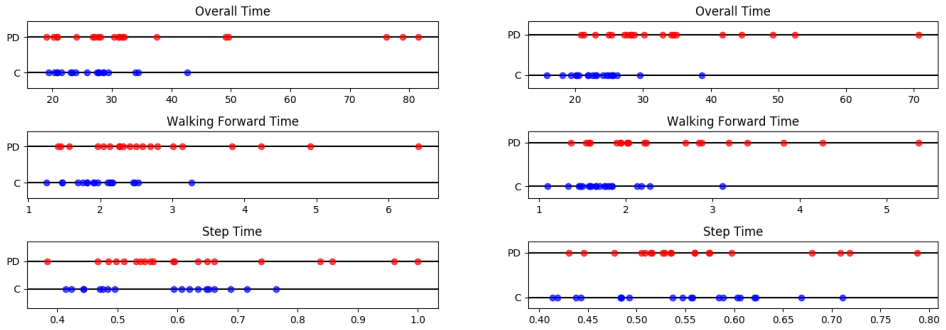
Figure 3.3: Feature selection workflow

Because the number of features extracted exceed the number of observation it is necessary to perform feature selection in addition to significance testing to choose only a few features that would be able to capture the differences between the classes. What is more, keeping a lower amount of features would allow to make models interpretable. The feature selection process is visualised in Figure 3.3. Statistical testing was described in Section 3.2 and provided the knowledge of which features are significantly different for controls and PD patients. This section focuses on the second stage of the feature selection.

One of the methods that provide a good insight and ranking of the features for numerical data is Fisher score [46]. The Fisher score values range from 0 to 1, where higher number usually indicates features with more distinguishable power. The formula is provided below. k is the number of classes. p_j is the fraction of points of class j. μ_j and σ_j are mean and standard deviation of class j correspondingly. μ is the mean of the whole dataset for a given features [46]. In our case k is two for C and PD.

$$F = \frac{\sum_{j=1}^{k} p_j(\mu_j - \mu)^2}{\sum_{j=1}^{k} p_j \sigma_j^2} \tag{3.1}$$

Firstly, only the duration of individual steps, walking forward time and the entire task duration was examined. This represents the methodology currently used in the clinical environment, whereby the total time taken by the subject to perform the test is used for PD diagnosis. Figure 3.4 shows the separability of the data for different time intervals. The discriminative power decreases with splitting the time into smaller chunks. And in opposite, the differences in time between the two groups accumulate for longer time interval. Fisher scores for the test duration, walking forward phase duration and for average step duration are presented in Table 3.3.



(a) First week (b) Second week

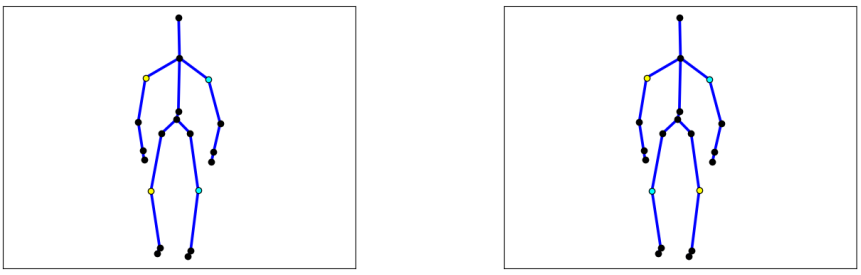
Figure 3.4: Visualisation of the data for different time intervals: duration of TUG test, duration of a walking forward phase and duration of a step. Each dot represents a PD patient (top row, red) or a healthy control (bottom row, blue), placed on a horizontal axis according to the total time taken to perform the task (top panels), the duration of the walking forward phase (middle panels) or the average duration of their steps (bottom panels)

Next, the predictive power of the MM parameters described in Chapter 2 Equations 2.1 was ranked in an attempt to find additional features that would aid in PD diagnostics. The top-scoring features for the first and second weeks, which were later used for classifier training, are presented in Tables 3.4 and 3.5, respectively (see full results in Appendix

Table 3.3: Fisher scores for different time intervals for 1st and 2nd weeks

Time period	1W Fisher Score	2W Fisher Score
TUG Test	0.112040	0.329513
Walk Forward Phase	0.176002	0.253020
Step	0.048310	0.015126

A.2).



(a) Sides features. One feature combines left shoulder and left knee, the other - right shoulder and right knee.

(b) Cross features. One feature combines left shoulder and right knee, the other - right shoulder and left knee.

Figure 3.5: Combination of joints. Different colors than black indicate that feature for this joint was used in the classifier. Joints coloured in the same color are combined together.

The Fisher score results are different for the two weeks under observation. For the first week, the best of the MM parameters were Vm followed by Tm. And the distribution of the good parameters between left and right side is almost equal. For the second week, however, there is no clear leader among the MM parameters. Tm, Vm and E have the highest discriminative power. The interesting observation from the results of the second week's data is than the right side joints perform significantly better that left side features.

Apart from Fisher score, a lot of attention is drawn towards medical interpretation of the results. For diagnosing PD symmetry of motor symptoms is an important feature, where the lack of it can sometimes indicate alternative diagnosis. The changes in motor functions and, therefore, the abnormalities can start from either of the sides - left or right. Even if analysis of one side gives better results on the given dataset, the method will not generalise. That is the reason why symmetrical points are always used together in the classifier despite some other candidates' better Fisher score. In addition, the combination of several joints are analysed. The main idea behind combining the joints is to see how upper-body joints and lower-body joints contribute to the final outcome. Also the same principle is used in left or right side joints. One of the combinations included the opposite side upper and lower-body joints (so called cross) 3.5b. That means features of

Table 3.4: Fisher scores of MM parameters for first week. The best MM parameter is Vm and there are equal number or similarly performing left and right side features.

Joint	Parameter	Fisher Score
HipRight	Vm	0.512484
HipLeft	Vm	0.510796
SpineMid	Vm	0.503264
Head	Vm	0.500898
SpineBase	Vm	0.492910
SpineShoulder	Vm	0.442011
ElbowRight	Vm	0.426837
HipRight	Tm	0.387325
ShoulderLeft	Vm	0.383611
HipLeft	Tm	0.375525
SpineMid	Tm	0.367021
SpineBase	Tm	0.362869
KneeRight	Vm	0.355398
ShoulderLeft	Am	0.353612
Head	Tm	0.346816
ElbowRight	Tm	0.327350
SpineShoulder	Tm	0.310331
ShoulderRight	Vm	0.302826
KneeLeft	Vm	0.293221
KneeRight	Tm	0.282561
ShoulderLeft	Tm	0.258273
HipRight	E	0.255310
KneeLeft	E	0.251511
ShoulderRight	E	0.251134
KneeLeft	Tm	0.234392
HipLeft	E	0.220955
KneeRight	E	0.219018
ShoulderRight	Tm	0.217997
ShoulderLeft	E	0.204045

Table 3.5: Fisher scores of MM parameters for second week. E, Vm and Tm perform equally well. However, right side features seem to have higher discriminative power.

Joint	Parameter	Fisher Score
ShoulderRight	Tm	0.690544
ElbowRight	Tm	0.675578
ShoulderRight	Vm	0.671453
ElbowRight	Vm	0.671087
KneeRight	E	0.654366
ShoulderRight	E	0.643410
SpineShoulder	Tm	0.641278
HipRight	Tm	0.636393
HipRight	Vm	0.632904
ElbowRight	E	0.629742
HipRight	E	0.621126
SpineShoulder	Vm	0.617846
SpineBase	Tm	0.597560
SpineMid	Tm	0.597017
SpineBase	Vm	0.593224
SpineMid	Vm	0.591947
SpineMid	E	0.579561
HipLeft	Tm	0.551512
HipLeft	Vm	0.545043
HipLeft	E	0.533522
ShoulderLeft	Tm	0.484933
ShoulderLeft	Vm	0.483145
KneeRight	Tm	0.475064
KneeLeft	E	0.467930
KneeRight	Vm	0.455969
ShoulderLeft	E	0.450723
ShoulderLeft	Am	0.351921
KneeLeft	Tm	0.348323
KneeLeft	Vm	0.332309

right shoulder were combined with left knee and left shoulder was summed with right knee. Another option combined the features of left side together and features of right side together 3.5a.

3.4 Model Training and Analysis

The high-level description of model training and evaluation process is presented in Figure 3.6.

After the feature selection process, the reasonable amount of parameters is available

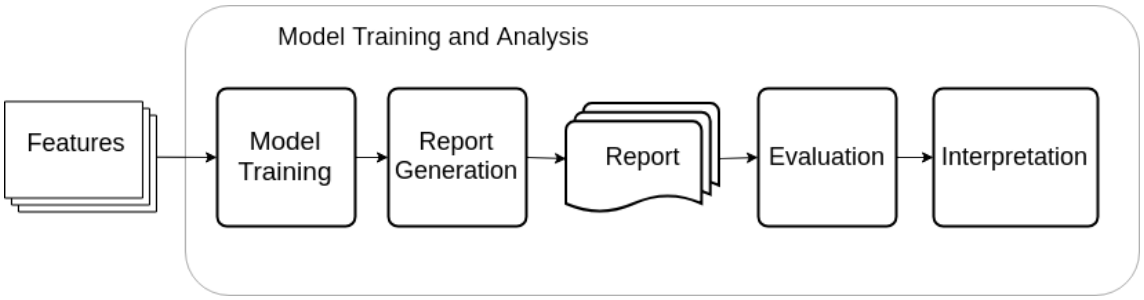


Figure 3.6: Model training and validation workflow

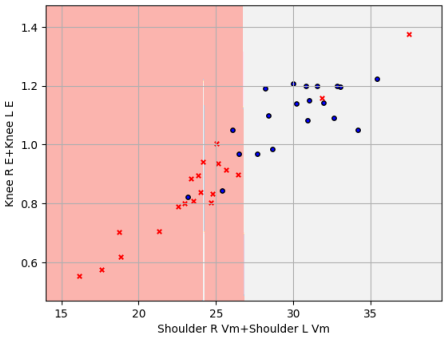
for classification. Because of need of interpretability, the following well-interpretable models are taken into account: DT, SVM, LogR and kNN. The SVC model is also used for comparison. The accuracy is computed for each fold of k-fold validation (k=4) and average value across the folds is reported.

Table 3.6 shows the accuracies achieved on classifiers with experiment duration. The baseline for the goodness of a classifier is a classifier using duration of TUG experiment as an only feature (also presented in Table 3.6, first row). The visualisation of the separability of the data based on experiment duration is presented in Figure 3.4. The controls and PD patients have a large overlap in the results. Therefore, it is unlikely that the experiment duration alone will have a good predictive power. Visual comparison in the charts makes it clear that the data from the first week should have a lower classification accuracy because the two groups have a larger overlap and are hard to distinguish.

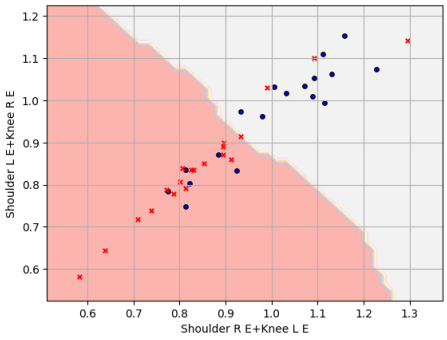
For the second week baseline kNN performed best, achieving accuracy of 82.5%. The baseline result for the first week is significantly lower: only 58.89% for Decision Tree classifier. For the data from second week, using the time in combination with other parameters does not perform much better than the baseline improving the result by 2.5 % only with Vm of hips. Most of the time, however, it improves accuracy of Logistic Regression making it as good as the baseline. For the first week, most of the classifiers are able to perform better than the baseline. For example, adding Vm of sides to time gives an increase of more than 19%.

The average accuracies for the classifiers that use only MM parameters are presented in Table 3.7. In this table, the best-performing classifiers are presented, alongside their counterparts which use the same features but a different algorithm. The accuracies not exceeding the baseline are written in grey font. The highest accuracies are highlighted in bold. Those are classifiers trained without duration of the experiment, only MM features. As a rule, k-nearest neighbours classifier performs the best on combined MM features. The first week's classifiers, having a considerably lower baseline to compare to,

are able to outperform time-only classifier with every combination of features presented in this work. The best achieved accuracy is 76.67% for SVM using combination of Tm of shoulder features. It is closely followed by kNN with two features: Vm of shoulders and E of knees having 76.39%. That is total increase in 17% compared to baseline. The best classifier for the second week was able to increase the accuracy by 5% compared to baseline. The best classifier is kNN using two features: Shoulder R E+Knee L E, Shoulder L E+Knee R E (accuracy 87.5%). There are also several classifiers that were able to achieve accuracy of 85%. It is also possible to compare the outcomes to the corresponding classifier with time as presented in Table 3.6. As a rule, classifier without time perform slightly better than the ones with time.



(a) kNN (features: Shoulder R Vm+Shoulder L Vm, Knee R E+Knee L E) Trained on first week, tested on second week.



(b) kNN (features: Shoulder R E+Knee L E, Shoulder L E+Knee R E) Trained on second week, tested on first week.

Figure 3.7: Decision boundary visualisation for test data. Right, lighter area indicates the healthy people. Left, darker color shows PD patients. PD patients are depicted with red cross and the symbol for controls is blue circle. Despite the fact the decision boundary divides the whole space into two classes, if the new observations would fall far away from the representatives depicted on the figure, they should be investigated further.

To test how the classifiers generalise to previously unseen data, best classifiers are trained on one week's worth of data and tested on the other week. The same evaluation process was repeated using the duration of the experiment as the only feature to get a baseline performance. For this, kNN classifier was chosen because it is generally performing better. The best classifier for the second week (features: Shoulder R E+Knee L E, Shoulder L E+Knee R E) was trained on data from second week and tested on the data from first week. The resulting accuracy is 73.6%. Decision boundary is depicted in Figure 3.7b and confusion matrix is presented in Table 3.8. The corresponding result of the duration-only classifier was 60.5%, which is notably lower.

The best classifier for the first week (features: Shoulder R Vm+Shoulder L Vm,Knee

R E+Knee L E) was trained on data from first week and tested on the data from second week. The resulting accuracy was 82.5%. Visual representation of the classification is shown in Figure 3.7a. Confusion matrix for the test data can be found in Table 3.9. Time-only classifier performed 17.5% worse (65%).

All in all, the results presented in this Chapter showed that test duration does not have enough discriminative power do distinguish between healthy individuals and people with PD. The baselines established by the test duration are 58.89 % and 82.5 % for the first and second week correspondingly. It is worth noting that the results from the first week are worse than those from the second. However, classifiers using features with MM parameters are able to outperform the baseline on both weeks. The best achieved accuracies not containing duration of an experiment with k-fold validation technique are 76.67 % for the first week and 87.5 % for the second. In addition, it is shown that adding time of an experiment to MM-based features does not improve the accuracy of a classifier. However, adding time significantly improves the accuracy of Logistic Regression Classifier.

Table 3.6: Accuracy for classifiers with time feature and MM parameters for 1st and 2nd weeks. The best or near best classifiers are marked in bold. The accuracies not exceeding the baseline are written in grey font.

Features	Classifier	1W accuracy	2W accuracy
t	DT	**58.89**	80
	kNN	53.06	**82.5**
	LogR	47.78	67.5
	SVM	55.84	77.5
Shoulder R E+Knee L E	DT	44.16	77.5
Shoulder L E+Knee R E	kNN	55.56	**82.5**
t	LogR	**63.34**	**82.5**
	SVM	58.61	80
Shoulder R Vm+Knee L Vm	DT	50	77.5
Shoulder L Vm+Knee R Vm	kNN	**68.61**	**82.5**
t	LogR	66.11	77.5
	SVM	50	67.5
Shoulder R E+Knee R E	DT	39.44	77.5
Shoulder L E+Knee L E	kNN	55.56	**82.5**
t	LogR	**63.34**	**82.5**
	SVM	58.61	80
Shoulder R Vm+Knee R Vm	DT	55.56	75.0
Shoulder L Vm+Knee L Vm	kNN	68.84	**82.5**
t	LogR	**78.61**	80
	SVM	47.50	62.5
Hip R Vm+Hip L Vm	DT	55.28	77.5
t	kNN	63.06	82.5
	LogR	**73.89**	**85**
	SVM	55.56	75
Knee R E+Knee L E	DT	45	77.5
t	kNN	53.06	**82.5**
	LogR	**58.06**	**82.5**
	SVM	**58.61**	80
Shoulder R Vm+Shoulder L Vm	DT	55.56	77.5
Knee R E+Knee L E	kNN	60.56	**82.5**
t	LogR	**76.39**	**82.5**
	SVM	60.28	80
Shoulder R Tm+Shoulder L Tm	DT	57.22	65
t	kNN	53.06	**82.5**
	LogR	**60.56**	**82.5**
	SVM	58.61	80

Table 3.7: Accuracy for best classifiers for 1st and 2nd weeks. The best or near best classifiers are marked in bold. The accuracies not exceeding the baseline are written in grey font.

Features	Classifier	1W accuracy	2W accuracy
Shoulder R E+Knee L E Shoulder L E+Knee R E	DT	58.06	70
	kNN	66.11	**87.5**
	LogR	50.56	70
	SVM	53.06	80
Shoulder R Vm+Knee L Vm Shoulder L Vm+Knee R Vm	DT	58.34	67.5
	kNN	64.17	**85**
	LogR	50.28	67.5
	SVM	55.28	72.5
Shoulder R E+Knee R E Shoulder L E+Knee L E	DT	50.28	80
	kNN	66.11	**85**
	LogR	50.56	70
	SVM	53.06	77.5
Shoulder R Vm+Knee R Vm Shoulder L Vm+Knee L Vm	DT	61.11	72.5
	kNN	63.89	**85**
	LogR	52.78	57.5
	SVM	53.06	65
Hip R Vm+Hip L Vm	DT	57.78	77.5
	kNN	**74.17**	80
	LogR	58.33	70
	SVM	66.39	82.5
Knee R E+Knee L E	DT	50.28	**85**
	kNN	**74.17**	82.5
	LogR	47.78	70
	SVM	58.61	80
Shoulder R Vm+Shoulder L Vm Knee R E+Knee L E	DT	60.28	77.5
	kNN	71.11	82.5
	LogR	50.83	67.5
	SVM	**76.39**	82.5
Shoulder R Tm+Shoulder L Tm	DT	68.61	75
	kNN	**76.67**	**85**
	LogR	47.78	65
	SVM	55.56	**85**

Table 3.8: Confusion matrix for kNN with Shoulder R E+Knee L E, Shoulder L E+Knee R E for test data (first week).

Actual \ Predicted	C	PD
C	12	6
PD	4	16

Table 3.9: Confusion matrix for kNN with Shoulder R Vm+Shoulder L Vm,Knee R E+Knee L E for test data (second week).

Actual \ Predicted	C	PD
C	16	4
PD	3	17

4. Discussion

The amount of data available for this study makes it hard to give definitive answers about how well this system will perform in a clinical setting. However, this preliminary study already revealed a number of important aspects of PD diagnosis, both using ML methods and in general. Strikingly, it turned out that the duration of the experiment, widely used for diagnosis in the clinic, is not the best feature to distinguish between the PD patients and healthy people. This was proven by the comparatively low Fisher scores and classification accuracies especially for the first week's data.

This work proposed an extended set of parameters that can be calculated for the classical setting of the Up-and-Go test. Using some of those features, it is possible to detect the PD patients with higher accuracy than with the test duration alone. Furthermore, unlike other approaches utilizing gait analysis tests [18–20], this set of parameters provides a deeper insight into the movement abnormalities by allowing to analyse the upper body joints. It was demonstrated that the shoulders possess a high discriminative power and contribute significantly to distinguishing movements of healthy elderly people and those suffering from Parkinson's disease. Interestingly, different classifiers - differing in both the ML algorithm and the parameters used - performed best on the data from the first and the second week. However, taking into account several top-performing classifiers, it is evident that the kNN algorithm worked best in general (Table 3.7).

A preference for the features of the knees, shoulders and hips can be observed (Tables 3.1, 3.2, 3.5, 3.5). This preference might have a physical explanation: perhaps a change in their movement is easiest to spot in PD patients. A number of reasons may lead to this, e.g. those joints being the most affected by a change in the patients' gait, or even the disease directly influencing their movement the most. Although obviously further investigations are needed to back up any of those speculations, they illustrate well how ML can be used as a tool in medical research. Interestingly, when analysing the data from the second week, the features describing the right-side joints had a notably higher discriminative power than the left-side ones. Since this was only observed for one of

the datasets, it could reflect some technical issue with that particular experiment, which lead to better detection of the right-side features. Therefore, it is always important to do "sanity checks" on the data sets and keep in mind that the results need to be medically interpretable.

It was shown that splitting the walking forward phase down to steps and analysing those smaller periods improves discriminative power of MM parameters. It can be deduced from the Fisher scores presented in this work (Tables 3.4 and 3.5) and those enclosed in Appendix A.3 that show the results for the whole walking forward phase. The features that had the highest predictive power among the MM parameters were E, Vm and Tm, i.e. the distance between the start and the end positions a joint, the sum of the velocities at each intermediate point, and the length of its trajectory within a step (Tables 3.1, 3.2, 3.5, 3.5; Figure 2.2; Equations 2.1). In a sense, those are the easiest parameters to measure, and therefore this preference may mean that their values are detected/computed by Kinect more accurately.

The difference between the two weeks - before and after therapy - might indicate that at different stages different set of parameters may be used to better differentiate between healthy and diseased people. However, these discrepancies might also be the result of the sample size of each of the datasets not being large enough. Therefore, this statement, as well as other conclusions of this study, need to be investigated on a larger dataset and a greater number of repeats to be able to draw reliable conclusions.

An important advantage of this approach is that it uses the setting widely used for diagnosing PD. The procedure of the test is not changed and the equipment needed is only extended by one Kinect sensor and a computer to perform the calculations. Therefore, the cost of learning this new technology by therapists is low, making it a plausible method to be adopted in the clinic in the near future. What is more, timed Up-and-Go test is used to detect numerous other conditions characterized by gait abnormalities. Therefore, it may be possible to use the same setup to diagnose other neurological diseases. This would also improve the reliability of the diagnostics for Parkinson's disease by increasing the ability to distinguish the different Parkinsonian diseases with similar symptoms.

5. Conclusion

This work is devoted to Parkinson's disease diagnostics problem. Diagnosing the disease is challenging because of the variability of the symptoms depending on the stage of PD and on the lifestyle of a patient. In addition, symptoms of Parkinsonian diseases are similar which makes it even harder to give a conclusive diagnosis on early stages. However, early diagnosis is crucial and can significantly improve the way of life of a patient. It is important to start the consulting and therapy as early as possible to be able to control the progression of the symptoms for a number of years.

One of the most commonly used tests in clinical setting is gait mobility test - Up-and-Go test. The only measurable parameter considered is usually duration of a test. It was proven that it is possible to distinguish between the healthy elderly people and those suffering of PD using only time of the test. Current work utilises the commonly used Up-and-Go test with only minor changes that enable data collection.

Motion Mass parameter notion was adapted in this work to extend the set of parameters calculated from the data. The main difference from works using the same notion is extraction of the features from steps. This showed better results comparing to using the duration of a phase as a main time interval for the calculation of MM.

The main distinction from other gait analysis approaches is the fact that despite using gait analysis as a base to diagnose the disease, current approach enables analysis of upper-body joints. The results showed that abnormalities in upper-body movements contribute significantly to understanding and diagnosing the disease. It is also possible to study the movement of groups of joints and compare the sides. The results showed that using Motion Mass parameters calculated on steps can be used to support the therapists in evaluation. The models outperformed the test duration based classifier - the baseline - proving to provide a better quantifiable outcomes for the TUG. The highest accuracy 87.5% was achieved on data from the second week using k-nearest neighbours classifier with cross joints' features.

This work can be considered a pilot study to justify the need to collect a larger dataset. Testing this approach using more data will give an opportunity to make more definitive conclusions about the goodness of the classifiers and establish a reasonable expectation to utilising the system in medical environment.

TUG test is also used for diagnosing other neurological diseases and detecting movement abnormalities. It can be possible to use the same setting for those purposes as well. However, it is necessary to conduct more experiments and collect data for people with different diagnosis to be able to distinguish between different illnesses. It is possible that by studying the differences in features across other diseases the reliability of this set up for detecting Parkinson's disease can improve due to a better classification of symptoms.

Acknowledgments

I am deeply grateful to my supervisor Sven Nõmm for being a constant source of inspiration, support and knowledge throughout the course of this work. I would also like to thank my family and close friends for their patience and support.

References

[1] Sven Nõmm, Aaro Toomela, Martti Vaske, Dan Uvarov, and Pille Taba. An alternative approach to distinguish movements of parkinson disease patients. *IFAC-PapersOnLine*, 49(19):272 – 276, 2016. ISSN 2405-8963. doi: https://doi.org/10.1016/j.ifacol.2016.10.546. URL http://www.sciencedirect.com/science/article/pii/S240589631632136X. 13th IFAC Symposium on Analysis, Design, and Evaluation ofHuman-Machine Systems HMS 2016.

[2] Paul Johns. Chapter 13 - parkinson's disease. In Paul Johns, editor, *Clinical Neuroscience*, pages 163 – 179. Churchill Livingstone, 2014. ISBN 978-0-443-10321-6. doi: https://doi.org/10.1016/B978-0-443-10321-6.00013-8. URL http://www.sciencedirect.com/science/article/pii/B9780443103216000138.

[3] Alfredo Berardelli, GK Wenning, Angelo Antonini, Daniela Berg, BR Bloem, Vincenzo Bonifati, Dina Brooks, DJ Burn, Carlo Colosimo, A Fanciulli, et al. Efns/mdses recommendations for the diagnosis of parkinson's disease. *European Journal of Neurology*, 20(1):16–34, 2013.

[4] J. B. Lehosit and L. J. Cloud. Early parkinsonism: Distinguishing idiopathic parkinson's disease from other syndromes. *Journal of Clinical Outcomes Management*, 22 (6):257–265, 2015. URL www.scopus.com. Cited By :1.

[5] J Jankovic. Parkinson's disease: clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry*, 79(4):368–376, 2008. ISSN 0022-3050. doi: 10.1136/jnnp.2007.131045. URL https://jnnp.bmj.com/content/79/4/368.

[6] Makoto Naoi and Wakako Maruyama. Cell death of dopamine neurons in aging and parkinson's disease. *Mechanisms of Ageing and Development*, 111(2):175 – 188, 1999. ISSN 0047-6374. doi: https://doi.org/10.1016/S0047-6374(99)00064-

0. URL http://www.sciencedirect.com/science/article/pii/S0047637499000640.

[7] Ronald B. Postuma and Daniela Berg. Chapter four - the new diagnostic criteria for parkinson's disease. In Kailash P. Bhatia, K. Ray Chaudhuri, and Maria Stamelou, editors, *Parkinson's Disease*, volume 132 of *International Review of Neurobiology*, pages 55 – 78. Academic Press, 2017. doi: https://doi.org/10.1016/bs.irn.2017.01.008. URL http://www.sciencedirect.com/science/article/pii/S0074774217300089.

[8] Addie Middleton and Stacy Fritz. Assessment of gait, balance, and mobility in older adults: Considerations for clinicians. *Current Translational Geriatrics and Experimental Gerontology Reports*, 2:205–214, 12 2013. doi: 10.1007/s13670-013-0057-2.

[9] Susan Morris, Meg E Morris, and Robert Iansek. Reliability of Measurements Obtained With the Timed "Up &; Go" Test in People With Parkinson Disease. *Physical Therapy*, 81(2):810–818, 02 2001. ISSN 0031-9023. doi: 10.1093/ptj/81.2.810. URL https://dx.doi.org/10.1093/ptj/81.2.810.

[10] Mary Thompson and Ann Medley. Performance of individuals with parkinson's disease on the timed up & go. *Journal of Neurologic Physical Therapy*, 22, 01 1998. doi: 10.1097/01253086-199822010-00009.

[11] Kinect for Windows SDK Beta. Features, 2011. URL https://www.microsoft.com/en-us/research/project/kinect-for-windows-sdk-beta/?from=http%3A%2F%2Fresearch.microsoft.com%2Fen-us%2Fum%2Fredmond%2Fprojects%2Fkinectsdk%2Fabout.aspx. Accessed 2019-04-21.

[12] Kinect for Windows SDK 2.0. Features, 2013. URL https://docs.microsoft.com/en-us/previous-versions/windows/kinect/dn782025(v%3dieb.10). Accessed 2019-04-21.

[13] A. P. Rocha, H. Choupina, J. M. Fernandes, M. J. Rosas, R. Vaz, and J. P. S. Cunha. Kinect v2 based system for parkinson's disease assessment. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1279–1282, Aug 2015. doi: 10.1109/EMBC.2015.7318601.

[14] Anna Krajuškina. An alternative approach for gait analysis of parkinson's disease patients. Bachelor's thesis, Tallinn University of Technology, 2017.

[15] Physiopedia. Timed up and go test (tug) — physiopedia,, 2019. URL https://www.physio-pedia.com/index.php?title=Timed_Up_and_Go_Test_(TUG)&oldid=206345. [Online; accessed 18-March-2019].

[16] Sarah J. Crane. Eldercare technology for clinical practitioners. *Mayo Clinic Proceedings*, 84(11):1050, 2009. ISSN 0025-6196. doi: https://doi.org/10.4065/84.11.1045-i. URL http://www.sciencedirect.com/science/article/pii/S0025619611606897.

[17] C. Caramia, D. Torricelli, M. Schmid, A. Muñoz-Gonzalez, J. Gonzalez-Vargas, F. Grandas, and J. L. Pons. Imu-based classification of parkinson's disease from gait: A sensitivity analysis on sensor location and feature selection. *IEEE Journal of Biomedical and Health Informatics*, 22(6):1765–1774, Nov 2018. ISSN 2168-2194. doi: 10.1109/JBHI.2018.2865218.

[18] Deepak Joshi, Aayushi Khajuria, and Pradeep Joshi. An automatic non-invasive method for parkinson's disease classification. *Computer Methods and Programs in Biomedicine*, 145:135 – 145, 2017. ISSN 0169-2607. doi: https://doi.org/10.1016/j.cmpb.2017.04.007. URL http://www.sciencedirect.com/science/article/pii/S0169260716306484.

[19] U. K. Pun, H. Gu, Z. Dong, and N. S. Artan. Classification and visualization tool for gait analysis of parkinson's disease. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2407–2410, Aug 2016. doi: 10.1109/EMBC.2016.7591215.

[20] Wei Zeng, Fenglin Liu, Qinghui Wang, Ying Wang, Limin Ma, and Yu Zhang. Parkinson's disease classification using gait analysis via deterministic learning. *Neuroscience Letters*, 633:268 – 278, 2016. ISSN 0304-3940. doi: https://doi.org/10.1016/j.neulet.2016.09.043. URL http://www.sciencedirect.com/science/article/pii/S0304394016307261.

[21] H.H. Manap, N.M. Tahir, and A.I.M. Yassin. Statistical analysis of parkinson disease gait classification using artificial neural network. In *Signal Processing and Information Technology (ISSPIT), 2011 IEEE International Symposium on*, pages 060–065, Dec 2011. doi: 10.1109/ISSPIT.2011.6151536.

[22] S. Nomm, A. Toomela, and J. Borushko. Alternative approach to model changes of human motor functions. In *2013 European Modelling Symposium*, pages 169–174, Nov 2013. doi: 10.1109/EMS.2013.30.

[23] N. Dao, Y. Zhang, J. Zheng, and J. Cai. Kinect-based non-intrusive human gait analysis and visualization. In *2015 IEEE 17th International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–6, Oct 2015. doi: 10.1109/MMSP.2015.7340804.

[24] Moataz Eltoukhy, Christopher Kuenze, Jeonghoon Oh, Marco Jacopetti, Savannah Wooten, and Joseph Signorile. Microsoft kinect can distinguish differences in overground gait between older persons with and without parkinson's disease. *Medical Engineering & Physics*, 44:1 – 7, 2017. ISSN 1350-4533. doi: https://doi.org/10.1016/j.medengphy.2017.03.007. URL http://www.sciencedirect.com/science/article/pii/S1350453317300887.

[25] Ross A. Clark, Benjamin F. Mentiplay, Emma Hough, and Yong Hao Pua. Three-dimensional cameras and skeleton pose tracking for physical function assessment: A review of uses, validity, current developments and kinect alternatives. *Gait & Posture*, 68:193 – 200, 2019. ISSN 0966-6362. doi: https://doi.org/10.1016/j.gaitpost.2018.11.029. URL http://www.sciencedirect.com/science/article/pii/S0966636218311913.

[26] Kinect for Windows, 2018. URL https://developer.microsoft.com/en-us/windows/kinect. Accessed 2019-04-21.

[27] S. Nömm, I. Masharov, A. Toomela, K. Medijainen, and P. Taba. Interpretable quantitative description of the digital clock drawing test for parkinson's disease modelling. In *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pages 1839–1844, Nov 2018. doi: 10.1109/ICARCV.2018.8581074.

[28] S. Nõmm, K. Bardõš, I. Mašarov, J. Kozhenkina, A. Toomela, and T. Toomsoo. Recognition and analysis of the contours drawn during the poppelreuter's test. In *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 170–175, Dec 2016. doi: 10.1109/ICMLA.2016.0036.

[29] S. Aich, M. Sain, J. Park, K. Choi, and H. Kim. A mixed classification approach for the prediction of parkinson's disease using nonlinear feature selection technique based on the voice recording. In *2017 International Conference on Inventive Computing and Informatics (ICICI)*, pages 959–962, Nov 2017. doi: 10.1109/ICICI.2017.8365279.

[30] S. Aich, K. Younga, K. L. Hui, A. A. Al-Absi, and M. Sain. A nonlinear decision tree based classification approach to predict the parkinson's disease using different

feature sets of voice data. In *2018 20th International Conference on Advanced Communication Technology (ICACT)*, pages 638–642, Feb 2018. doi: 10.23919/ICACT.2018.8323864.

[31] M. Su and K. Chuang. Dynamic feature selection for detecting parkinson's disease through voice signal. In *2015 IEEE MTT-S 2015 International Microwave Workshop Series on RF and Wireless Technologies for Biomedical and Healthcare Applications (IMWS-BIO)*, pages 148–149, Sep. 2015. doi: 10.1109/IMWS-BIO.2015.7303822.

[32] A. Gümüşçü, K. Karadağ, M. E. Tenekecı, and İ. B. Aydılek. Genetic algorithm based feature selection on diagnosis of parkinson disease via vocal analysis. In *2017 25th Signal Processing and Communications Applications Conference (SIU)*, pages 1–4, May 2017. doi: 10.1109/SIU.2017.7960384.

[33] A. Benba, A. Jilbab, A. Hammouch, and S. Sandabad. Voiceprints analysis using mfcc and svm for detecting patients with parkinson's disease. In *2015 International Conference on Electrical and Information Technologies (ICEIT)*, pages 300–304, March 2015. doi: 10.1109/EITech.2015.7163000.

[34] S. Chandrayan, A. Agarwal, M. Arif, and S. S. Sahu. Selection of dominant voice features for accurate detection of parkinson's disease. In *2017 Third International Conference on Biosignals, Images and Instrumentation (ICBSII)*, pages 1–4, March 2017. doi: 10.1109/ICBSII.2017.8082297.

[35] L. Jeancolas, H. Benali, B. Benkelfat, G. Mangone, J. Corvol, M. Vidailhet, S. Lehericy, and D. Petrovska-Delacrétaz. Automatic detection of early stages of parkinson's disease through acoustic voice analysis with mel-frequency cepstral coefficients. In *2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pages 1–6, May 2017. doi: 10.1109/ATSIP.2017.8075567.

[36] M. S. Wibawa, H. A. Nugroho, and N. A. Setiawan. Performance evaluation of combined feature selection and classification methods in diagnosing parkinson disease based on voice feature. In *2015 International Conference on Science in Information Technology (ICSITech)*, pages 126–131, Oct 2015. doi: 10.1109/ICSITech.2015.7407790.

[37] M. S. Islam, I. Parvez, , and P. Goswami. Performance comparison of heterogeneous classifiers for detection of parkinson's disease using voice disorder (dysphonia). In *2014 International Conference on Informatics, Electronics Vision (ICIEV)*, pages 1–7, May 2014. doi: 10.1109/ICIEV.2014.6850849.

[38] W. Caesarendra, M. Ariyanto, J. D. Setiawan, M. Arozi, and C. R. Chang. A pattern recognition method for stage classification of parkinson's disease utilizing voice features. In *2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES)*, pages 87–92, Dec 2014. doi: 10.1109/IECBES.2014.7047636.

[39] N. Lizbeth, C. J. Pérez, J. Martín, and F. Calle-Alonso. A new asymmetric link-based binary regression model to detect parkinson's disease by using replicated voice recordings. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 1182–1186, Sep. 2018. doi: 10.23919/EUSIPCO.2018.8553292.

[40] Max A. Little, Patrick E. McSharry, Stephen J. Roberts, Declan AE Costello, and Irene M. Moroz. Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *BioMedical Engineering OnLine*, 6(1):23, Jun 2007. ISSN 1475-925X. doi: 10.1186/1475-925X-6-23. URL `https://doi.org/10.1186/1475-925X-6-23`.

[41] Jundong Li, Kewei Cheng, Suhang Wang, Fred Morstatter, Trevino Robert, Jiliang Tang, and Huan Liu. Feature selection: A data perspective. *arXiv:1601.07996*, 2016.

[42] Hamed Sarbolandi, Damien Lefloch, and Andreas Kolb. Kinect range sensing: Structured-light versus time-of-flight kinect. *Computer Vision and Image Understanding*, 139:1 – 20, 2015. ISSN 1077-3142. doi: https://doi.org/10.1016/j.cviu.2015.05.006. URL `http://www.sciencedirect.com/science/article/pii/S1077314215001071`.

[43] M. Samir, E. Golkar, and A. A. A. Rahni. Comparison between the kinect™ v1 and kinect™ v2 for respiratory motion tracking. In *2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, pages 150–155, Oct 2015. doi: 10.1109/ICSIPA.2015.7412180.

[44] Jan-Joonas Bernstein. Context based registration and analysis of human motions. Bachelor's thesis, Tallinn University of Technology, 2017.

[45] A. Krajushkina, S. Nomm, A. Toomela, K. Medijainen, E. Tamm, M. Vaske, D. Uvarov, H. Kahar, M. Nugis, and P. Taba. Gait analysis based approach for parkinson's disease modeling with decision tree classifiers. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 3720–3725, Oct 2018. doi: 10.1109/SMC.2018.00630.

[46] Charu C. Aggarwal. *Data Mining: The Textbook*. Springer International Publishing, 2015.

A. Appendices

A.1 Statistical test

Below are presented all p-values for all the step-based features for both weeks one and two.

t-test Results

1st week			
Joint	Param	p-value	t-statistic
HipRight	Vm	0.000114	4.331885
HipLeft	Vm	0.000118	4.317425
SpineMid	Vm	0.000127	4.295227
Head	Vm	0.000138	4.265985
SpineBase	Vm	0.000146	4.248369
SpineShoulder	Vm	0.000302	4.000785
ElbowRight	Vm	0.000386	3.919672
HipRight	Tm	0.000591	3.768417
ShoulderLeft	Vm	0.000649	3.734449
HipLeft	Tm	0.000704	3.70669
SpineMid	Tm	0.00078	3.672206
SpineBase	Tm	0.000832	3.648851
Head	Tm	0.001056	3.563405
ShoulderLeft	Am	0.001359	3.514941
KneeRight	Vm	0.001364	3.519113
ElbowRight	Tm	0.001543	3.430895
SpineShoulder	Tm	0.001834	3.36415
ShoulderRight	Vm	0.001912	3.366591
KneeLeft	Vm	0.002594	3.242248
ElbowRight	E	0.003003	3.183453
SpineShoulder	E	0.003045	3.178296
ElbowLeft	Vm	0.003408	3.138019
KneeRight	Tm	0.003656	3.141354
ShoulderLeft	Tm	0.004071	3.06872
SpineMid	E	0.004285	3.049255
HipRight	E	0.004306	3.047474
ShoulderRight	E	0.004533	3.028032
KneeLeft	E	0.004563	3.025357
WristRight	E	0.004668	3.024271
HandRight	E	0.005008	3.000951
SpineBase	E	0.005283	2.969566
Head	E	0.006232	2.905912
WristRight	Vm	0.006328	2.900048
KneeLeft	Tm	0.006445	2.89707
AnkleLeft	E	0.007186	2.850232
ShoulderRight	Tm	0.007292	2.855015
KneeRight	E	0.007616	2.827512
AnkleRight	E	0.007626	2.82723
HipLeft	E	0.007685	2.824928
ShoulderLeft	E	0.010102	2.716117
ShoulderLeft	Jm	0.011202	2.685289
HandRight	Vm	0.011893	2.649771
SpineMid	Am	0.012128	2.687136
FootLeft	E	0.012544	2.628262
HipRight	Am	0.012579	2.635399
Head	Am	0.012984	2.664981
HipLeft	Am	0.014558	2.626781
SpineBase	Am	0.015099	2.588378
SpineMid	Jm	0.015613	2.585066
SpineBase	Jm	0.016906	2.547432
FootRight	E	0.017556	2.492583

2nd week			
Joint	Param	p-value	t-statistic
ShoulderRight	Tm	9E-06	5.122567
ElbowRight	Tm	1.1E-05	5.06675
ShoulderRight	Vm	1.1E-05	5.051259
ElbowRight	Vm	1.1E-05	5.049883
KneeRight	E	1.4E-05	4.986571
ShoulderRight	E	1.6E-05	4.944651
SpineShoulder	Tm	1.6E-05	4.936453
HipRight	Tm	1.7E-05	4.917614
HipRight	Vm	1.8E-05	4.904116
ElbowRight	E	1.9E-05	4.89185
HipRight	E	2.1E-05	4.85827
SpineShoulder	Vm	2.1E-05	4.845427
SpineBase	Tm	2.8E-05	4.765217
SpineMid	Tm	2.8E-05	4.76305
SpineBase	Vm	2.9E-05	4.747894
SpineMid	Vm	3E-05	4.742782
SpineMid	E	3.4E-05	4.692902
SpineBase	E	3.5E-05	4.686157
SpineShoulder	E	4E-05	4.645735
HipLeft	Tm	4.9E-05	4.57793
HipLeft	Vm	5.3E-05	4.551003
HipLeft	E	6.2E-05	4.502647
FootRight	E	6.5E-05	4.485094
WristRight	E	6.7E-05	4.478054
AnkleRight	E	7E-05	4.460577
HandRight	E	8.9E-05	4.385519
AnkleLeft	E	9E-05	4.380694
Head	E	9.2E-05	4.374511
Head	Tm	9.6E-05	4.358204
FootLeft	E	9.7E-05	4.354813
Head	Vm	0.000108	4.318978
ShoulderLeft	Tm	0.000117	4.292722
ShoulderLeft	Vm	0.00012	4.284801
KneeRight	Tm	0.000134	4.248813
KneeLeft	E	0.000148	4.216791
KneeRight	Vm	0.000174	4.162549
ShoulderLeft	E	0.000187	4.138537
WristRight	Tm	0.000195	4.125027
HipLeft	Am	0.000206	4.106511
WristRight	Vm	0.000266	4.020375
ShoulderLeft	Am	0.00077	3.656913
KneeLeft	Tm	0.000813	3.638171
ElbowLeft	E	0.000815	3.637056
KneeLeft	Vm	0.001035	3.553553
ElbowLeft	Tm	0.001045	3.550092
SpineBase	Am	0.001207	3.499416
HandRight	Tm	0.001288	3.476575
SpineMid	Am	0.001343	3.461516
ElbowLeft	Vm	0.001354	3.458707
SpineMid	Jm	0.001531	3.414885
SpineBase	Jm	0.00163	3.392641

t-test Results

ShoulderRight	Am	0.018792	2.474641	HandRight	Vm	0.001796	3.35781
WristRight	Tm	0.019538	2.444425	HipRight	Am	0.00189	3.339467
HipRight	Jm	0.02069	2.433611	HipLeft	Jm	0.002908	3.182649
ElbowLeft	Tm	0.021504	2.404268	ShoulderLeft	Jm	0.003383	3.126716
AnkleLeft	Vm	0.021872	2.406202	HipRight	Jm	0.00546	2.946832
ElbowRight	Am	0.022224	2.410724	WristRight	E/Am	0.006231	2.896306
AnkleRight	Vm	0.022943	2.42275	WristLeft	E	0.007538	2.82265
Head	Jm	0.023331	2.40635	WristLeft	Tm	0.007755	2.811581
ElbowRight	Jm	0.027606	2.334201	ShoulderRight	Am	0.009084	2.74959
ElbowLeft	E	0.032874	2.219143	Head	Am	0.010072	2.708706
HandRight	Tm	0.033269	2.213777	WristLeft	Vm	0.011953	2.640238
HipLeft	Jm	0.03926	2.184565	SpineShoulder	Am	0.013745	2.583711
ElbowLeft	Am	0.041341	2.116017	HandLeft	E	0.014213	2.570053
AnkleRight	Tm	0.042221	2.144033	AnkleRight	Tm	0.018506	2.461095
HandRight	Am	0.047712	2.086336	FootRight	Tm	0.021987	2.38855
AnkleLeft	Tm	0.05123	2.023598	ShoulderRight	Jm	0.026635	2.306416
FootRight	Vm	0.057113	2.000226	AnkleLeft	Tm	0.026877	2.302494
FootLeft	Vm	0.057549	1.971436	FootLeft	Tm	0.027347	2.294991
HipRight	E/Am	0.070281	-1.869171	HandLeft	Tm	0.030595	2.246034
SpineMid	E/Am	0.070366	-1.865617	Head	Jm	0.031171	2.23785
ShoulderRight	E/Am	0.07183	-1.855096	ElbowRight	Jm	0.032118	2.224676
ElbowLeft	Jm	0.08465	1.773257	KneeRight	Am	0.034043	2.19892
WristRight	Am	0.085544	1.78572	AnkleRight	Vm	0.035042	2.186058
HandRight	Jm	0.090149	1.756024	AnkleLeft	Vm	0.042242	2.101932
KneeRight	Am	0.090238	1.769731	FootLeft	Vm	0.045292	2.070065
WristLeft	Vm	0.105019	1.662933	HandLeft	Vm	0.045549	2.067475
KneeLeft	Jm	0.108916	1.643922	HandRight	E/Am	0.046662	2.056368
FootRight	Tm	0.110144	1.660859	FootRight	Vm	0.047082	2.052234
KneeLeft	Am	0.126355	1.567452	AnkleRight	E/Am	0.048995	2.033818
FootLeft	Tm	0.126553	1.569565	AnkleLeft	E/Am	0.05743	1.959414
SpineBase	E/Am	0.143375	-1.498453	ElbowLeft	Am	0.058239	1.952787
KneeRight	Jm	0.143954	1.511049	ElbowRight	Am	0.077479	1.814642
ShoulderLeft	E/Am	0.146002	-1.498101	FootRight	E/Tm	0.078048	1.811028
WristLeft	E	0.149661	1.472191	FootRight	E/Am	0.084001	1.774451
SpineShoulder	Am	0.157841	1.453684	SpineBase	E/Am	0.120831	-1.586864
HandLeft	Vm	0.158568	1.440154	ShoulderLeft	E/Am	0.121981	-1.581815
ShoulderRight	Jm	0.171687	1.398635	KneeRight	Jm	0.122846	1.578042
WristRight	Jm	0.171844	1.39983	ElbowLeft	Jm	0.124872	1.569286
ElbowLeft	E/Am	0.173151	-1.390532	KneeLeft	E/Am	0.125439	1.566858
ShoulderLeft	t	0.187238	-1.346475	SpineShoulder	Jm	0.128557	1.553654
SpineBase	t	0.187238	-1.346475	ShoulderLeft	E/Tm	0.138052	-1.514982
SpineShoulder	t	0.187238	-1.346475	FootLeft	E/Am	0.138738	1.512268
ShoulderRight	t	0.187238	-1.346475	ElbowRight	E/Am	0.157755	1.441072
WristRight	t	0.187238	-1.346475	WristLeft	Jm	0.158951	1.436824
WristLeft	t	0.187238	-1.346475	WristRight	Jm	0.168595	1.403474
SpineMid	t	0.187238	-1.346475	FootLeft	E/Tm	0.172724	1.389649
AnkleLeft	t	0.187238	-1.346475	KneeLeft	Am	0.182772	1.35705
ElbowRight	t	0.187238	-1.346475	SpineMid	E/Am	0.191547	-1.329693
KneeRight	t	0.187238	-1.346475	WristLeft	Am	0.217485	1.254048
ElbowLeft	t	0.187238	-1.346475	HandRight	E/Tm	0.224412	1.235001
HandRight	t	0.187238	-1.346475	SpineBase	E/Tm	0.232598	-1.21305
HipLeft	t	0.187238	-1.346475	SpineMid	E/Tm	0.243377	-1.184997
HandLeft	t	0.187238	-1.346475	HipLeft	E/Am	0.250104	-1.167947

t-test Results

AnkleRight	t	0.187238	-1.346475	HipLeft	E/Tm	0.259712	-1.144164
Head	t	0.187238	-1.346475	KneeLeft	Jm	0.269915	1.11959
HipRight	t	0.187238	-1.346475	HandLeft	E/Am	0.274947	1.107715
KneeLeft	t	0.187238	-1.346475	AnkleLeft	E/Tm	0.302088	1.046178
FootRight	t	0.187238	-1.346475	AnkleRight	E/Tm	0.315865	1.016404
FootLeft	t	0.187238	-1.346475	ShoulderRight	E/Tm	0.338406	-0.969542
Head	E/Tm	0.188086	-1.362529	HandRight	Jm	0.349356	0.947529
HandLeft	Am	0.203669	1.296146	WristRight	E/Tm	0.351622	0.943032
HandLeft	E	0.228989	1.223789	HandLeft	Jm	0.412858	0.827968
HipLeft	E/Tm	0.229075	-1.245365	SpineShoulder	t	0.453046	-0.758136
AnkleLeft	Jm	0.233945	1.211517	WristLeft	t	0.453046	-0.758136
HipRight	E/Tm	0.25172	-1.183503	SpineMid	t	0.453046	-0.758136
SpineBase	E/Tm	0.272205	-1.132138	WristRight	t	0.453046	-0.758136
AnkleRight	Jm	0.277905	1.109927	SpineBase	t	0.453046	-0.758136
SpineMid	E/Tm	0.279637	-1.114109	AnkleLeft	t	0.453046	-0.758136
FootRight	E/Tm	0.279779	1.10193	KneeRight	t	0.453046	-0.758136
AnkleRight	Am	0.282854	1.099225	ShoulderRight	t	0.453046	-0.758136
AnkleLeft	Am	0.293343	1.06961	ElbowRight	t	0.453046	-0.758136
WristLeft	Tm	0.296088	1.060265	HipRight	t	0.453046	-0.758136
HipLeft	E/Am	0.312884	-1.027432	FootLeft	t	0.453046	-0.758136
HandLeft	Jm	0.31972	1.009003	HipLeft	t	0.453046	-0.758136
ShoulderLeft	E/Tm	0.331162	-0.993744	Head	t	0.453046	-0.758136
AnkleRight	E/Am	0.347433	0.952737	KneeLeft	t	0.453046	-0.758136
FootRight	E/Am	0.355316	0.938118	ShoulderLeft	t	0.453046	-0.758136
ElbowRight	E/Tm	0.358772	-0.935797	HandRight	t	0.453046	-0.758136
HandLeft	Tm	0.381507	0.886045	HandLeft	t	0.453046	-0.758136
FootLeft	E/Tm	0.409137	0.836167	FootRight	t	0.453046	-0.758136
Head	E/Am	0.410358	-0.833337	ElbowLeft	t	0.453046	-0.758136
WristLeft	Am	0.424207	0.808498	AnkleRight	t	0.453046	-0.758136
FootLeft	E/Am	0.442147	0.777153	WristLeft	E/Am	0.480343	0.71277
FootLeft	Jm	0.467152	0.734988	FootLeft	Jm	0.483525	0.707578
KneeRight	E/Tm	0.498842	-0.685899	HandRight	Am	0.495695	0.687898
ElbowRight	E/Am	0.502935	-0.676694	KneeRight	E/Tm	0.508678	0.667197
FootRight	Jm	0.525961	0.642391	HipRight	E/Tm	0.539638	-0.618957
HandLeft	E/Am	0.542075	-0.615587	KneeLeft	E/Tm	0.54642	0.608586
ElbowLeft	E/Tm	0.555175	-0.596495	FootLeft	Am	0.554502	0.596314
FootRight	Am	0.557057	0.595215	HipRight	E/Am	0.556963	-0.592594
FootLeft	Am	0.579335	0.559853	ElbowLeft	E/Am	0.563295	0.583064
AnkleRight	E/Tm	0.604001	0.524623	Head	E/Tm	0.570521	0.572253
WristLeft	E/Tm	0.623551	0.495478	HandLeft	Am	0.574414	0.566457
WristLeft	Jm	0.634058	0.480108	AnkleLeft	Jm	0.586572	0.548477
WristRight	E/Tm	0.644812	0.465	KneeRight	E/Am	0.631551	0.483449
HandLeft	E/Tm	0.686888	0.406392	AnkleLeft	Am	0.665615	0.435568
HandRight	E/Tm	0.687657	0.405367	ElbowRight	E/Tm	0.668696	-0.431288
KneeLeft	E/Tm	0.737817	0.337664	ElbowLeft	E/Tm	0.684983	0.408798
KneeLeft	E/Am	0.764118	-0.30236	Head	E/Am	0.6981	0.390838
ShoulderRight	E/Tm	0.778705	-0.283167	AnkleRight	Am	0.71017	0.374424
HandRight	E/Am	0.79978	-0.255555	SpineShoulder	E/Am	0.729482	0.348374
WristLeft	E/Am	0.824802	0.223006	HandLeft	E/Tm	0.734192	0.342059
WristRight	E/Am	0.824946	-0.222881	WristRight	Am	0.75088	0.319789
SpineShoulder	Jm	0.836152	0.208472	FootRight	Am	0.786738	0.272467
AnkleLeft	E/Tm	0.846966	0.19463	SpineShoulder	E/Tm	0.834005	-0.211013
SpineShoulder	E/Tm	0.858143	-0.18048	ShoulderRight	E/Am	0.83621	0.208168

t-test Results

AnkleLeft	E/Am	0.872929	0.161144	AnkleRight	Jm	0.897591	0.129569
KneeRight	E/Am	0.873907	-0.159836	FootRight	Jm	0.988474	0.014542
SpineShoulder	E/Am	0.919323	-0.102022	WristLeft	E/Tm	0.99777	0.002813

A.2 Fisher Scores for Step-based MM

Below are presented the full results for Fisher scoring algorithm for both weeks one and two.

Fisher Scores for step-based MM

1st week		
Joint	Parameter	Fisher Score
HipRight	Vm	0.512484
HipLeft	Vm	0.510796
SpineMid	Vm	0.503264
Head	Vm	0.500898
SpineBase	Vm	0.49291
SpineShoulder	Vm	0.442011
ElbowRight	Vm	0.426837
HipRight	Tm	0.387325
ShoulderLeft	Vm	0.383611
HipLeft	Tm	0.375525
SpineMid	Tm	0.367021
SpineBase	Tm	0.362869
KneeRight	Vm	0.355398
ShoulderLeft	Am	0.353612
Head	Tm	0.346816
ElbowRight	Tm	0.32735
SpineShoulder	Tm	0.310331
ShoulderRight	Vm	0.302826
KneeLeft	Vm	0.293221
KneeRight	Tm	0.282561
ElbowRight	E	0.279876
SpineShoulder	E	0.27582
ElbowLeft	Vm	0.273426
ShoulderLeft	Tm	0.258273
WristRight	E	0.25647
SpineMid	E	0.255603
HipRight	E	0.25531
HandRight	E	0.253739
KneeLeft	E	0.251511
ShoulderRight	E	0.251134
SpineBase	E	0.242906
KneeLeft	Tm	0.234392
WristRight	Vm	0.23194
Head	E	0.230981
AnkleLeft	E	0.222815
HipLeft	E	0.220955
KneeRight	E	0.219018
AnkleRight	E	0.218412
ShoulderRight	Tm	0.217997
SpineMid	Am	0.21102
Head	Am	0.208543
HipLeft	Am	0.204114
ShoulderLeft	Jm	0.204105
ShoulderLeft	E	0.204045
SpineMid	Jm	0.196187
HipRight	Am	0.195892
SpineBase	Am	0.195011
HandRight	Vm	0.193333
SpineBase	Jm	0.190083
FootLeft	E	0.188836
AnkleRight	Vm	0.173375

2nd week		
Joint	Parameter	Fisher Score
ShoulderRight	Tm	0.690544
ElbowRight	Tm	0.675578
ShoulderRight	Vm	0.671453
ElbowRight	Vm	0.671087
KneeRight	E	0.654366
ShoulderRight	E	0.64341
SpineShoulder	Tm	0.641278
HipRight	Tm	0.636393
HipRight	Vm	0.632904
ElbowRight	E	0.629742
HipRight	E	0.621126
SpineShoulder	Vm	0.617846
SpineBase	Tm	0.59756
SpineMid	Tm	0.597017
SpineBase	Vm	0.593224
SpineMid	Vm	0.591947
SpineMid	E	0.579561
SpineBase	E	0.577896
SpineShoulder	E	0.56797
HipLeft	Tm	0.551512
HipLeft	Vm	0.545043
HipLeft	E	0.533522
FootRight	E	0.52937
WristRight	E	0.52771
AnkleRight	E	0.523599
HandRight	E	0.506126
AnkleLeft	E	0.505013
Head	E	0.503588
Head	Tm	0.499841
FootLeft	E	0.499063
Head	Vm	0.490883
ShoulderLeft	Tm	0.484933
ShoulderLeft	Vm	0.483145
KneeRight	Tm	0.475064
KneeLeft	E	0.46793
KneeRight	Vm	0.455969
ShoulderLeft	E	0.450723
WristRight	Tm	0.447785
HipLeft	Am	0.443774
WristRight	Vm	0.425353
ShoulderLeft	Am	0.351921
KneeLeft	Tm	0.348323
ElbowLeft	E	0.34811
KneeLeft	Vm	0.332309
ElbowLeft	Tm	0.331662
SpineBase	Am	0.322261
HandRight	Tm	0.318068
SpineMid	Am	0.315318
ElbowLeft	Vm	0.314807
SpineMid	Jm	0.30688
SpineBase	Jm	0.302895

Fisher Scores for step-based MM

Head	Jm	0.169763	HandRight	Vm	0.296708
HipRight	Jm	0.16889	HipRight	Am	0.293475
FootRight	E	0.167625	HipLeft	Jm	0.266559
ElbowRight	Am	0.167466	ShoulderLeft	Jm	0.257272
WristRight	Tm	0.164382	HipRight	Jm	0.228522
AnkleLeft	Vm	0.164223	WristRight	E/Am	0.220752
ShoulderRight	Am	0.16215	WristLeft	E	0.209667
ElbowRight	Jm	0.160287	WristLeft	Tm	0.208026
ElbowLeft	Tm	0.159944	ShoulderRight	Am	0.198954
HipLeft	Jm	0.142319	Head	Am	0.193081
AnkleRight	Tm	0.136313	WristLeft	Vm	0.183444
ElbowLeft	E	0.134743	SpineShoulder	Am	0.175673
HandRight	Tm	0.13425	HandLeft	E	0.17382
HandRight	Am	0.129244	AnkleRight	Tm	0.159394
ElbowLeft	Am	0.123539	FootRight	Tm	0.150136
FootRight	Vm	0.119091	ShoulderRight	Jm	0.139988
AnkleLeft	Tm	0.116342	AnkleLeft	Tm	0.139513
FootLeft	Vm	0.111286	FootLeft	Tm	0.138605
HipRight	E/Am	0.09861	HandLeft	Tm	0.132754
SpineMid	E/Am	0.096736	Head	Jm	0.131789
ShoulderRight	E/Am	0.093753	ElbowRight	Jm	0.130242
KneeRight	Am	0.093693	KneeRight	Am	0.127243
WristRight	Am	0.093493	AnkleRight	Vm	0.125759
HandRight	Jm	0.089946	AnkleLeft	Vm	0.116266
ElbowLeft	Jm	0.08615	FootLeft	Vm	0.112768
FootRight	Tm	0.082237	HandLeft	Vm	0.112486
WristLeft	Vm	0.076193	HandRight	E/Am	0.11128
KneeLeft	Jm	0.074551	FootRight	Vm	0.110833
FootLeft	Tm	0.070543	AnkleRight	E/Am	0.108853
KneeLeft	Am	0.069401	AnkleLeft	E/Am	0.101034
KneeRight	Jm	0.067901	ElbowLeft	Am	0.100352
ShoulderLeft	E/Am	0.065895	ElbowRight	Am	0.086656
SpineBase	E/Am	0.063497	FootRight	E/Tm	0.086311
SpineShoulder	Am	0.062038	FootRight	E/Am	0.08286
WristLeft	E	0.059393	SpineBase	E/Am	0.066267
HandLeft	Vm	0.057621	ShoulderLeft	E/Am	0.065846
WristRight	Jm	0.056525	KneeRight	Jm	0.065532
Head	E/Tm	0.056327	ElbowLeft	Jm	0.064807
ElbowLeft	E/Am	0.052136	KneeLeft	E/Am	0.064606
ShoulderRight	Jm	0.051623	SpineShoulder	Jm	0.063522
WristLeft	t	0.04831	ShoulderLeft	E/Tm	0.060399
ShoulderLeft	t	0.04831	FootLeft	E/Am	0.060183
KneeRight	t	0.04831	ElbowRight	E/Am	0.05465
SpineMid	t	0.04831	WristLeft	Jm	0.054328
ShoulderRight	t	0.04831	WristRight	Jm	0.051835
SpineShoulder	t	0.04831	FootLeft	E/Tm	0.050819
SpineBase	t	0.04831	KneeLeft	Am	0.048463
WristRight	t	0.04831	SpineMid	E/Am	0.046529
FootLeft	t	0.04831	WristLeft	Am	0.041385
AnkleRight	t	0.04831	HandRight	E/Tm	0.040138
FootRight	t	0.04831	SpineBase	E/Tm	0.038723
HandLeft	t	0.04831	SpineMid	E/Tm	0.036953
ElbowRight	t	0.04831	HipLeft	E/Am	0.035897

Fisher Scores for step-based MM

ElbowLeft	t	0.04831	HipLeft	E/Tm	0.03445
HandRight	t	0.04831	KneeLeft	Jm	0.032986
Head	t	0.04831	HandLeft	E/Am	0.03229
HipLeft	t	0.04831	AnkleLeft	E/Tm	0.028802
HipRight	t	0.04831	AnkleRight	E/Tm	0.027186
AnkleLeft	t	0.04831	ShoulderRight	E/Tm	0.024737
KneeLeft	t	0.04831	HandRight	Jm	0.023627
HipLeft	E/Tm	0.047727	WristRight	E/Tm	0.023403
HandLeft	Am	0.047374	HandLeft	Jm	0.01804
HipRight	E/Tm	0.042969	SpineBase	t	0.015126
HandLeft	E	0.041256	ShoulderRight	t	0.015126
AnkleLeft	Jm	0.041239	WristLeft	t	0.015126
SpineBase	E/Tm	0.039336	SpineMid	t	0.015126
SpineMid	E/Tm	0.038085	SpineShoulder	t	0.015126
AnkleRight	Jm	0.036543	WristRight	t	0.015126
AnkleRight	Am	0.036004	AnkleRight	t	0.015126
FootRight	E/Tm	0.035311	HandLeft	t	0.015126
AnkleLeft	Am	0.033001	KneeRight	t	0.015126
WristLeft	Tm	0.030921	FootRight	t	0.015126
HipLeft	E/Am	0.030679	ShoulderLeft	t	0.015126
ShoulderLeft	E/Tm	0.029648	KneeLeft	t	0.015126
HandLeft	Jm	0.028082	HipRight	t	0.015126
ElbowRight	E/Tm	0.026038	AnkleLeft	t	0.015126
AnkleRight	E/Am	0.025572	ElbowRight	t	0.015126
FootRight	E/Am	0.025178	HipLeft	t	0.015126
HandLeft	Tm	0.021728	Head	t	0.015126
FootLeft	E/Tm	0.019877	FootLeft	t	0.015126
Head	E/Am	0.019474	ElbowLeft	t	0.015126
WristLeft	Am	0.018205	HandRight	t	0.015126
FootLeft	E/Am	0.016615	WristLeft	E/Am	0.013369
FootLeft	Jm	0.014985	FootLeft	Jm	0.013175
KneeRight	E/Tm	0.013834	HandRight	Am	0.012453
ElbowRight	E/Am	0.012509	KneeRight	E/Tm	0.011715
FootRight	Jm	0.012054	HipRight	E/Tm	0.010082
HandLeft	E/Am	0.010512	KneeLeft	E/Tm	0.009747
FootRight	Am	0.010472	FootLeft	Am	0.009358
ElbowLeft	E/Tm	0.010208	HipRight	E/Am	0.009241
FootLeft	Am	0.00888	ElbowLeft	E/Am	0.008946
AnkleRight	E/Tm	0.008022	Head	E/Tm	0.008618
WristLeft	E/Tm	0.006528	HandLeft	Am	0.008444
WristLeft	Jm	0.006305	AnkleLeft	Jm	0.007917
WristRight	E/Tm	0.005831	KneeRight	E/Am	0.006151
HandLeft	E/Tm	0.004483	AnkleLeft	Am	0.004993
HandRight	E/Tm	0.004443	ElbowRight	E/Tm	0.004895
KneeLeft	E/Tm	0.003252	ElbowLeft	E/Tm	0.004398
KneeLeft	E/Am	0.002506	Head	E/Am	0.00402
ShoulderRight	E/Tm	0.002236	AnkleRight	Am	0.003689
HandRight	E/Am	0.001821	SpineShoulder	E/Am	0.003194
WristRight	E/Am	0.001396	HandLeft	E/Tm	0.003079
WristLeft	E/Am	0.001379	WristRight	Am	0.002691
SpineShoulder	Jm	0.001235	FootRight	Am	0.001954
AnkleLeft	E/Tm	0.001089	SpineShoulder	E/Tm	0.001172
SpineShoulder	E/Tm	0.000955	ShoulderRight	E/Am	0.00114

Fisher Scores for step-based MM

AnkleLeft	E/Am	0.000731	AnkleRight	Jm	0.000442
KneeRight	E/Am	0.000706	FootRight	Jm	6E-06
SpineShoulder	E/Am	0.000291	WristLeft	E/Tm	0

A.3 Fisher Scores for Walking Forward Phase MM

Fisher Scores for Walk Forward Phase

1st week		
Joint	Parameter	Fisher Score
WristRight	t	0.176002
SpineShoulder	t	0.176002
HipLeft	t	0.176002
HipRight	t	0.176002
HandRight	t	0.176002
KneeLeft	t	0.176002
KneeRight	t	0.176002
HandLeft	t	0.176002
FootRight	t	0.176002
FootLeft	t	0.176002
ShoulderLeft	t	0.176002
ElbowRight	t	0.176002
ShoulderRight	t	0.176002
SpineBase	t	0.176002
SpineMid	t	0.176002
ElbowLeft	t	0.176002
Head	t	0.176002
WristLeft	t	0.176002
AnkleRight	t	0.176002
AnkleLeft	t	0.176002
HipRight	E/Am	0.106269
ShoulderLeft	E/Am	0.072126
FootRight	Tm	0.070023
FootRight	Am	0.064041
SpineMid	E/Am	0.063669
FootRight	Jm	0.063495
HandLeft	Tm	0.061694
SpineBase	E/Am	0.060947
FootLeft	E/Tm	0.056698
WristLeft	Tm	0.054267
FootRight	E/Am	0.051565
FootRight	E	0.050686
ShoulderRight	E/Am	0.050152
HipLeft	E/Am	0.049479
FootLeft	E/Am	0.045714
SpineShoulder	Jm	0.043878
AnkleRight	E	0.043717
FootRight	Vm	0.043613
AnkleRight	E/Am	0.040869
WristLeft	Jm	0.04059
FootLeft	Am	0.039875
HandLeft	Vm	0.03978
FootRight	E/Tm	0.039031
Head	E	0.037359
WristLeft	E	0.032595
WristRight	Tm	0.032107
HandRight	Tm	0.031965
AnkleLeft	E/Am	0.03153
FootLeft	Jm	0.031461
WristLeft	Am	0.031228
FootLeft	Tm	0.030227

2nd week		
Joint	Parameter	Fisher Score
WristRight	E/Am	0.282255
WristRight	t	0.25302
AnkleRight	t	0.25302
FootLeft	t	0.25302
HandLeft	t	0.25302
ElbowRight	t	0.25302
HandRight	t	0.25302
Head	t	0.25302
HipLeft	t	0.25302
FootRight	t	0.25302
HipRight	t	0.25302
KneeLeft	t	0.25302
KneeRight	t	0.25302
ShoulderLeft	t	0.25302
ShoulderRight	t	0.25302
ElbowLeft	t	0.25302
AnkleLeft	t	0.25302
SpineBase	t	0.25302
SpineMid	t	0.25302
SpineShoulder	t	0.25302
WristLeft	t	0.25302
AnkleLeft	E/Am	0.242143
FootLeft	E/Am	0.234233
WristRight	Am	0.184063
HandRight	E/Am	0.170394
FootRight	E/Am	0.168738
HandLeft	Am	0.158427
FootLeft	E/Tm	0.154286
HandLeft	E/Am	0.151061
HipLeft	E/Am	0.147527
HandRight	Am	0.145014
SpineBase	E/Am	0.140421
WristLeft	E/Am	0.137421
AnkleRight	E/Am	0.134943
SpineMid	E/Am	0.131786
FootLeft	Jm	0.13156
HandLeft	Jm	0.13043
FootLeft	Am	0.128749
WristLeft	Am	0.126922
HandRight	Jm	0.126693
FootRight	E/Tm	0.122912
FootRight	Jm	0.122245
HipLeft	Am	0.118269
FootRight	Am	0.116859
SpineBase	Am	0.109431
WristRight	Jm	0.105353
HipRight	Am	0.103284
FootLeft	Vm	0.103268
WristLeft	Jm	0.102212
HipRight	E/Am	0.100049
FootLeft	Tm	0.096472

Fisher Scores for Walk Forward Phase

ElbowLeft	E	0.029948	SpineMid	Am	0.095395
Head	E/Am	0.029864	HandLeft	Vm	0.092808
WristLeft	Vm	0.029522	FootRight	Vm	0.089057
HandLeft	E	0.0287	AnkleLeft	Am	0.087028
HandLeft	Jm	0.02781	FootRight	Tm	0.083107
SpineShoulder	Am	0.025004	HandLeft	Tm	0.081472
ShoulderLeft	E	0.024598	KneeRight	E/Am	0.071147
SpineShoulder	E	0.023329	AnkleLeft	Jm	0.069848
HandLeft	Am	0.022538	ElbowLeft	Vm	0.069521
SpineShoulder	Tm	0.020993	AnkleLeft	E/Tm	0.069439
WristRight	E/Am	0.019475	WristLeft	Vm	0.067017
ShoulderRight	Tm	0.019389	Head	Vm	0.061607
ShoulderRight	E	0.018518	WristLeft	Tm	0.057977
ShoulderLeft	Am	0.017617	ShoulderLeft	E/Am	0.057897
ElbowLeft	Tm	0.017231	AnkleLeft	Vm	0.055887
ShoulderLeft	Tm	0.016679	KneeLeft	E/Am	0.055527
Head	Tm	0.016429	Head	Jm	0.054389
HandLeft	E/Tm	0.014994	AnkleRight	Am	0.053487
AnkleLeft	E/Tm	0.014472	ElbowLeft	Tm	0.053356
KneeLeft	E/Am	0.014168	SpineShoulder	Jm	0.051049
Head	E/Tm	0.013963	ElbowLeft	Jm	0.048864
HipRight	Am	0.013792	AnkleLeft	Tm	0.047841
WristRight	Jm	0.013374	ShoulderLeft	Vm	0.047328
ElbowRight	E	0.013004	AnkleRight	Jm	0.047006
SpineMid	Tm	0.012983	KneeRight	Jm	0.045632
FootLeft	Vm	0.01294	ElbowLeft	E	0.043991
AnkleRight	Tm	0.012774	HandRight	Vm	0.043925
AnkleRight	Am	0.011881	Head	Tm	0.043594
SpineMid	E	0.011505	ShoulderLeft	E	0.041716
WristRight	E	0.011424	Head	E	0.039082
WristRight	Am	0.011267	KneeRight	Am	0.039021
AnkleRight	Jm	0.011219	KneeRight	E/Tm	0.038337
WristLeft	E/Am	0.010964	WristLeft	E	0.037602
HipLeft	E/Tm	0.010862	SpineBase	Jm	0.036801
HandRight	Vm	0.01067	AnkleRight	E/Tm	0.036366
SpineBase	Tm	0.010449	SpineShoulder	Vm	0.034695
KneeLeft	Tm	0.010159	ElbowRight	Jm	0.034079
SpineBase	E	0.010132	ElbowLeft	Am	0.033769
KneeRight	E	0.009855	ElbowRight	E/Am	0.033209
WristLeft	E/Tm	0.009715	HipRight	Jm	0.032578
WristRight	E/Tm	0.00948	HandLeft	E	0.031836
ElbowLeft	E/Am	0.009354	ShoulderLeft	Tm	0.031113
WristRight	Vm	0.009299	AnkleRight	Vm	0.031078
HipLeft	E	0.009089	AnkleRight	Tm	0.027194
HipRight	E/Tm	0.00891	Head	E/Tm	0.026555
AnkleLeft	Am	0.008659	ElbowRight	Am	0.026263
HandRight	E	0.008193	HandRight	Tm	0.025925
KneeRight	E/Tm	0.008	HandRight	E/Tm	0.025348
HipRight	E	0.007719	KneeRight	Vm	0.024627
SpineShoulder	E/Am	0.00732	WristRight	Vm	0.024561
ElbowRight	E/Am	0.007181	HandLeft	E/Tm	0.024029
SpineBase	E/Tm	0.007065	SpineMid	Vm	0.023276
HipRight	Tm	0.006976	SpineShoulder	Tm	0.023117

Fisher Scores for Walk Forward Phase

HandRight	E/Tm	0.006943	SpineShoulder	E	0.022302
AnkleLeft	Tm	0.006892	HipLeft	Jm	0.021175
HipLeft	Tm	0.00657	ShoulderLeft	Am	0.020906
SpineMid	Jm	0.006275	SpineMid	Jm	0.019812
ShoulderLeft	E/Tm	0.006172	KneeLeft	Am	0.018627
ElbowRight	E/Tm	0.005692	HipLeft	Vm	0.01805
HandRight	Jm	0.005628	KneeLeft	Jm	0.017881
HandLeft	E/Am	0.005607	ElbowLeft	E/Tm	0.016322
AnkleLeft	Jm	0.005472	SpineBase	Vm	0.015856
AnkleRight	E/Tm	0.005225	KneeRight	Tm	0.015652
SpineMid	E/Tm	0.005196	KneeLeft	Vm	0.014426
ElbowLeft	E/Tm	0.005185	ElbowRight	Vm	0.013589
HipLeft	Am	0.005108	SpineMid	E	0.013455
SpineBase	Am	0.004953	KneeLeft	E/Tm	0.013136
ShoulderRight	Jm	0.00477	SpineMid	Tm	0.011331
ElbowRight	Jm	0.00474	WristRight	Tm	0.010875
ElbowRight	Am	0.004644	SpineShoulder	Am	0.010687
HandRight	Am	0.004284	HipLeft	E	0.010397
ElbowRight	Vm	0.003844	HipRight	Vm	0.009354
ElbowLeft	Jm	0.003691	SpineBase	E	0.00856
AnkleRight	Vm	0.003623	KneeLeft	Tm	0.008324
KneeLeft	E	0.003435	ShoulderRight	E/Tm	0.008228
ElbowRight	Tm	0.003424	WristRight	E/Tm	0.008021
SpineBase	Jm	0.003423	WristLeft	E/Tm	0.007975
ShoulderRight	Am	0.002753	SpineBase	E/Tm	0.007581
KneeRight	Vm	0.002681	ElbowRight	E	0.007516
KneeRight	Am	0.00245	Head	Am	0.007381
SpineMid	Am	0.002295	HipLeft	Tm	0.007307
HipRight	Jm	0.002248	ShoulderLeft	E/Tm	0.006645
SpineShoulder	Vm	0.001568	SpineShoulder	E/Tm	0.006491
KneeRight	E/Am	0.001385	HipLeft	E/Tm	0.006185
ShoulderLeft	Jm	0.001302	SpineBase	Tm	0.005991
HipLeft	Jm	0.001287	ShoulderRight	Vm	0.00578
KneeRight	Tm	0.001193	ElbowLeft	E/Am	0.005696
HipLeft	Vm	0.001111	ShoulderRight	E	0.005429
FootLeft	E	0.000965	HandRight	E	0.004985
ShoulderRight	E/Tm	0.000857	ElbowRight	Tm	0.004869
AnkleLeft	E	0.000841	WristRight	E	0.004836
KneeLeft	E/Tm	0.000822	ShoulderRight	Jm	0.004447
HipRight	Vm	0.000821	HipRight	E/Tm	0.00438
SpineShoulder	E/Tm	0.000797	Head	E/Am	0.003743
KneeLeft	Am	0.000736	HipRight	E	0.003423
AnkleLeft	Vm	0.00051	KneeLeft	E	0.003274
KneeLeft	Jm	0.000477	ElbowRight	E/Tm	0.003205
ElbowLeft	Vm	0.000353	FootLeft	E	0.002214
HandRight	E/Am	0.000348	HipRight	Tm	0.002123
ElbowLeft	Am	0.000339	SpineMid	E/Tm	0.001473
KneeLeft	Vm	0.000306	AnkleLeft	E	0.001338
Head	Vm	0.000193	ShoulderRight	Tm	0.001074
KneeRight	Jm	0.000183	ShoulderRight	E/Am	0.001054
ShoulderRight	Vm	0.000118	ShoulderRight	Am	0.001033
SpineBase	Vm	0.000105	AnkleRight	E	0.000848
ShoulderLeft	Vm	4.4E-05	SpineShoulder	E/Am	0.00049

Fisher Scores for Walk Forward Phase

Head	Am	1.5E-05	FootRight	E	0.000427
Head	Jm	2E-06	KneeRight	E	3.3E-05
SpineMid	Vm	0	ShoulderLeft	Jm	1.2E-05

A.4 Main Workflow

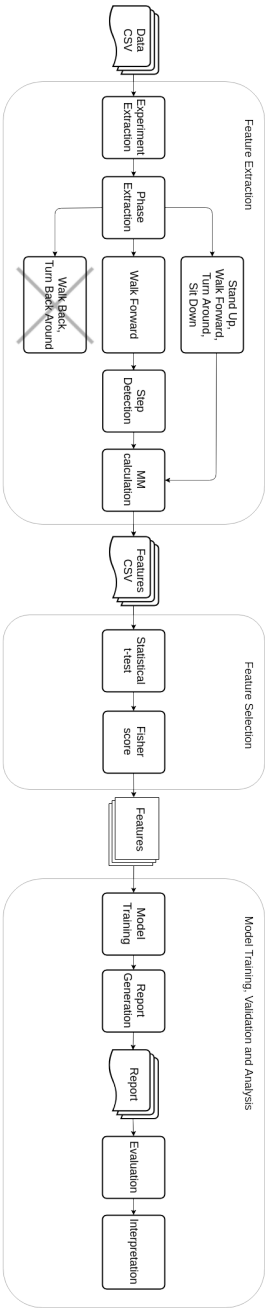


Figure A.1: Full schema of the main workflow of the current work.

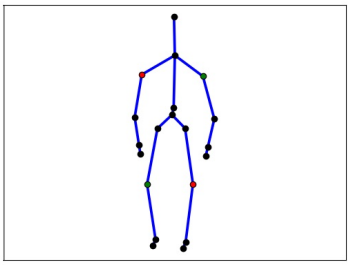
A.5 Report

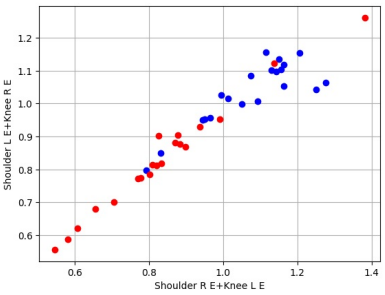
Below is given an example of how a report looks like.

Features:

Shoulder R E+Knee L E
Shoulder L E+Knee R E

KT_12, KT_18, KT_11, KT_20, KT_06, KT_02, KT_10, KT_15, KT_16, KT_05, KT_09, KT_24, KT_03, KT_01, KT_07, KT_23, KT_13, KT_04, KT_19, KT_08, SG03, SG07, SG04, SG15, SG24, SG16, SG08, SG01, SG02, SG11, SG09, SG21, SG17, SG13, SG23, SG14, SG19, SG22, SG05, SG18





Average Accuracy

	Average Accuracy
DecisionTreeClassifier	70.0
KNeighborsClassifier	87.5
LogisticRegression	70.0
SVC	80.0

Average Confusion Matrices

	tn	fp	fn	tp	Precision	Recall	F1
DecisionTreeClassifier	3.00	2.00	1.00	4.00	0.67	0.80	0.73
KNeighborsClassifier	4.50	0.50	0.75	4.25	0.89	0.85	0.87
LogisticRegression	4.75	0.25	2.75	2.25	0.90	0.45	0.60
SVC	4.50	0.50	1.50	3.50	0.88	0.70	0.78

Average Report for LogisticRegression

	precision	recall	f1-score	support
C	0.64	0.96	0.76	5.0
PD	0.94	0.47	0.60	5.0
macro_avg	0.79	0.72	0.68	10.0
micro_avg	0.70	0.70	0.70	10.0
weighted_avg	0.80	0.70	0.68	10.0

Average Report for KNeighborsClassifier

	precision	recall	f1-score	support
C	0.85	0.90	0.87	5.0
PD	0.90	0.84	0.87	5.0
macro_avg	0.88	0.87	0.87	10.0
micro_avg	0.88	0.88	0.88	10.0

weighted_avg	0.88	0.88	0.88	10.0

Average Report for SVC

	precision	recall	f1-score	support
C	0.74	0.9	0.81	5.0
PD	0.89	0.7	0.78	5.0
macro_avg	0.82	0.8	0.80	10.0
micro_avg	0.80	0.8	0.80	10.0
weighted_avg	0.82	0.8	0.80	10.0

Average Report for DecisionTreeClassifier

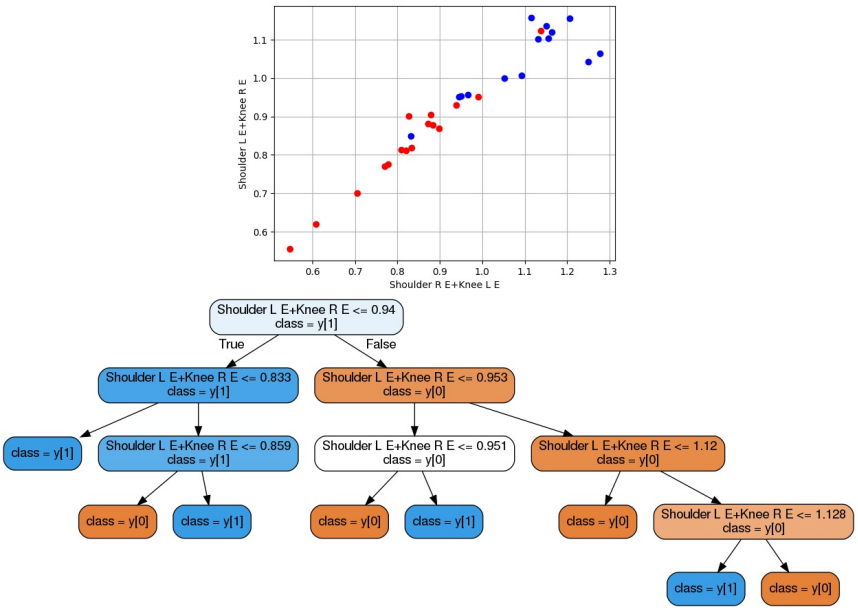
	precision	recall	f1-score	support
C	0.73	0.58	0.64	5.0
PD	0.68	0.80	0.73	5.0
macro_avg	0.70	0.69	0.68	10.0
micro_avg	0.70	0.70	0.70	10.0
weighted_avg	0.71	0.70	0.69	10.0

Accuracy for all of the folds

	0	1	2	3
LogisticRegression	80.0	70.0	70.0	60.0
KNeighborsClassifier	80.0	90.0	100.0	80.0
SVC	80.0	80.0	90.0	70.0
DecisionTreeClassifier	80.0	60.0	70.0	70.0

0 fold

	tn	fp	fn	tp	Precision	Recall	F1
DecisionTreeClassifier	2	2	1	5	0.71	0.83	0.77
KNeighborsClassifier	3	1	1	5	0.83	0.83	0.83
LogisticRegression	4	0	4	2	1.00	0.33	0.50
SVC	3	1	2	4	0.80	0.67	0.73



Report for LogisticRegression

	precision	recall	f1-score	support
C	0.83	0.83	0.83	6.0
PD	0.75	0.75	0.75	4.0
macro_avg	0.79	0.79	0.79	10.0
micro_avg	0.80	0.80	0.80	10.0
weighted_avg	0.80	0.80	0.80	10.0

Report for KNeighborsClassifier

	precision	recall	f1-score	support
C	0.83	0.83	0.83	6.0
PD	0.75	0.75	0.75	4.0
macro_avg	0.79	0.79	0.79	10.0
micro_avg	0.80	0.80	0.80	10.0
weighted_avg	0.80	0.80	0.80	10.0

Report for SVC

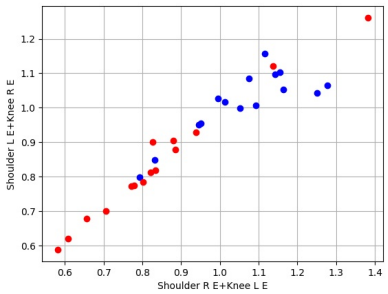
	precision	recall	f1-score	support
C	0.83	0.83	0.83	6.0
PD	0.75	0.75	0.75	4.0
macro_avg	0.79	0.79	0.79	10.0
micro_avg	0.80	0.80	0.80	10.0
weighted_avg	0.80	0.80	0.80	10.0

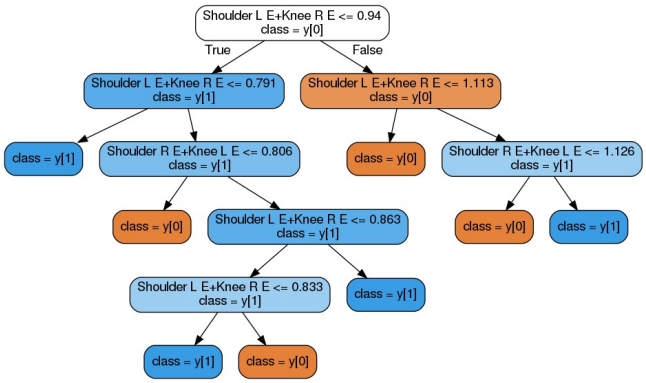
Report for DecisionTreeClassifier

	precision	recall	f1-score	support
C	0.83	0.83	0.83	6.0
PD	0.75	0.75	0.75	4.0
macro_avg	0.79	0.79	0.79	10.0
micro_avg	0.80	0.80	0.80	10.0
weighted_avg	0.80	0.80	0.80	10.0

1 fold

	tn	fp	fn	tp	Precision	Recall	F1
DecisionTreeClassifier	2	2	1	5	0.71	0.83	0.77
KNeighborsClassifier	3	1	1	5	0.83	0.83	0.83
LogisticRegression	4	0	4	2	1.00	0.33	0.50
SVC	3	1	2	4	0.80	0.67	0.73





Report for LogisticRegression

	precision	recall	f1-score	support
C	0.62	1.0	0.77	5.0
PD	1.00	0.4	0.57	5.0
macro_avg	0.81	0.7	0.67	10.0
micro_avg	0.70	0.7	0.70	10.0
weighted_avg	0.81	0.7	0.67	10.0

Report for KNeighborsClassifier

	precision	recall	f1-score	support
C	0.83	1.0	0.91	5.0
PD	1.00	0.8	0.89	5.0
macro_avg	0.92	0.9	0.90	10.0
micro_avg	0.90	0.9	0.90	10.0
weighted_avg	0.92	0.9	0.90	10.0

Report for SVC

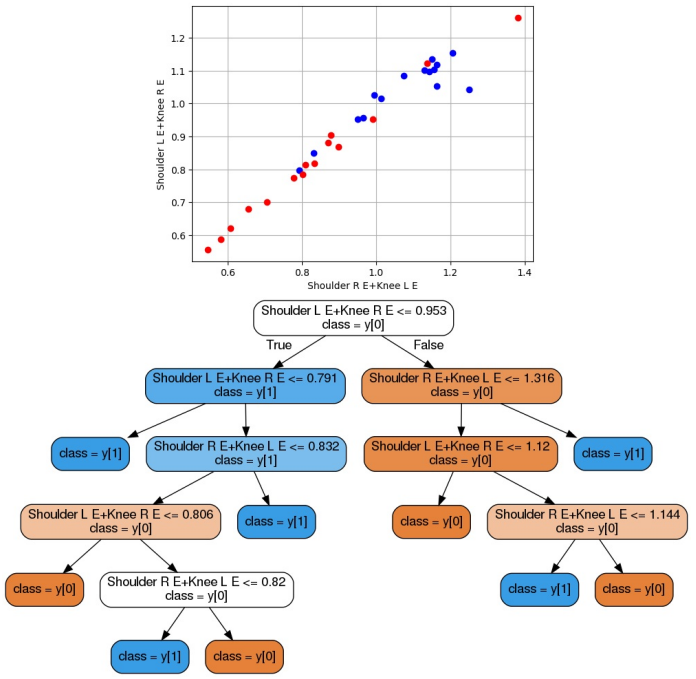
	precision	recall	f1-score	support
C	0.71	1.0	0.83	5.0
PD	1.00	0.6	0.75	5.0
macro_avg	0.86	0.8	0.79	10.0
micro_avg	0.80	0.8	0.80	10.0
weighted_avg	0.86	0.8	0.79	10.0

Report for DecisionTreeClassifier

	precision	recall	f1-score	support
C	0.67	0.4	0.50	5.0
PD	0.57	0.8	0.67	5.0
macro_avg	0.62	0.6	0.58	10.0
micro_avg	0.60	0.6	0.60	10.0
weighted_avg	0.62	0.6	0.58	10.0

2 fold

	tn	fp	fn	tp	Precision	Recall	F1
DecisionTreeClassifier	2	2	1	5	0.71	0.83	0.77
KNeighborsClassifier	3	1	1	5	0.83	0.83	0.83
LogisticRegression	4	0	4	2	1.00	0.33	0.50
SVC	3	1	2	4	0.80	0.67	0.73



Report for LogisticRegression

	precision	recall	f1-score	support
C	0.62	1.0	0.77	5.0
PD	1.00	0.4	0.57	5.0
macro_avg	0.81	0.7	0.67	10.0
micro_avg	0.70	0.7	0.70	10.0
weighted_avg	0.81	0.7	0.67	10.0

Report for KNeighborsClassifier

	precision	recall	f1-score	support
C	1.0	1.0	1.0	5.0
PD	1.0	1.0	1.0	5.0
macro_avg	1.0	1.0	1.0	10.0
micro_avg	1.0	1.0	1.0	10.0
weighted_avg	1.0	1.0	1.0	10.0

Report for SVC

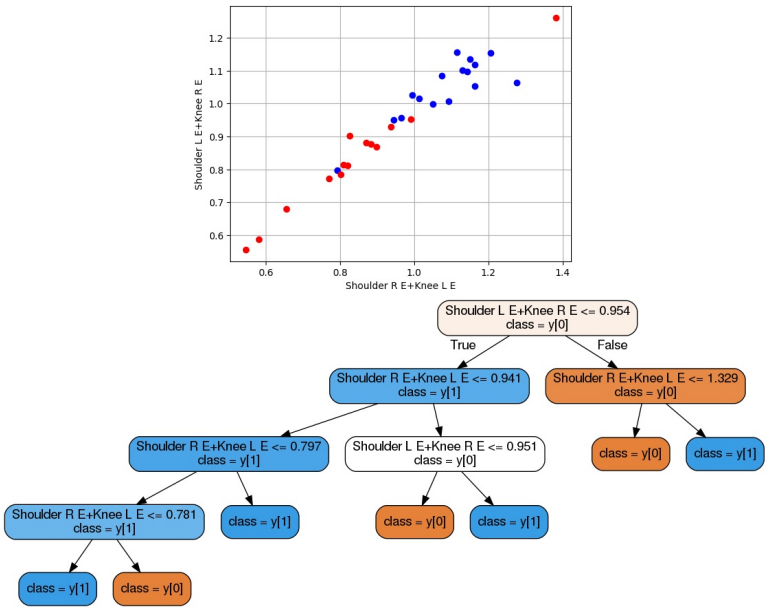
	precision	recall	f1-score	support
C	0.83	1.0	0.91	5.0
PD	1.00	0.8	0.89	5.0
macro_avg	0.92	0.9	0.90	10.0
micro_avg	0.90	0.9	0.90	10.0
weighted_avg	0.92	0.9	0.90	10.0

Report for DecisionTreeClassifier

	precision	recall	f1-score	support
C	0.75	0.6	0.67	5.0
PD	0.67	0.8	0.73	5.0
macro_avg	0.71	0.7	0.70	10.0
micro_avg	0.70	0.7	0.70	10.0
weighted_avg	0.71	0.7	0.70	10.0

3 fold

	tn	fp	fn	tp	Precision	Recall	F1
DecisionTreeClassifier	2	2	1	5	0.71	0.83	0.77
KNeighborsClassifier	3	1	1	5	0.83	0.83	0.83
LogisticRegression	4	0	4	2	1.00	0.33	0.50
SVC	3	1	2	4	0.80	0.67	0.73



Report for LogisticRegression

	precision	recall	f1-score	support
C	0.50	1.00	0.67	4.0
PD	1.00	0.33	0.50	6.0
macro_avg	0.75	0.67	0.58	10.0
micro_avg	0.60	0.60	0.60	10.0
weighted_avg	0.80	0.60	0.57	10.0

Report for KNeighborsClassifier

	precision	recall	f1-score	support
C	0.75	0.75	0.75	4.0
PD	0.83	0.83	0.83	6.0
macro_avg	0.79	0.79	0.79	10.0
micro_avg	0.80	0.80	0.80	10.0
weighted_avg	0.80	0.80	0.80	10.0

Report for SVC

	precision	recall	f1-score	support
C	0.60	0.75	0.67	4.0
PD	0.80	0.67	0.73	6.0
macro_avg	0.70	0.71	0.70	10.0
micro_avg	0.70	0.70	0.70	10.0
weighted_avg	0.72	0.70	0.70	10.0

Report for DecisionTreeClassifier

	precision	recall	f1-score	support
C	0.67	0.50	0.57	4.0
PD	0.71	0.83	0.77	6.0
macro_avg	0.69	0.67	0.67	10.0
micro_avg	0.70	0.70	0.70	10.0
weighted_avg	0.70	0.70	0.69	10.0