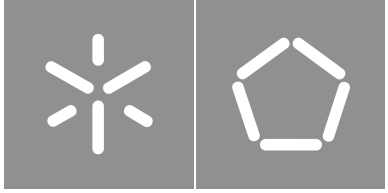
Universidade do Minho
Escola de Engenharia

Pedro Gonçalo Santos Pires Pinheiro

Artificial intelligence-based software for recognizing parkinsonian gait patterns based on wearable miniaturized sensors

⟨October⟩, ⟨2022⟩



Universidade do Minho
Escola de Engenharia

Pedro Gonçalo Santos Pires Pinheiro

Artificial intelligence-based software for recognizing parkinsonian gait patterns based on wearable miniaturized sensors

Master Thesis
Master in Informatics Engineering

Work developed under the supervision of:
Cristina Peixoto Santos

⟨October⟩, ⟨2022⟩

Acknowledgements

Firstly i want to thank my instructor Prof. Cristina Santos, for providing me with the opportunity of working on this thesis, and for all the guidance, valuable insights, as well as valuable research connections that helped me to grown and expand my knowledge provided during the development of this dissertation. This work wouldn't be possible without her contribution and research suggestions.

Thank you also to everyone involved in the +Sense project that helped me throughout the development of this thesis. I'm thankful for all the help and debated ideas.

I would also like to thank the Department of Informatics of University of Minho for letting me use their cluster as tool to make the big portion of this thesis have a faster development.

Thank you to Helena Raquel Gonçalves, for all the shared knowledge, tips, and help provided throughout the entirety of this thesis, without which this work would not have been possible.

I would also like to thank my friend Filipe Freitas, for their encouragement, support, and help throughout the entire process.

Last but not least, i would like to thank my friends that were not mentioned, as well as my family for their support, encouragement, providing motivation, and for listening to me and providing me with support during the development of this thesis.

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the Universidade do Minho.

Vila Verde, 29 August 2022

Pedro Gonçalo Santos Pires Pinheiro

Resumo

A Doença de Parkinson (DP) é uma doença degenerativa do sistema nervoso central, geralmente caracterizada por prejudicar vários aspetos da marcha dos pacientes, como bradicinesia, comprimento do passo encurtado e congelamento da marcha. As escalas de avaliação clínica são tipicamente usadas com base em exames para monitorizar esses sintomas motores associados à marcha. Além disso, estas avaliações são baseadas na memória dos pacientes e pesquisas subjetivas, fornecendo dados tendenciosos. Assim, são necessários dados de longo prazo sobre as atividades motoras diárias do paciente.

Avanços tecnológicos forneceram dispositivos sensores pequenos e vestíveis capazes de capturar dados de longo prazo, podendo ser utilizados em ambientes domiciliares permitindo a captura de dados precisos. A combinação desses sensores com inteligência artificial (IA) produz modelos capazes de biomarcar os níveis de doença, condições motoras e bem-estar dos pacientes, e de fornecer dados não tendenciosos sobre os padrões de marcha dos pacientes. A integração destes modelos num aplicativo para médicos facilitará gerir o estado de DP e tratamentos mais personalizados serão alcançados.

Tendo isto em conta, esta tese tem como objetivo usar dados de pacientes que apresentam deficiências de marcha para treinar modelos baseados em IA que sejam capazes de classificar níveis de doença, condições motoras e qualidade de vida desses pacientes.

Para isso, foram adquiridos dados de 40 pacientes com DP, com o objetivo de desenvolver 3 modelos de IA diferentes, um usado para classificar o nível de doença de um paciente na escala UPDRS-III, outro para classificar as condições motoras escala H&Y e outro usado para classificar a qualidade de vida. Esses modelos foram implementados numa APP para auxiliar os médicos durante as suas consultas.

Os resultados obtidos foram positivos. O modelo UPDRS-III conseguiu uma acurácia de 91,67%, uma sensibilidade de 90,43% e uma especificidade de 93,98%, enquanto o modelo H&Y alcançou uma acurácia de 88,98%, uma sensibilidade de 88,71%, e especificidade de 92,79%, sendo que o modelo PDQ-39 obteve acurácia de 84,19%, sensibilidade de 82,13% e especificidade de 90,24%.

Palavras-chave: Doença de Parkinson, Marcha, Biomarcar, Padrões, Sensores vestíveis, Inteligência Artificial

Abstract

Parkinson's Disease (PD) is a degenerative disease of the central nervous system, usually characterized by causing several gait impairment symptoms, such as bradykinesia, shortened stride length, shuffling gait and freezing of gait. Clinical assessment scales are typically used based on observational examinations to monitor these motor symptoms associated with gait. Further, these assessments are based on patients' memory recall, subjective surveys, medication phase, and mood during the appointment, providing biased data. Thus, long-term data regarding the patient's daily motor activities is required.

Technological advancements provided small and wearable sensor devices able to capture long-term acquisitions of data. Given their miniaturized size and portability, these sensors can be used in domiciliary environments enabling to capture accurate data. Combining these sensors with artificial intelligence (AI) produces models able to biomark patients' disease levels, motor conditions and well-being. These AI models can provide non-biased data about patients' gait-associated patterns. Integrating these AI-based solutions in a user-friendly clinic APP for physicians will facilitate PD management, and more personalized treatments will be achieved.

Taking this in mind, this thesis aims to use data from patients who show developed gait impairments to train AI-based models that are able to classify disease levels, motor conditions and the quality of life of said patients.

For that, data from 40 patients with PD was gathered. This data was then used to develop 3 different AI models, one used to classify a patient's disease level on the Unified Parkinson's Disease Rating Scale (UPDRS-III) scale, another to classify a patient's motor conditions on the Hoehn and Yahr (H&Y) scale, and another one used to classify a patient's quality of life (QoL). These models were then implemented in an easy to use APP to help the physicians during their appointments with the patients.

Positive results were obtained, being observed that. The UPDRS-III model manged to achieve achieve an accuracy of 91.67%, a sensitivity of 90.43%, and a specificity of 93.98%, while the H&Y model achieved an an accuracy of 88.98%, a sensitivity of 88.71%, and a specificity of 92.79%, and the Parkinson's Disease Questionnaire (PDQ-39) model achieved an accuracy of 84.19%, a sensitivity of 82.13%, and a specificity of 90.24%.

Keywords: Parkinson's Disease, Gait, Biomark, Patterns, Wearable sensors, Artificial Intelligence

Contents

List of Figures

List of Tables

1 Introduction

This dissertation presents the work carried out over the past year, integrated in the scope of the Master's Degree in Computer Science at the Biomedical Robotic Devices Lab (BIRDLAB) included in the Center of MicroElectroMechanicalSystems (CMEMS), a research center of the Department of Industrial Electronics (DEI) of University of Minho.

This project's main goal was to create an easy to use APP capable of helping physicians assessing a patient's PD disease level, motor conditions and quality of life, using data of 40 PD patients, gathered using inertial measurement units (IMU).

The project was divided into 4 main phases: **(i)** an in-depth research of state-of-the-art studies was carried out in order to better understand the scientific contributions about wearable devices combined with AI for recognition of parkinsonian gait patterns; **(ii)** the development, preparation and processing of dataset to be used as input for the Deep Learning (DL) models; **(iii)** the development of 3 different AI-based models, capable of classifying a patient on a motor scale, their disease stage and quality of life; **(iv)** integrating these models into an easy-to-use clinical APP that can aid physicians on assessing patient's disease stage, motor conditions and quality of life.

1.1 Motivation

PD is a degenerative disease of the central nervous system, characterized by causing several disabling motor symptoms associated with the mobility of patients [1]. It is often associated with a vast list of gait-associated disabilities for which there is still limited pharmacological/surgical treatment efficacy. These disabilities considerably increase the risk of falls and limit the quality of life and autonomy of the patient, who become dependent on third parties for the most trivial and daily activities [2]. Bradykinesia, shortened stride length, shuffling gait and freezing of gait are some of these prototypical gait-associated signs in PD [3], [4], [5]. PD is the second-most common neurodegenerative disorder that affects 2–3% of the population with more than 65 years of age, and even though the main focus of this thesis is on the disease's gait patterns, PD also shows non-motor symptoms like cognitive decline and autonomic failure [6], caused by the loss of neurons in the substantia nigra [7]. As seen in [8], its prevalence ranges from

1 to 2 per 1000 in unselected population people and affects 1% of the population above 60 years. PD is rare before the age of 50 years and reaches a prevalence of 4% in the highest age groups. An estimated 6.1 million individuals globally had a PD diagnosis in 2016 [9]. Some risk factors include age as the most important risk factor, male gender, certain pesticides and family history [10].

For monitoring these motor symptoms associated with gait, clinical assessment scales are typically used based on observational and surveys on the well-being of patients. The most common used scales include the UPDRS-III, H&Y and PDQ-39 scales. These assessments are based on patients' memory recall, subjective surveys, medication phase, and mood during the appointment. Thus, these tools provide biased data [11], and there are still weak links between gait patterns and the degree of illness. During routine appointments, doctors are limited to the information indicated by the patient, so objective and continuous metrics of the patient's gait patterns cannot be obtained [12]. Even more, there are a number of neurologic conditions that mimic the disease, making it difficult to diagnose in its early stages [13].

Due to technological advancements, downsized and wear-ability sensor devices have been used for monitoring PD motor symptoms. The small size of these devices, portability, low-cost and reduced power-consumptions allow their usage in everyday activities without interfering with the patients movement [14]. Wearable sensors, mainly, inertial sensors, have been used in the PD domain to: **(i)** estimate gait-associated metrics, such as step/stride time/length, velocity, cadence and gait asymmetries/variabilities; **(ii)** measure episodes of motor blockages, known as freezing of gait; **(iii)** capture postural deviations associated with poor walking performance; and **(iv)** capture kinematic-driven data. Advantageously, these data can be captured along long-term acquisitions, in home-scenarios and be used as input for other assistance/rehabilitation devices. In fact, if physicians could access these data, they would obtain more objective, reliable and continuous information about the actual stage of their patients' motor conditions. Additionally, gait deviations could be detected in the early-disease stage, facilitating the appointments of illness diagnosis where the motor symptoms are not so clear. Based on that information, treatments will be more personalized, and patients will benefit from a more closed disease management.

Combining inertial data with AI allows it to be possible to biomark patients' disease levels, motor conditions and well-being. In fact, statistical approaches have shown that for a considerable sample of patients, correlated with clinical scales, it is observed a stratification between each stage. Supported by these statistical findings, some researchers applied AI to recognize PD gait patterns and used that information to study patients' motor function [15]. Despite the positive results observed by the field-related scientific community, further research is required to: **(i)** better understand which wearable sensors body configuration and number can better produce meaningful information about patients' walking condition; **(ii)** study the possibility of using raw inertial data aiming to improve computational consumptions; **(iii)** and explore which dataset preparation methodologies, features selection and AI-based models can have better performance to describe gait parkinsonian patterns along disease levels. Therefore, a systematic approach that allows to overcome the gaps previously identified will be followed.

1.2 Problem Statement

Certain gait patterns, such as reduced gait speed [4], vary between different stages of the disease. Data collected using IMUs such as accelerometers and gyroscopes will therefore vary. The aim of this thesis is to see how data collected from patients with varying gait patterns, using inertial sensors, reflect on the disease status and quality of life.

To this end, this dissertation is inserted in the +sense project that aims to present front-end high-tech solutions based on wearable biofeedback devices which rely on acquisition, interpretation and feedback of patients' sensorimotor information. One of the technologies developed by this project is a wearable, with an integrated sensor, an instrumented strap. This project encompasses the production of a dataset with data gathering of 40 PD patients that ware asked to walk at a comfortable speed, and their repesctive clinical evaluations on the UPDRS-III and H&Y scales and the PDQ-39 questionnaire. Thus, an emergin contribution to this project are the models used to classify patients on the UPDRS-III, H&Y scales and PDQ-39 questionnaire and an APP that integrates these developed models, using data acquired by wearable sensory systems. These models will be used as a support tool for these patients' physicians. In order to achieve this, an easy-to-use APP shall be developed in which the 3 different AI models previously developed will be integrated. Along with these models, there will be other utilities this APP will provide to the physicians about the patients' PD state, such as gait related metrics.

1.3 Goals

The main goal of this thesis is the development of AI-based models able to automatically classify patients' the UPDRS-III and H&Y scales and on the PDQ-39 questionnaire, by using a dataset with data gathered from 40 PD patients using wearable sensory systems.

There are four major objectives that it will allow to pursue the ultimate goal of this master thesis, which are outlined below:

Goal 1: Identification and analysis of similar work in the literature and patents

This goal aims to complete an intensive and extensive research on PD, with great emphasis on its typical gait patterns. During this phase, a critical review of the technologies based on AI for recognizing gait patters in neurological diseases will also be carried out. The goal of this review is to identify the limitations of current systems to contribute in an innovative way to the scientific panorama. This goal is adressed in Chapter 2.

Goal 2: Dataset preparation

For the second step of this thesis, the preparation and development of a dataset to be used for developing the previously mentioned AI models is expected. The data will be captured using the +sMotion module. The final dataset will be used to the train the different DL models. This goal is adressed in

Chapter 3.

Goal 3: Implementation of parkinsonian gait patterns recognition algorithms based on signals from wearable miniaturized sensors

This goal consists of the development of an AI model capable of classifying PD patients in the UPDRS-III, H&Y and PDQ-39 scales. To this end, the training of several AI models with the aim of selecting the one with the best performance is expected. It is also expected to improve the results found in [16] that were able to reach an accuracy of 99.4% in the H&Y scale, and the results found in [17] that were able to reach an accuracy of accuracy of 75.3% in the UPDRS-III scale. This goal is adressed in Chapter 4.

Goal 4: Model integration on user-friendly APP and APP validation

The fourth and last step of this thesis consists in the development of a user-friendly app that implements the model with the best performance, aiming to be used by physician and researchers in the area. This goal is adressed in Chapter 5.

1.4 Research Questions

Considering the ultimate goal of this thesis and the step-goals presented, relevant research questions were identified, as follows:

- **RQ 1:** Which scales allow a more comprehensive assessment of the patient? This question relates to **Goal 1** and is answered in Section 2.1.
- **RQ 2:** Which and how many sensors are used, and where are these sensors placed on the patient's body? This question relates to **Goal 2** and the answers can be found in Section 2.3.3.
- **RQ 3:**What are most the commonly used AI-based models? This question relates to **Goal 3** and is answered in Section 2.3.4.
- **RQ 4:** How to integrate the previously developed models in the APP? This questionnaire relates to **Goal 4** and is answered in Section 5.3.

1.5 Contributions To Knowledge

The development of this thesis will result in the development of 3 different AI-based models and an easy-to-use APP which integrates these models. Each of these models will have its purpose, such as classifying a PD patient in the UPDRS-III scale, classifying a PD patient in the H&Y scale and classifying a PD patient in the PDQ-39 questionnaire. The APP will be able to help the patients' physicians and it will serve as a PD monitoring tool.

1.6 Dissertation Structure

The first chapter of this manuscript presents an introductory section that explains PD, some of the associated gait patterns, and why it is important to monitor these patterns. It also describes limitations of today's current monitoring systems/exams and why wearable sensors combined with AI are an ideal system to describe patients' gait patterns. After this section, it is explained the main goals of this project, such as developing an AI model and implementing it in a user-friendly APP that physicians can use.

The second chapter provides a review on the state-of-the-art studies. In this chapter, gait patterns, data acquisition methods, AI-based models and their performance evaluation are explained in more depth.

In the next chapter, it is presented an overview of the proposed solution framed on the context of the project that this dissertation is integrated. This project's name is +Sense: Sensory biofeedback devices for patients with PD.

The fourth chapter presents the pipeline used for the development of the AI-based models. It presents the data input-preparation process used in the data gathered using the +sMotion module from the +Sense project, the methodology for validation, training and testing, and the implemented metrics for model evaluation. Also, in this chapter, the obtained results are shown, comparing these with the outcomes presented on the reviewed state-of-the-art of Chapter 2.

The fifth chapter presents a guide for the developed APP. The previously developed models were integrated in this APP so that it can be used as a PD monitoring tool by the patients' physicians.

The sixth and final chapter culminates with the conclusions of this thesis as well as provide future directions for this project.

2

Literature Review

2.1 Introduction

There are multiple gait-debilitating diseases, being the most known besides PD, Amyotrophic Lateral Sclerosis (ALS) and Huntington's Disease (HD). The specific characteristics of gait disorders may differ across different neurological diseases [18]. These diseases have some gait patterns in common, such as reduced stride length, step cadence and walking speed. However, some of the gait patterns of a disease do not show in another disease. For example, for patients with ALS, their average stride interval is significantly longer than that of healthy controls or of patients with PD or HD [19]. Also, patients with ALS have less steady gait between successive stride intervals [20], while patients with HD show slowed execution of movements in the upper limbs [21]. As for PD, some of the most common gait impairments are shuffling gait, freezing of gait, impaired balance and postural instability. Patients that suffer from PD can show multiple features of these debilitating gait patterns.

As previously stated, the traditional methods of assessing a patient's disease state are usually based on biased data [11], because physicians, during routine appointments are limited to the information indicated by the patient [9], [22], so objective and continuous metrics of the patient's gait patterns cannot be obtained. This means that a front-end monitoring system able to provide continuous and non-biased data is needed so the physicians can present a close follow-up of their patients.

Wearable sensors, such as accelerometers, gyroscopes and magnetometers have been tremendously used for motion monitoring in PD as seen in [23], [24]. They provide continuous and objective data, comprising non-intrusive tech easily integrated in the patients' daily activities. The most common sensors used are IMUs, such as accelerometers, gyroscopes or pressure sensors [25], [26], [27]. The number of sensors used for this acquisition may vary, being still not clear the ideal number and body configuration. Some of the most common body locations to place these sensors are the arms, legs and trunk of the patients. On the other side, other type of wearable sensors, such as pressure sensors are placed on insoles. The data collected from these sensors are then used to create datasets that will serve as the input for the AI-based models. Some of these datasets are then released to the public following an open-source basis, such as the Gait in Parkinson's Disease by PhysioNet [28], [11], [29],[16] which contains ground

reaction force data obtained during walking from 93 participants with mild to intermediate PD, and 73 Healthy Controls, or the Gutenberg Gait database [30], which includes data about 350 healthy individuals recorded in laboratory over the past seven years, as of 2021. However, most of these datasets did not present a clear and standard explanation about data input type, protocols and sensors configurations.

Researchers have developed automatic classifiers of parkinsonian gait patterns by combining data from these wearable sensors with AI models [31], [28], [31]. However, the scientific challenge of applying AI on inertial data remains to be addressed, as further investigations are required on what sensory information should be used to identify gait patterns, and which AI models can classify these gait patterns with the most accuracy.

In light of the need to better understand the state-of-the-art, a comprehensive review was accomplished on the scientific contributions of wearable devices combined with AI for recognition of parkinsonian gait patterns. From this critical review, the following questions were investigated and answered: (i) Which parkinsonian prototypical gait patterns have been recognized by combining AI-based models with sensory acquired data? (ii) Which type of data was used in AI models input for recognition of parkinsonian prototypical gait patterns? and (iii) Which AI models and how were they implemented in parkinsonian gait patterns recognition systems?

Most of the reviewed papers used the Support Vector Machine (SVM) [29], Decision Trees [16], or Artificial Neural Networks (ANN) [15], [32] as the preferred algorithm for the development of an AI model. However, some other algorithms used are Naïve-Bayes [16], Logistic Regression [33], Hidden Markov [14] and Long-Short Term Memory (LSTM) models [34]. Most of the developed models don't use Real-Time processing or Feature Selection. MATLAB [28], [29] and Python [34] are the most common used programming languages for the development of Machine Learning and AI models. In addition to this, the validation metrics used the most were sensitivity, specificity and accuracy. Combining data from these wearable sensors with these AI models makes it possible to develop automatic classifiers of parkinsonian gait patterns. A tool such as this will be able to support physicians by providing continuous, non-biased and clinically related data of patients' gait patterns. Thus, a more in-depth analysis is achieved, and the physician can treat the patients in the most efficient way.

2.2 Methods

2.2.1 Data sources, search strategy and study selection

An electronic systematical search was carried out on databases such as Google Scholar and Scopus, looking for studies related to the use of AI models to classify or predict gait-associated disorders in PD patients. The literature search was performed according to the guidelines of Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), as depicted on Figure 1. For that purpose, keywords matching headings were used: ["Parkinson's Disease AND Gait Patterns"]; ["Parkinson's Disease AND

Gait patterns AND Artificial Intelligence"]; ["Parkinson's Disease AND Gait patterns AND Machine Learning"]; ["Parkinson's Disease AND Machine Learning"]; ["Parkinson's Disease AND Artificial Intelligence"].

Studies were included if they fulfilled the following inclusion criteria: (i) studies of idiopathic PD, (ii) usability of AI-based models to recognize prototypical parkinsonian gait patterns, (iii) data used in the paper was collected from healthy patients or PD patients or both, (iv) results were published in the English language and within the past 10 years. The exclusion criteria were: (i) not using inertial sensors data, (ii) not using AI-based models in the classification of PD gait patterns. Some studies reference lists were searched for additional support.

2.3 Results

In the following sections it is described the method of eligibility of studies to be reviewed, the prototypical parkinsonian gait patterns, datasets and the AI models implemented in current related state-of-the-art.

2.3.1 General Results

A total of 162 studies were identified through Google Scholar (n=139) and Scopus (n=23) databases. Duplicates were removed (n=85). Some of the remaining studies were excluded after reviewing their titles (n=15) and abstracts (n=27). From the 77 titles and abstracts retrieved, 55 full-text studies were assessed for eligibility. Studies that did not meet the predefined inclusion criteria were excluded. 9 studies met the eligibility criteria and were included in this review. This approach was represented in section 2.3.1 in Figure 1.

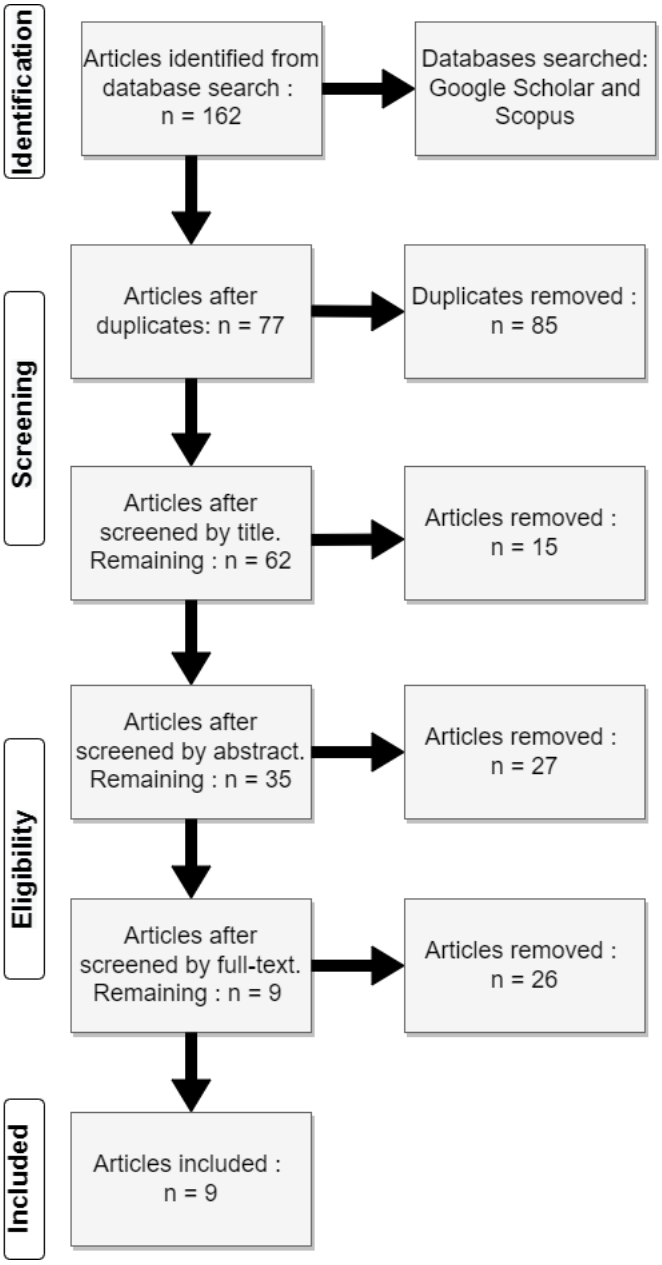


Figure 1: Flowchart for the search strategy based on PRISMA

2.3.2 Parkinsonian prototypical motor patterns

After reading and analyzing the selected articles, these were divided into three categories according to their underlying goal. All of these articles shared a common purpose: to help the early diagnosis of patients with PD through wearable sensors. Despite that, the metrics identified were used for :

1. **Disease progression** - articles that seek to find a correlation between the data and clinical scales used to establish the severity of the disease or perceive how the disease is evolving over time in a patient.

2. **Diagnosis** - articles that used the data collected to distinguish PD patients from healthy controls.
3. **Detection of Freezing of gait** - articles that explore the detection and prediction of Freezing of gait

Table 1 presents the main goal and stated gait patterns in the articles that were selected.

Table 1: Prototypical gait patterns found in the review

Paper	Goal	Gait Patterns	Scale
[11]	Disease Progression, Diagnosis & Detection of Freezing of gait	Tremor, bradykinesia, rigidity and postural instability	H&Y & UPDRS-III
[28]	Diagnosis	Resting tremor, and bradykinesia	Not indicated
[29]	Diagnosis	Resting tremor, muscle rigidity, bradykinesia, and postural instability	Not indicated
[34]	Detection of Freezing of gait	Freezing of gait	Not indicated
[32]	Diagnosis & Disease Progression	Bradykinesia, rigidity, impaired balance, and postural control	H&Y
[15]	Diagnosis	Reduced cadence, step length and walking speed	Not indicated
[16]	Classify motor disability	Bradykinesia, worse balance and posture	H&Y
[31]	Detection of Freezing of gait	Freezing of gait	Not indicated
[35]	Detection of Freezing of gait	Freezing of gait, slower and shorter stride lengths	Not indicated

The most common scale used to automatically classify a patient in, is the H&Y scale. This scale has become the most commonly and widely used scale to estimate the severity of PD in a patient [36] by quantifying the disease stage [37]. Progression in HY stages has been found to correlate with motor decline, deterioration in quality of life, and neuroimaging studies of dopaminergic loss [38]. However, another important scale referenced in the state-of-theart is the UPDRS-III scale, which is considered to be the gold standard clinical rating scale for PD [39]. The UPDRS is a scale that was developed as an effort to incorporate elements from existing scales to provide a comprehensive, efficient and flexible way of measuring and monitoring PD-related disability and impairment [40]. Taking this into account, these are the scales that this thesis will aim to automatically classify patients in using AI-based models.

2.3.3 Data Input

Table 2 indicates the used sensors for data acquisition, such as gyroscopes, accelerometers and pressure sensors. Also, the number of sensors used is presented. The number of sensors used varies from 2 to 22 and typically the placement of these sensors is in the feet. In most cases, the protocol used for data gathering was to invite the patients to walk at a self-selected speed for a certain amount of time. The number of patients used for data gathering was between 10 and 168, and these groups of patients most of the time included both patients with PD and healthy controls. These procedures usually take place in a laboratory.

Table 2: Data acquisition for AI-based models according to the reviewed studies

Paper	Sensors			Data Acquisition			Dataset	
	Which one?	**How many?**	**Where?**	**Participants**	**Protocol**	**Setting**	**Online/Authors**	**Preparation**
[11]	Pressure Sensors	16	Feet	93 patients with PD and 73 healthy controls	Walking at self-selected walking pace for 2 minutes	Laboratory	Goldberger, A., Amaral, L.	NI
[28]	Pressure Sensors	16	Feet	93 patients with PD and 73 healthy controls	Walking at self-selected walking pace for 2 minutes	Laboratory	Goldberger, A., Amaral, L.	NI
[29]	Pressure Sensors	16	Feet	93 patients with PD and 73 healthy controls	Walking at self-selected walking pace for 2 minutes	Laboratory	Goldberger, A., Amaral, L.	NI
[34]	Accelerometers, gyroscopes	3	Legs and hips	10 patients with PD	Patients performed 3 different walking tasks	Laboratory	Plotnik, M., Roggen, D.	NI

Continued on next page

Table 2 – *Continued from previous page*

Paper	Sensors			Data Acquisition			Dataset	
	Which one?	**How many?**	**Where?**	**Participants**	**Protocol**	**Setting**	**Online/Authors**	**Preparation**
[32]	Retroreflective markers	22	Seventh cervical vertebra, sacrum, bilaterally over the acromion, anterior superior iliac spine, greater trochanter, lateral femoral condyle, fibula head, lateral malleolus, and metatarsal head	76 patients with PD and 67 healthy individuals	Walking at self-selected walking pace for 10 meters	Laboratory	Varrecchia, T., Castiglia, S.	Walking speed, cadence, step width, step length
[15]	Pressure Sensors	16	Feet	93 patients with PD and 73 healthy controls	Walking at self-selected walking pace for 2 minutes	Laboratory	Goldberger, A., Amaral, L.	NI
[16]	Pressure Sensors	16	Feet	93 patients with PD and 73 healthy controls	Walking at self-selected walking pace for 2 minutes	Laboratory	Goldberger, A., Amaral, L.	NI
[31]	Accelerometer	2	Knees	36 patients with PD and 15 healthy controls	Walking at a comfortable speed	Laboratory	Aich, S., Pradhan P.	Preprocessed data was filtered using a low pass filter

Continued on next page

Table 2 – *Continued from previous page*

Paper	Sensors			Data Acquisition			Dataset	
	Which one?	**How many?**	**Where?**	**Participants**	**Protocol**	**Setting**	**Online/Authors**	**Preparation**
[35]	Accelerom-eter	2	Leg	11 pa-tients with PD	Walking in a straight line. TUG test was also performed	Labo-ratory	Borzi, L., Mazzetta, I.	NI

2.3.4 AI-based Models

The information gathered about which AI-based models was used in the reviewed papers, the programming language used, the performance of these models and which features were used can be seen in Table 3.

Table 3: Information gathered about AI-based models according to the reviewed studies

Paper	Model	Features	Feature selection/extraction	Programming Language	Performance	Real-Time?
[11]	Support Vector Machine, Artificial Neural Network	Tremor frequency, rhythmicity, frequency and amplitude of repetitive movements	Was used data from both upper and lower extremities to extract features	MATLAB	Accuracy: 95%, Specificity: 96%, Sensitivity: 97%	No Real-Time performance
[28]	Classification and Regression Trees	NI	Vibes algorithm and the OneRAttributeEval algorithm for feature selection	MATLAB	Accuracy: 96,39%, Sensitivity: 96,77%, Specificity: 95,89%	No Real-Time performance
[29]	Support Vector Machine	Several gait features are extracted as the feature vectors	Extracted four different features to create feature vectors for the classification	MATLAB	Accuracy: 93.37%, Sensitivity: 92.47%, Specificity: 94.52%	No Real-Time performance
[34]	Long Short-Term Memory	NI	No feature selection/extraction	Python	Accuracy: More than 90%	No Real-Time performance
[32]	Artificial Neural Network	NI	A Principal Component Analysis was used to define a subset of features	MATLAB	Accuracy: 66.16%, Sensitivity: 66%, Specificity: 85%	No Real-Time performance
[15]	Artificial Neural Network	Average angle of the hip, knee, and ankle	Some features were extracted	NI	Accuracy: 95.63%	No Real-Time performance

Continued on next page

Table 3 – *Continued from previous page*

Paper	Model	Features	Feature selection/extraction	Programming Language	Performance	Real-Time?
[16]	Decision Tree, Support Vector Machine and Bayes Classifier	NI	No feature selection/extraction	MATLAB	Accuracy: 99.4%, Sensitivity: 99.6%, Specificity: 99.8%	No Real-Time performance
[31]	Support Vector Machine, Decision Tree	Step time, stride time, step length, stride length and walking speed	Five gait parameters are calculated from accelerometer data	MATLAB	Accuracy: 88%, Sensitiviy:90.89%, Specificity:91.21%	No Real-Time performance
[35]	Support Vector Machine	NI	Exploited Decision Trees for feature selection	NI	Accuracy: 86.1%, Sensitiviy:86.3%, Specificity:85.5%	Has Real-Time performance

In [11] both SVM and ANN algorithms were used. The features used were tremor frequency, rhythmicity, frequency and amplitude of repetitive movements. These features were extracted from data from both upper and lower extremities. An accuracy of 95%, sensitivity of 96%, and specificity of 97% was obtained.

In [28], Classification and Regression Trees were used. The features used weren't specified, but they obtained them using the Vibes algorithm and the OneRAttributeEval algorithm. An accuracy of 96,39%, sensitivity of 96,77%, and specificity of 95,89% was obtained.

In [29], the SVM algorithm was used. The features used weren't specified, but they extracted four different features to create feature vectors for the classification. An accuracy of 93,37%, sensitivity of 92,47%, and specificity of 94,52% was obtained.

In [34], the LSTM algorithm was used. The features used weren't specified. An accuracy of more than 90% was obtained.

In [32], the ANN algorithm was used. The features used weren't specified, however a Principal Component Analysis was used to define a subset of features. An accuracy of 71,68%, sensitivity of 71,5%, and specificity of 88% was obtained.

In [15], the ANN algorithm was used. The features used were average angle of the hip, knee and ankle. For this a Principal Component Analysis was used to define a subset of features. An accuracy of 95,63% was obtained.

In [16], the Decision Tree, SVM and Bayes Classifier algorithms were used. The features used weren't specified. An accuracy of 97,6%, sensitivity of 99.6%, and specificity of 99.8% was obtained.

In [31], the SVM and Decision Tree algorithms were used. The features used were step time, stride time, step length, stride length and walking speed. These features were calculated from the collected data of an accelerometer. An accuracy of 88%, sensitivity of 90,89%, and specificity of 91,21% was obtained.

In [35], the SVM algorithm was used. The features used weren't specified, but a Decision Tree was exploited for the feature selection process . An accuracy of 86,1%, sensitivity of 86,3%, and specificity of 85,5% was obtained.

The preferred programming languages used were MATLAB and Python. Only one of these articles implemented Real-Time performance [35].

2.4 Discussion

2.4.1 Which parkinsonian prototypical gait patterns have been recognized by combining AI-based models with sensory acquired data?

Along with the most common PD gait patterns presented before, such as reduced cadence, stride length and freezing of gait, there is the possibility of also monitoring some other patterns, such as postural instability, insufficient heel strike, toe clearance and asymmetric stride duration, resulting in a larger array of monitored gait patterns that can be integrated in a rehabilitation tool. This would result in a tool that can provide a more personalized treatment and a consistent, feasible and objective way for the physician to monitor patients' motor symptoms. It was observed that most of the studies compared if a prototypical gait parkinsonian was observed by comparing healthy subjects with PD control groups. However, it will benefit and will be innovative if the PD groups could be stratified by disease stage, disability level or even well-being state (as on H&Y, UPDRS-III or PDQ-39). This will enable to use wearable sensors data combined with AI to complement clinical examinations about patients' profile with a holistic point of view (motor and non-motor assessment). The selected studies did not contain AI-based models capable of automatically classifying a patient on the PDQ-39 questionnaire. By implementing this in the APP, the physician can perform a more complete assessment of the patients' PD, as well as their disease stage and motor disability.

2.4.2 Which type of data was used in AI models input for recognition of parkinsonian prototypical gait patterns?

It was observed that the most common method of obtaining data for the training of an AI model was using inertial sensors, such as the gyroscope and the accelerometer and foot pressure sensors. Gyroscopes and accelerometers were usually placed on the arms, legs and trunk of the patients, and the foot pressure sensors were usually placed on the feet of the patient or on the floor. The number of sensors used for the foot pressure sensors is usually 2, one on each foot. However, it may vary, with maximum number of sensors being 22, using the same number of pressure sensors on each foot. Regarding the inertial sensors (accelerometer and gyroscope), the most common sensors configurations included the use of 2 to 5 sensors. It is required to find a trade-of between the number and body location of sensors to use without losing meaningful information and use intrusive systems. As refered in [41], the placement of just one inertial sensor in the patients' lower back is able to measure a complete gait cycle, while using foot pressure sensors requires placing the sensors in both feet to acquire the data related to a complete gait cycle. Thus, less processing time is required. Using the lower back as the location to place the inertial sensor allows the measurement of important sensory information about the central segment of the body (centre of mass), which may be important for characterizing balance and postural stability.

Data collection settings only considered protocols on controlled environments, such as laboratories, under the supervision of physician. Commonly patients were invited to walk on level ground for a certain distance along a predefined track and perform realistic daily living activities, such as fetching coffee or opening doors, or walking with numerous turns. This setting allows for the physicians to control the entire procedure, making sure the data that is obtained is entirely non-biased. However, there is no standard protocol to capture significant motion information about patients' motor condition in daily life motor tasks. Furthermore, it is required more clinical evidence, aiming to improve datasets.

After the data acquisition, researchers have the option of making the developed dataset public. This might help some developers to extrapolate data input signals to other related investigations and accelerate the development of technological solutions in PD field or even for other neurological diseases.

In the related state-of-the-art studies, in most cases a big part of the development of the project consisted in big amounts of processing the datasets. Thus, a higher computational power is needed to eventually use these datasets to train AI-based models, which translates into a larger time-window required to complete this process. This means that the minimum processing of the dataset used in this thesis is required. Ideally, no processing is needed in order to hasten the training process.

2.4.3 Which AI models and how they were implemented in parkinsonian gait patterns recognition systems?

The review provided insights into which AI algorithms were used to classify gait patterns in PD. It was verified that the most common AI-based models, included the use of SVM, Support Vector Regression, Naïve-Bayes, Logistic Regression and ANN. However, if the dataset being used is a time series, the aforementioned algorithms will not deliver an accuracy as good as a Recurrent Neural Networks, like the LSTM. Advantageously, DL models such as the LSTM, are able to be fed with raw input data, while other Machine Learning algorithms do not provide such functionalities [42]. Even more, DL is suited for analyzing and extracting useful knowledge from large amounts of data [43].

The use of these models to classify gait patterns in patients with PD prevents the biased diagnosis of a clinician. Meaning, the development of a model with a great gait pattern classification accuracy can prevent the misdiagnosis of PD and other diseases alike. By implementing this model into an easy-to-use APP, it could then be of great help to the physician when diagnosing a patient. In fact, it is required to assess the level of acceptiblity and usability of these clinical APPs.

2.5 Conclusions and Future Directions

A literature review about the use of AI-based models to classify PD gait patterns was carried out. With it, came in-depth knowledge about PD gait patterns, being the most common patterns the slower gait speed, shortened stride length, shuffling gait and freezing of gait.

These gait-associated impairments are usually identified using either an IMU or a camera motion analysis system. IMUs are lower cost/power computation technology, being able to be used on patients' home scenarios. The most used IMUs integrated are accelerometers and gyroscopes and some other common sensors were foot pressure sensors. The number of sensors used for data acquisition varies from project to project, but if we're talking about foot pressure sensors it usually varies from 2 to 22, with each foot having the same number of sensors. For the IMUs the most common amounts vary between 2 and 5. These sensors are commonly placed in the arms, legs and trunk of the patients.

SVM, Support Vector Regression, Naïve-Bayes, Logistic Regression and ANN are some common algorithms used to develop AI models. These models are usually implemented using either MATLAB or Python. The most common validation metric is the accuracy. However, sensitivity and specificity are also commonly used. High validation metrics, such as the ones stated previously, can be reached with well developed AI models. This means the purpose of this thesis can be accomplished.

Despite technological and scientific advancements, some limitations were found. Table 4 summarizes the identified limitations regarding technological, adopted strategies, and validation methodology issues and it is also provided guidelines for their mitigation. Therefore, a systematic approach will be followed to identify the requirements of the system, from the point of view of the user and technologies, considering the limitations identified in the literature review, allowing to move on to the next dissertation tasks:

Table 4: Limitations identified in the literature review

Limitation	End user requirements	Guidelines
Lack on parkinsonian gait recognition	Holistic patients' assessment	Assess patients by different scales which include disease, motor and well-being assessment
Non clear body configuration of WS	Portability, comfort, easy set-up	Find a trade-off between the number/location of sensors without losing significant data
No data acquisition from home-based conditions or inclusion of daily motor tasks	Personalized treatments	Perform experimental tests including daily tasks in home-based scenarios
No assessment of acceptability of clinical APPs based on AI-based models integrated applied to inertial data	Acceptability of the device	Include the users' opinion in the validation of the proposed solution and assess its acceptability and usability

3

Solution Overview

3.1 Introduction

The main goal of this thesis is the development of AI-based models to automatically classify patients' disease stage, using data acquired by wearable inertial sensors. It is expected to recognize the presence of gait patterns to biomark the disease level, motor disability and quality of life level. Further, it is planned to integrate the AI-based models into a user-friendly clinical APP. This dissertation frames in project research titled by +SENSE: Sensory biofeedback devices for patients with Parkinson's Disease, which aims towards high-tech solutions to mitigate motor symptoms in PD. Therefore, with this dissertation it is expected to contribute with a clinical decision support tool, to complement physicians' examinations of patients' motor conditions with more objective and reliable data.

Data input of the AI models was recorded with a wearable motion lab, +SENSE device, an instrumented waistband, which has integrated an IMU to capture patients' lower trunk kinematic-driven data. This IMU has integrated an accelerometer and gyrospoce providing information about patients' lower trunk acceleration and angular velocity. The recorded inertial data measured the patients' gait patterns which will feed the AI models. The dataset contains data from 40 patients with PD who were asked to walk at a comfortable speed three times for a distance of 10 meters at a comfortable speed. Besides the motion-related data captured, clinical and sociodemographic data were recorded during the experimental tests. Thus, patients were also assessed considering their (i) disease level, using H&Y scale; (ii) motor condition assessed by UPDRS-III; and (iii) QoL level using a specialized scale to PD, the PDQ-39.

It is expected to implement the required dataset processing methodologies, used DL models given its ability biomark:

1. Disease level biomarker by classifying the patient's H&Y stage;
2. Motor disability by automatically rating UPDRS-III score;
3. Quality of life pointing PDQ-39 score;

It is expected with these advanced models to produce an holistic information about a patient with PD, regarding three key assessment levels, illness, motor and well-being domains.

The ultimate goal includes the integration of these AI-based models into an user-friendly APP to be used by physicians. Loading the captured inertial data with the wearable +sense device in the expected AI-based APP, physicians can complement their traditional examinations of patients' motor behaviours, with an extra objective assessment and smart classification. If the inertial data was captured in home-scenarios more reliable and feasible data are obtained about patients' motor conditions during their quotidian. Figure 2 depicts the conceptual overview delineated for this dissertation.

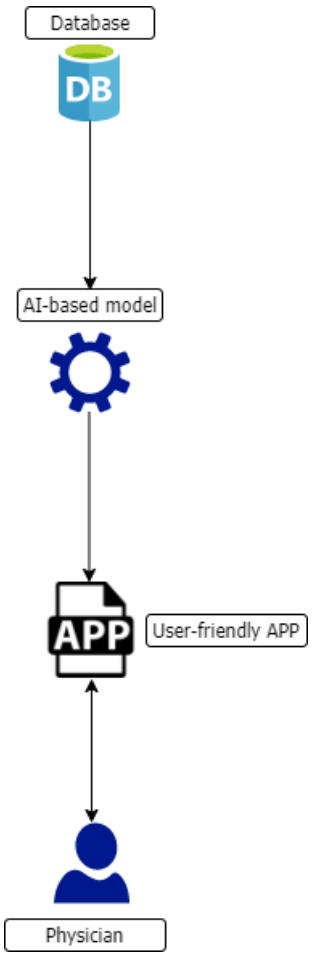


Figure 2: Diagram explaining the APP workflow

In this chapter, it is presented the project in which this dissertation frames, the +sense project. To that end, an explanation of the goal of this project is presented, followed by a brief description of the +sMotion module that is responsible for the data acquisition used for the development of this thesis. This is followed by an introduction to the +sC-Support, which is the module of this project responsible for the development of the clinical APP, making it the most important module for the development of this thesis.

3.2 +sense

This thesis is integrated into the +sense project. +sense presents high-tech front-end solutions based on wearable biofeedback devices which rely on the acquisition, interpretation and feedback of patients' motor information. The project envisions improving patients' quality of life, being less dependent on third parties by promoting their motor autonomy. The project comprises three main technologies: an instrumented waistband, smartphone and desktop APPs, and mixed reality strategies. These technologies are used by the four modules that comprise +sense, as shown in Figure 3: (1) +sBiofeedback; (2) +sMotion; (3) +sC-Support and (4) +sImmersive. This dissertation used the functionalities of +sMotion module (mainly the acquired data with the instrumented waistband) and contributed to the +sC-Support module.

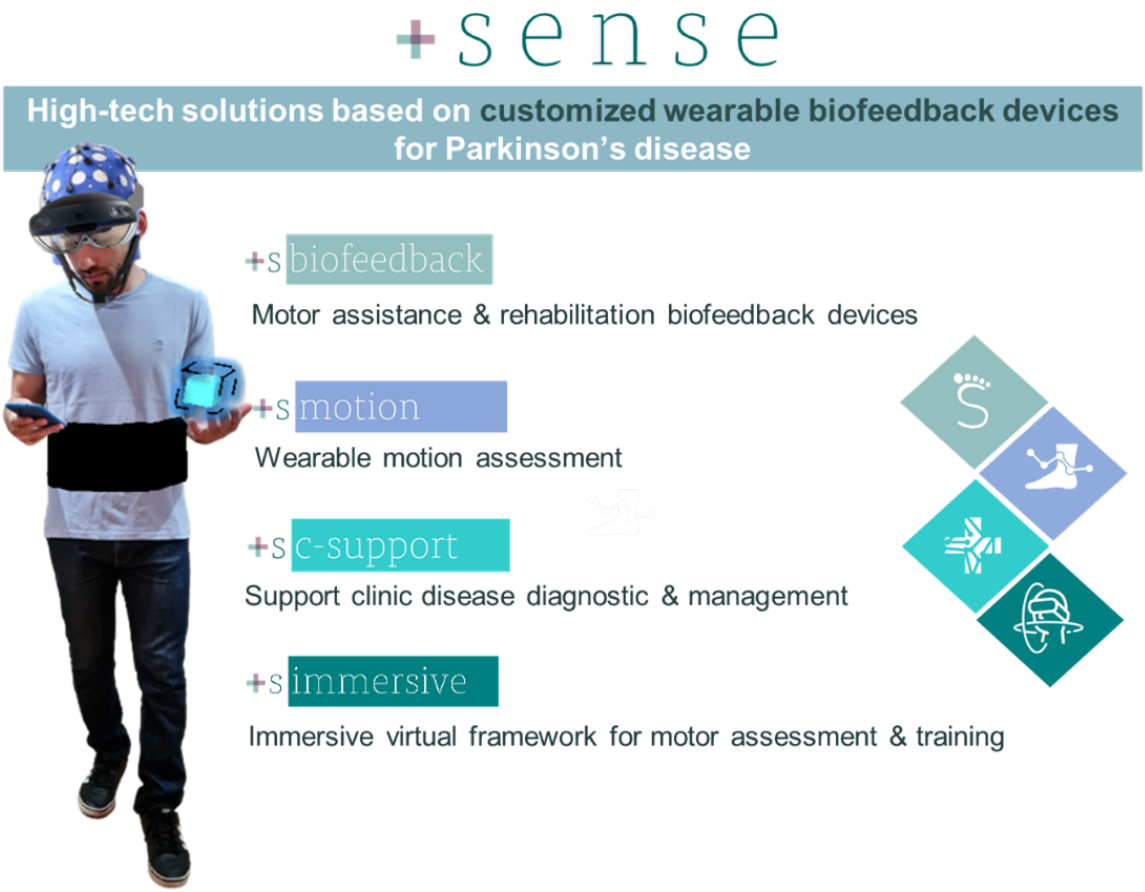


Figure 3: +Sense Project Description

3.3 +sMotion

The +sMotion module is responsible for acquiring and monitoring lower trunk inertial signals, providing real-time gait segmentation, post-processing gait analysis and gait-associated metrics estimation. This module uses the instrumented waistband which acts as the gait analysis LAB. The device comprises a 1) Sensory Acquisition Unit; 2) Processing Unit; 3) Data Storage Unit; 4) Mobile APP; and 5) +SDesktop GUI, as depicted in Figure 4

Figure 4: +sMotion Description

Sensory acquisition relies on the use of the MPU-6050 Inertial Measurement Unit to acquire acceleration and angular velocity data. The processing unit comprises a STM32F4-Discovery to receive the acquired data from the sensory acquisition unit and run in real-time a gait event detection algorithm based on heuristic rules with adaptive thresholds and ranges to segment a gait cycle from both legs into: initial contact (IC)/Heel-strike (HS), foot-flat (FF), mid-stance (MSt), final contact (FC)/toe-off (TO) and heel-off (HO). Acquired inertial data and identified events are saved in the Data Storage Unit, a On The Go (OTG) USB driver. The Mobile APP is an Android APP that wirelessly communicates with the processing unit, via Bluetooth, enabling to start/stop data acquisition, control operability settings and plotting the acquired data. +SDesktop GUI is an interface developed in MATLAB© able to read the data saved on the USB driver and estimate the gait-associated metrics. Given the AI-based algorithms were accomplished in python environments, this dissertation addressed the conversion of +sDesktop GUI to this environment. Thus, it was developed a new +sDesktop GUI also able to load, visualize and reprocess inertial data. +sCsupport module will complement this APP with the AI-based models.

This waistband advantageously uses a sensor capable of measuring an entire gait cycle. It also can be adapted to different people with different physical features and can be used under the patients' clothing, with the potential to be used in people's homes, thus gathering data from daily activities.

3.4 +sC-Support

+sC-support uses the outcomes measured with the instrumented waistband described in the previous +sMotion chapter to apply AI models able to accomplish a better PD management.

In this way, through a single sensor, on the patient's waist, it is possible to capture gait patterns in the inertial data. For example, it is expected that patients with an advanced motor disability can describe lower magnitudes of inertial data given their limitation on mobility [44]. This pattern can translate that these patients are in a more illness severe phase, measured by higher scores of UPRDS-III, requiring more medication, which when applied in AI are able to diagnose PD disease or even stratify its levels. In this way, +sC-support is able to complement physicians in the evaluation of patients.

This dissertation has an impact contribution to this module. An extensive statistical study was conducted to verify if gait metrics vary between patients and non-patients, and between different levels of UPDRS-III, PDQ-39 and H&Y. Next, various AI methods were applied in order to be able to obtain good results in distinguishing healthy from sick and the various levels of the UPDRS-III.

3.5 Conclusions

This thesis aimed to contribute to the +sense project. Specifically, it contributed to +sC-support with the help of +sMotion, more precisely, the instrumented waistband to use its wearable sensor.

For the +sC-support module, this dissertation contributed an extensive AI study to stratify and diagnose PD, using clinical scales such as UPDRS-III, PDQ-39 and H&Y.

4

Deep-Learning Frameworks

4.1 Introduction

In this section of the thesis, there will be an in-depth analysis of the development of the AI-based models. For this, a description as well as an exploration of the dataset is presented, followed by the processing of the input dataset. After this, the model training pipeline is explained, followed by a description about how the models were evaluated. The results of these evaluations are then presented for every AI-based model developed. Afterwards the results obtained are critically discussed and compared with the related state-of-the-art.

Both the dataset processing and the model training pipeline were developed using Python. The main libraries used for the dataset processing were Numpy, Pandas and Sklearn, while the package used for the development of the AI-based model was Tensorflow.

4.2 Data Preperation

4.2.1 Dataset Description

The input data of the proposed AI models contained gait measures of 40 idiopathic PD patients (21 male and 19 female, age: 66.83±9.52, height: 163.92±7.89, weight: 71.53±14.04,). This database also included measures of disease severity (i.e., H&Y, UPDRS-III, and PDQ-39 scales) for the PD patients. There were 16 patients with H&Y scores = 1, 15 patients with H&Y scores = 2 and 9 patients with H&Y scores = 3, mean: 1.83±0.87, as seen in Table . There were 11 patients with Low UPDRS-III scores, 15 patients with Mild UPDRS-III scores and 14 patients with High UPDRS-III scores, mean: 22.08±11.40. There were also 16 patients with High PDQ-39 scores, 14 patients with Mild PDQ-39 scores and 10 patients with Low PDQ-39 scores, mean: 37.34±22.55.

The acquired data was obtained by using the +sMotion module refered in 3.3. The subjects were asked to walk at their chosen pace for a distance of 10 meters. The output of the sensors consisted in 6 different features, 3 related to the accelerometer and 3 related to the gyroscope. Each of these 3

features represented the x, y and z axis. The accelerometer and gyroscope values gathered were divided by their resolutions (accelerometer resolution = 8192, gyroscope resolution = 65.5) in order to convert these values into their correct measurement units.

4.2.2 Dataset Exploration

After this, 3 different datasets were created. One with to be used for the classification of UPDRS-III, another for the classification of PDQ-39 and the last of for the classification of H&Y. Concluding this step, it was noticeable that every dataset was unbalanced, especially the H&Y and UPDRS-III datasets. So, all of these datasets needed to be balanced. Since the loss of data would most likely harm the results of the developed AI-based models by discarding some useful examples for the modeling of the classifier [45], the dataset would need to be oversampled. For this, the imbalanced-learn Python library, which contains the SMOTE(Synthetic Minority Oversampling Technique) technique, was used. This technique was used since it has been shown that SMOTE yields betters results for re-sampling [46]. SMOTE balances the data by over-sampling the minority class by taking each minority class sample and introducing new synthetic examples [47]. For this, an interpolation strategy is used to create these synthetic examples [48].

Figure 5 presents the observations per class of UPDRS-III, a total of 1046976 observations. It contains 221592 observations for the "Mild"class, 560224 observations for the "Moderate"class and 265160 observations for the "Severe"class, being observed an unbalanced dataset. To balance the number of observations for each class, the SMOTE technique was applied to the data. Thus, as observed in orange bars in Figure 5 a balanced dataset was obtained, with a total of 1680672 observations.

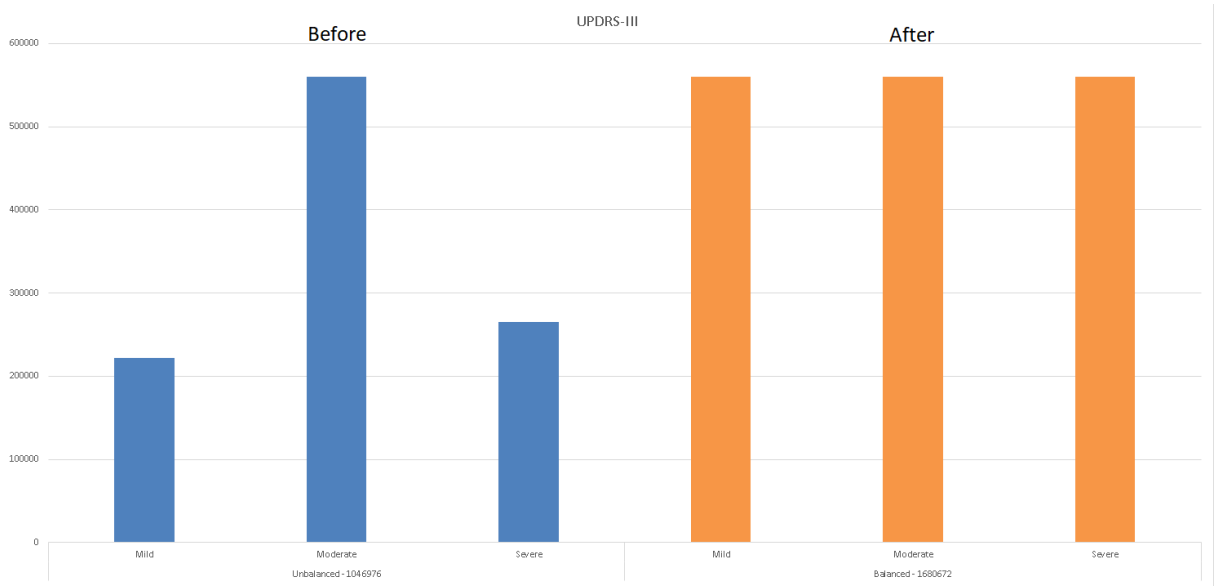


Figure 5: UPDRS-III Dataset - Blue columns represent before and orange columns represent after balancing

Figure 6 also presents an unbalanced dataset with 396144 observations for the "Mild"class, 512472 observations for the "Moderate"class and 129360 observations for the "Severe"class. To balance this

dataset, the SMOTE technique was again used on the data. Thus, as observed in orange bars, a balanced dataset was obtained, with a total of 1564416 observations.

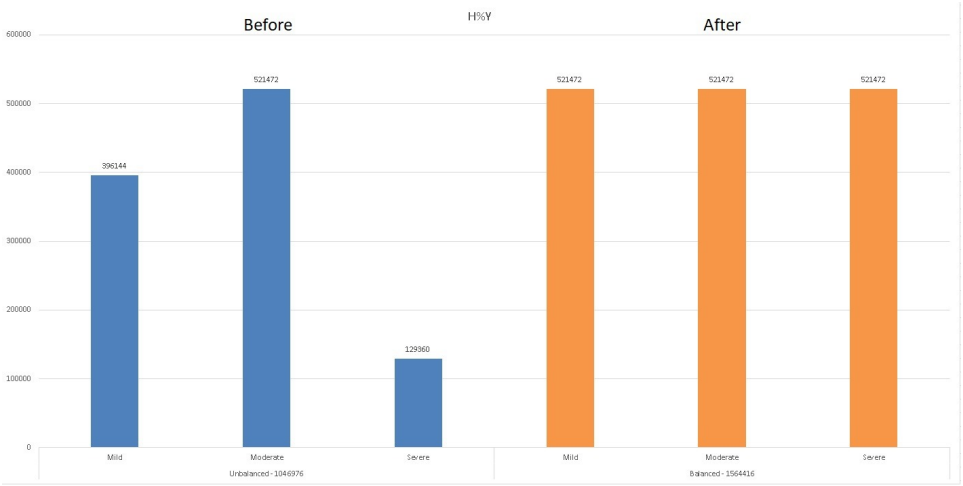


Figure 6: H&Y Dataset - Blue columns represent before and orange columns represent after balancing

Figure 7 once again shows an unbalanced dataset with a total of 325584 observations for the "Low"class, 372512 observations for the "Middle"class and 348880 observations for the "High"class. To balance this dataset, the SMOTE technique was once again used on the data. Thus, as observed in orange bars a balanced dataset was obtained, with a total of 1117536 observations.

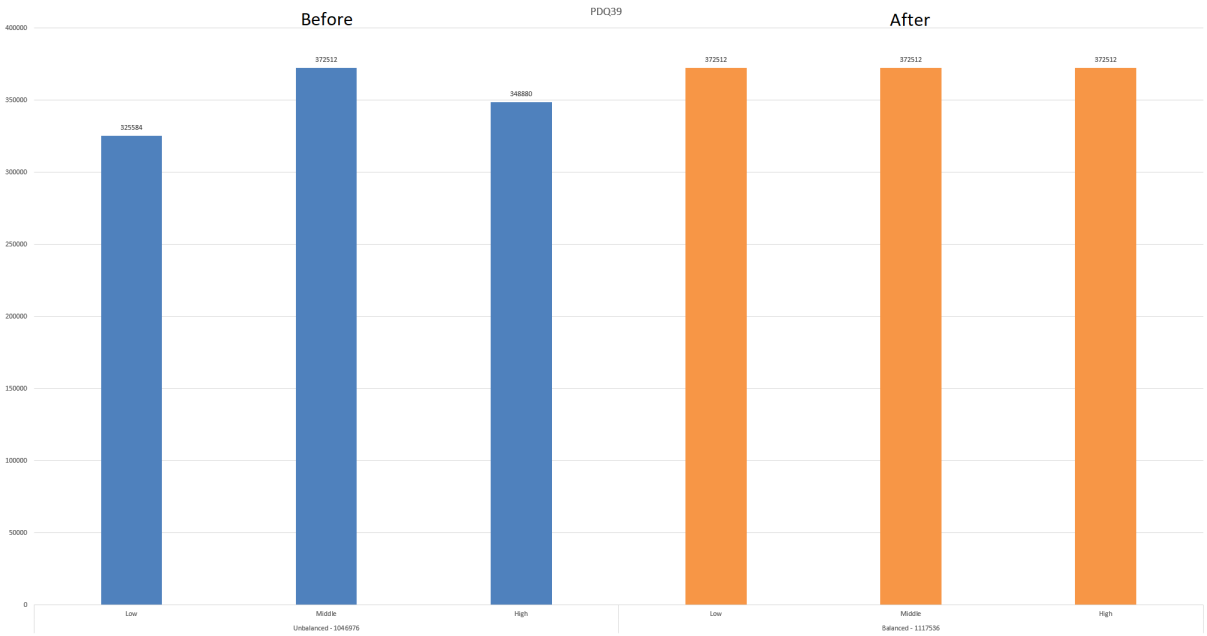


Figure 7: PDQ-39 Dataset - Blue columns represent before and orange columns represent after balancing

4.2.3 Data Transformation and Labelling

Given DL models were employed, it was necessary to transform the dataset from a 2D structure to a 3D structure. Thus, the shape of the dataset needs to be transformed from (number of observations, 6) into (number of observations, number of timesteps, 6). The original data was grouped into size windows of 56 samples, the mean step time previously measured in [41]. This originated a model data input with the shape (number of observations, 56, 6).

The dataset needs to be divided into 3 different datasets used for training, testing and validation. Forthis, the original dataset was randomly divided into a dataset with 20% of the data that was to be used to the testing phase, and another with the remaining 80%. The remaining data is going to be used for GridSearchCV hyperparameter tuning data and was once again randomly diveded into 2 more datasets, one used for the training of the models that has 80% of the data, and another used for the validation of the models. This division can be see in the Figure 8.

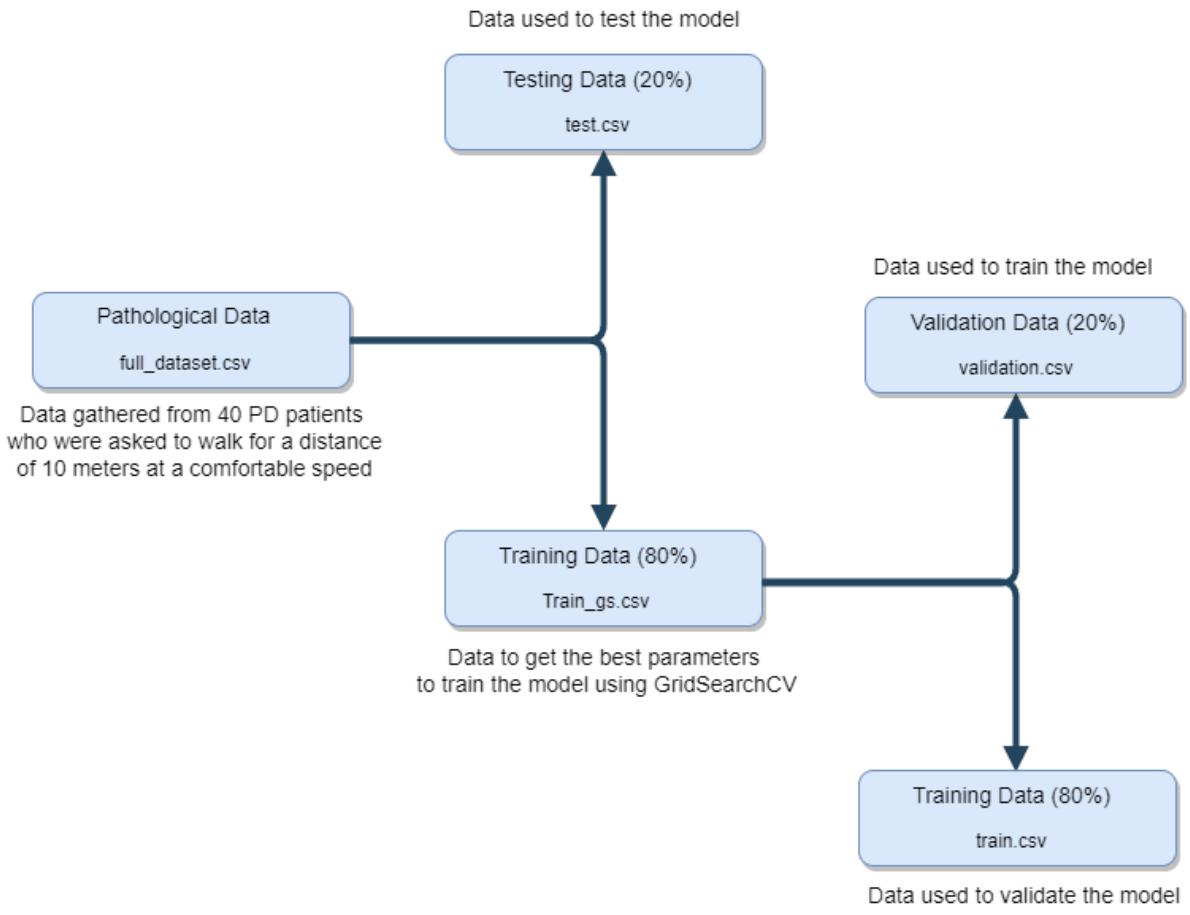


Figure 8: Dataset division into training, validation and testing datasets

Following this, Min-max normalization was applied to the data. This technique performs a linear transformation on the original data resulting in a scaled data in the range [0, 1].

4.3 DL Model Description and Hyperparameters

Recurrent Neural Networks (RNN) are neural sequences that achieve state-of-the-art performance on tasks such as language modeling, speech recognition, and machine translation [49]. Their designs resemble more or less a detailed analogy with biological brain modules. What distinguishes RNNs from other feedforward neural networks is the connection between the neurons possesses cycles. Thus, they obtain great results for time series processing applications [50]. However, RNNs are limited to look back in time for approximately 10 timesteps. This issue was addressed with LSTMs that are capable of learning more than 1000 timesteps [51] by enforcing constant error flow through constant error carousels (CECs) [52]. LSTMs can be used to complete tasks such as prediction, pattern classification, among other tasks. The ability of processing sequential data makes the LSTM an efficient tool in many fields, including medicine [53]. Advantageously, LSTMs are more efficient than other RNN algorithms [52].

Even though LSTMs show great performance, the training phase depends heavily on a set important hyperparameters such as batch size, number of neurons, learning rate [54], and dropout rate [55]. The number of epochs is also an important hyperparameter that needs to be tuned during the training phase. If not chosen carefully, the model may either overfit on the input data if the value is too high, or not reach its optimum performance if the value is too low.

4.4 Model Training Pipeline

The model chosen to be used on the development of this thesis is an LSTM. This model has an input layer with an input shape of (number of sequences, 56, 6) and an output shape of (number of sequences, 56, 6). This layer is followed by a LSTM layer with a input shape of (number of sequences, 56, 6) and output shape of (number of sequences, number of neurons). After this, a Droupout and Dense layers follow. Both of these layers have the same input and output shapes which is (number_of_sequences, number_of_neurons). The output layer is another Dense layer with an input shape of (number of sequences, number of neurons) and an output shape of (number of sequences, 3). The number 3 represents the different output possibilities.

This model configuration can be seen in Figure 9, and its output can be seen in Figure 10.

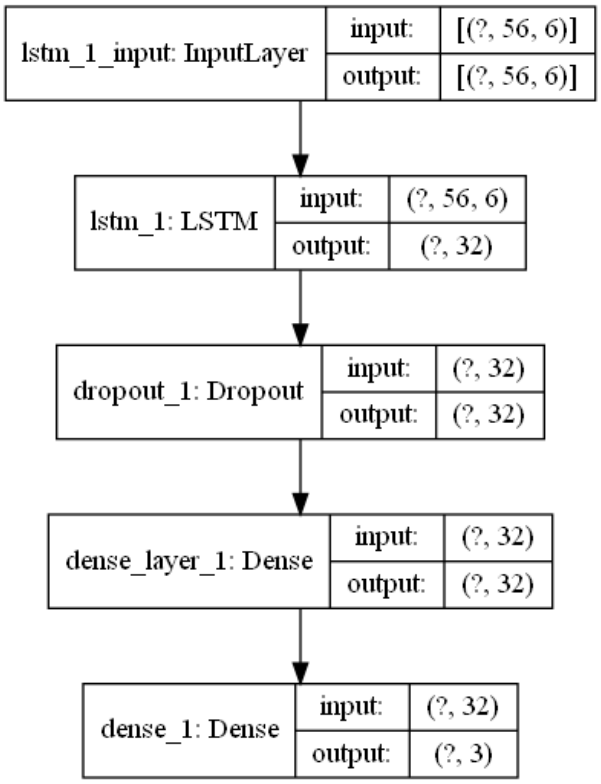


Figure 9: LSTM Configuration

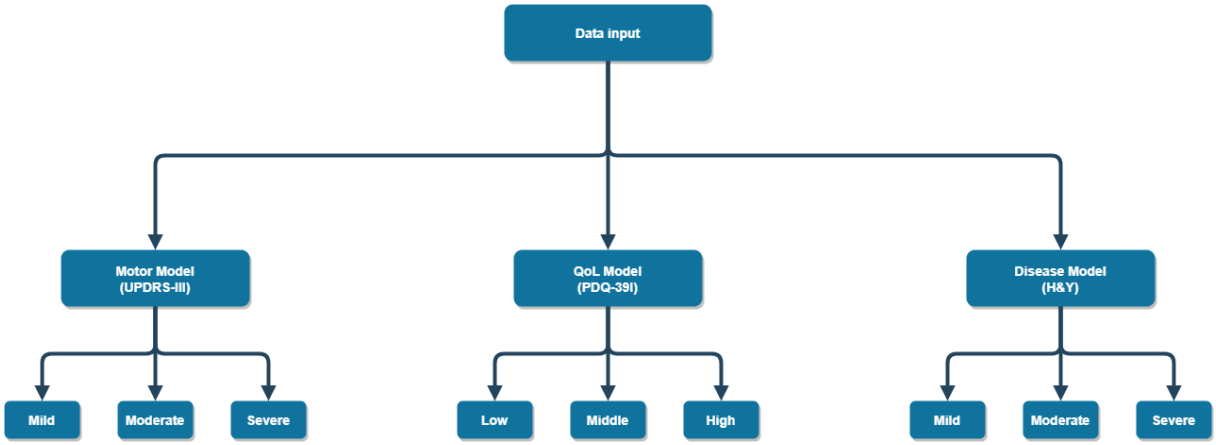


Figure 10: Model output

In order to optimize the results of this AI-based model, it was decided to use the GridSearchCV algorithm for hyperparameter tuning. GridSearchCV is a class provided by the scikit learning framework for adjusting the parameters applied by the estimator. The instructions usually define a dictionary to store the parameters that need to be looked up first, and then GridSearchCV will do all the necessary model adjustments and generate the best parameters. The parameters tuned were the neurons, which varied between 16, 32, 64 and 128, the learning rate which varied from 0.01 to 0.0001, the batch size whose two only options were 32 and 64, the epochs that could be 100, 150 or 200 and the dropout rate, that

varied between 0.2 and 0.35.

4.5 Model Evaluation and Loss

The performance metrics used for assessing the classifier efficacy are accuracy, specificity, recall (R) which is the same as sensitivity [28], precision (P), area under curve (AUC) and f1-score. Accuracy is the simplest and most common measure to assess the classifier and it is defined as the degree of correct predictions of the model based on true positive (TP), true negative (TN), false positive (FN) and false negative (FN) values [16].

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

The sensitivity is expressed as the ratio of true positives [31] and is defined as follows [56]:

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

The specificity is expressed as the ratio of true negatives defined as follows [35] :

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

Precision refers to the number of TP divided by the total number of positive. Model precision is the probability a predicted true label is indeed true and is defined can be defined as follows [57]:

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

The f1-score combines the precision and recall of a classifier into a single metric by taking their harmonic mean and it is defined as [58]:

$$F1 - score = \frac{2*(P*R)}{P+R}$$

The AUC can be interpreted as the ability of the model to separate classes [58].

The metric used to choose the model with the best results in the GridSearchCV hyperparameter tuning phase was the accuracy, since it is the most common metric found in related state-of-the-art articles, as seen in the metrics used by [59], [34] and [15].

4.6 Results

4.6.1 Training Performance

Furthermore, the TimeSeriesSplit cross-validator was used to apply a 10-fold cross-validation on the GridSearchCV hyperparameter tuning training data.

4.6.1.1 UPDRS-III

For the training and hyperparameter tuning of the UPDRS-III model, the results of the LSTM analysis are shown in the following Table 5. In this Table is shown the number of epochs used, and the values of the different tuned hyperparameters for training, and resulting Loss, Accuracy (Acc), Precision (Prec), Sensitivity (Sens), Specificity (Spec), F1-score and AUC.

Table 5: UPDRS-III AI-based model train results

DL Model	Epochs	Hyperparameters		Step	Loss	Acc	Prec	Sens	Spec	F1	AUC
LSTM	150	Batch size Neurons Dropout rate Learn rate	32 64 0.2 0.001	Train	0.1890	92.73%	93.88%	91.73%	94.75%	92.77%	99.0%
				Validation	0.2380	91.56%	92.81%	90.33%	93.96%	91.54%	98.43%

A plot of the obtained accuracy and loss of the developed model can be seen in the figures 11 and 12 shown below.

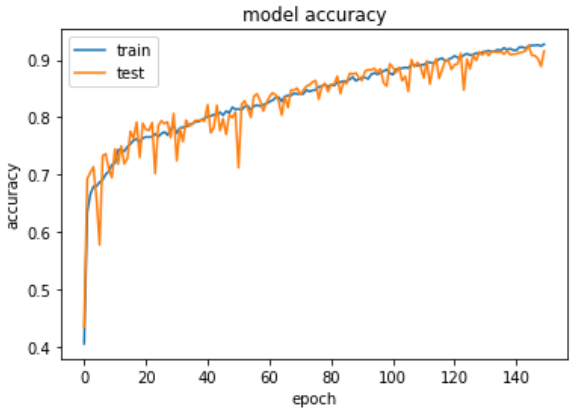


Figure 11: UPDRS-III Model Accuracy Plot

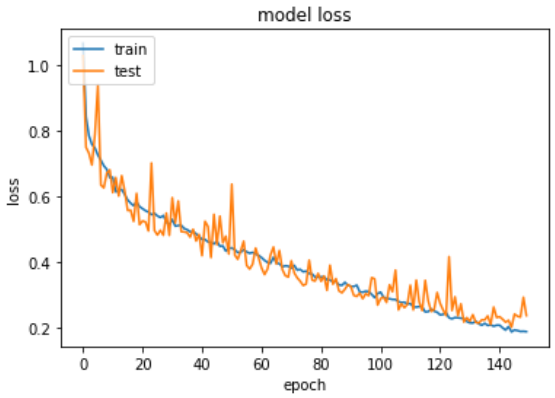


Figure 12: UPDRS-III Model Loss Plot

A confusion matrix testing the model performance with the validation data can be seen in Figure 13

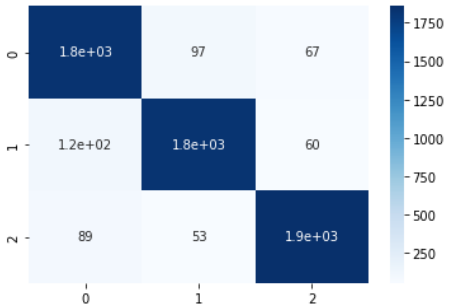


Figure 13: UPDRS-III Validation Matrix

4.6.1.2 H&Y

For the training and validation of the H&Y model, the results of the LSTM analysis are shown in the following Table 6. In this Table is shown the number of epochs used, and the values of the different tuned hyperparameters for training, and resulting Loss, Accuracy (Acc), Precision (Prec), Sensitivity (Sens), Specificity (Spec), F1-score and AUC.

Table 6: H&Y AI-based model train results

DL Model	Epochs	Hyperparameters		Step	Loss	Acc	Prec	Sens	Spec	F1	AUC
LSTM	100	Batch size Neurons Dropout rate Learn rate	32 128 0.2 0.001	Train	0.2490	90.27%	90.65%	89.87%	92.82%	90.25%	98.16%
				Validation	0.2908	89.05%	88.92%	88.63%	92.78%	88.92%	97.68%

A plot of the obtained accuracy and loss of the developed model can be seen in the figures 14 and 15 shown below.

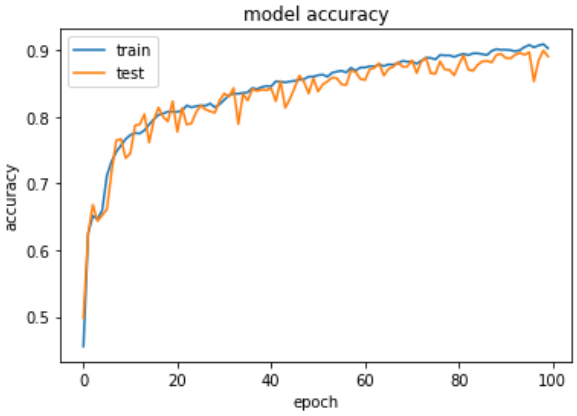


Figure 14: H&Y Model Accuracy Plot

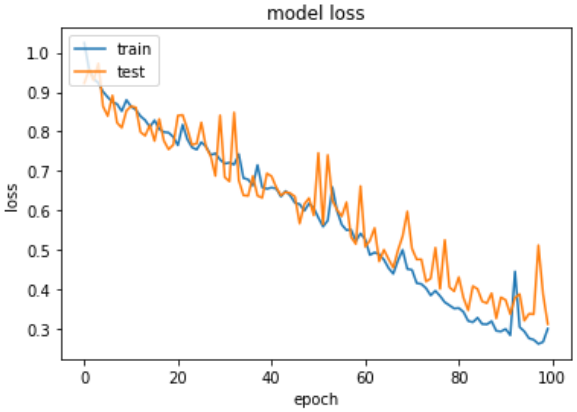


Figure 15: H&Y Model Loss Plot

A confusion matrix testing the model performance with the validation data can be seen in Figure 13

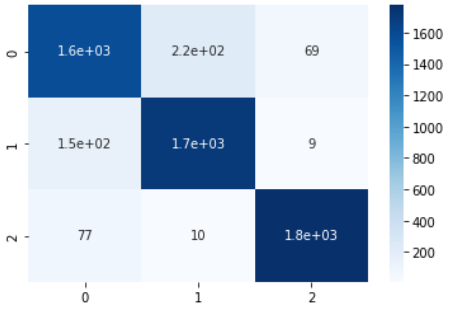


Figure 16: H&Y Validation Matrix

4.6.1.3 PDQ-39

For the training and validation of the PDQ-39 model, the results of the LSTM analysis are shown in the following table 7. In this Table is shown the number of epochs used, and the values of the different tuned hyperparameters for training, and resulting Loss, Accuracy (Acc), Precision (Prec), Sensitivity (Sens), Specificity (Spec), F1-score and AUC.

Table 7: PDQ-39 AI-based model train results

DL Model	Epochs	Hyperparameters		Step	Loss	Acc	Prec	Sens	Spec	F1	AUC
LSTM	100	Batch size Neurons Dropout rate Learn rate	32 128 0.2 0.001	Train	0.3002	87.86%	89.92%	85.63%	91.48%	87.71%	97.56%
				Validation	0.3121	88.0%	89.68%	86.03%	91.90%	87.71%	97.38%

A plot of the obtained accuracy and loss of the developed model can be seen in the figures 17 and 18 shown below.

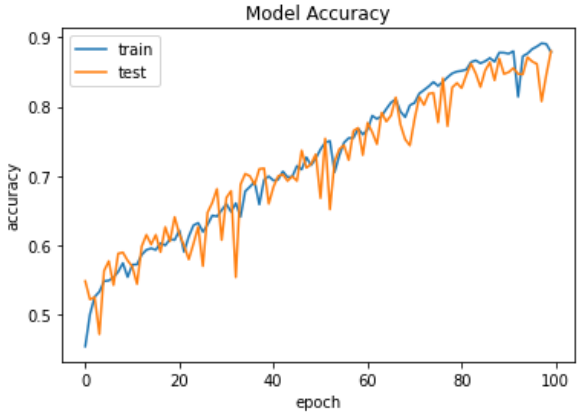


Figure 17: PDQ-39 Model Accuracy Plot

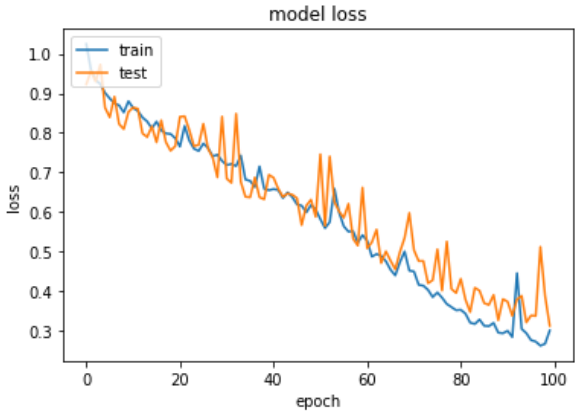


Figure 18: PDQ-39 Model Loss Plot

A confusion matrix testing the model performance with the validation data can be seen in Figure 19

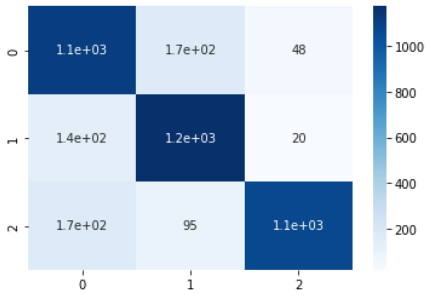


Figure 19: PDQ-39 Validation Matrix

4.6.2 Testing Evaluation

The next step after the training of the developed models, consists on using the test dataset to evaluate the performance of said models. For this, we used the evaluate method and also developed a confusion matrix for each of the models. A confusion matrix provides much more detailed information on the results of the test than the mere accuracy or loss [60]. The matrix shows which classes have been confused with which during the test. The obtained results can be seen in the following sections.

4.6.2.1 UPDRS-III

In the table below 8, the results of the evaluation of the UPDRS-III model can be seen.

Table 8: UPDRS-III AI-based model test results

Metrics	Test Results
Accuracy	91.67%

Continued on next page

Table 8 – *Continued from previous page*

Paper	**Goal**
Loss	0.2384
F1-score	91.58%
Precision	92.87%
Sensitivity	90.44%
Specificity	93.99%
AUC	98.41%

The following figure 20 shows the confusion matrix of the testing of the UPDRS-III model.

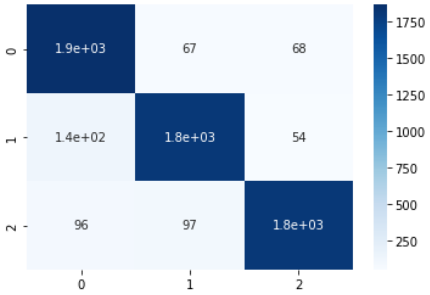


Figure 20: UPDRS-III Confusion Matrix

4.6.2.2 H&Y

In the table below 9, the results of the evaluation of the H&Y model can be seen.

Table 9: H&Y AI-based model test results

Metrics	Test Results
Accuracy	88.98%
Loss	0.2844
F1-score	89.05%
Precision	89.40%
Sensitivity	88.71%
Specificity	92.79%
AUC	97.72%

The following figure 21 shows the confusion matrix of the testing of the UPDRS-III model.

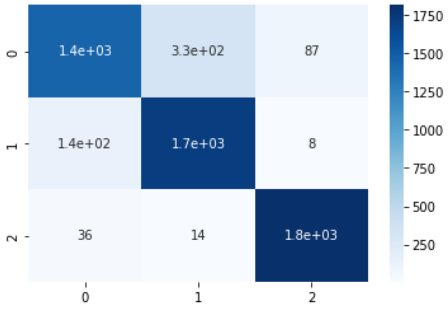


Figure 21: H&Y Confusion Matrix

4.6.2.3 PDQ-39

In the table below 10, the results of the evaluation of the H&Y model can be seen.

Table 10: PDQ-39 AI-based model test results

Metrics	Test Results
Accuracy	87.80%
Loss	0.3175
F1-score	87.60%
Precision	89.53%
Sensitivity	85.85%
Specificity	91.85%
AUC	97.33%

The following figure 22 shows the confusion matrix of the testing of the UPDRS-III model.

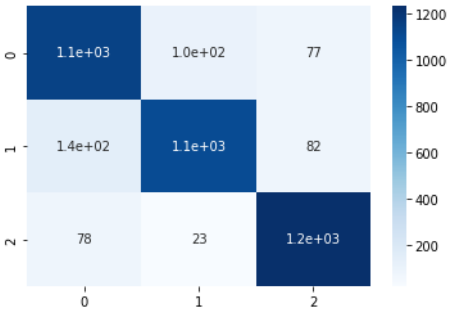


Figure 22: PDQ-39 Confusion Matrix

4.7 Discussion

In the case of the H%Y scale, an accuracy of 88.98%, a sensitivity of 88.71%, and a specificity of 92.79% was obtained. As we can see in [16] an overall accuracy of 99.4%, a sensitivity of 98.2, and a specificity of 99.3% was achieved with the use of a Decision Tree classifier. The H&Y AI-based model shows slightly worse results compared to [16], with an accuracy of 91.67% against an accuracy of 99.4%. However, [16] used a dataset that contains measures of gait from 93 patients with idiopathic PD. This means that a dataset with double the size of the dataset used for the development of this thesis was used. And as previously stated, the bigger the amounts of data, the better performance the developed models will have. In other words, if the dataset used for the development of this thesis had bigger amounts of data, the performance of the UPDRS-III model would most likely show better results, even though the results

obtained are already significant. Also, the data used in [16] consisted of vertical ground reaction force records of subjects, using 8 sensors under each foot, while the data used for the development of this thesis included data of a single IMU. This results in larger ammounts of data, which results in better performance for the models. Even more, the models used in [16] were Machine Learning models, which means a greater amount of data processing needed to be done, while DL models can directly receive raw data as input, reducing computational power.

In the case of the PDQ-39 scale in which we had an accuracy of 84.19%, a sensitivity of 82.13%, and a specificity of 90.24%, there was no data found in review so that our results can be compared to others.

The UPDRS-III AI-based model has the best overall results, followed by H%Y AI-based model, which is followed by the PDQ-39 AI-based model. The decrease of the accuracy in these models can be justified by the size of the dataset used for each model. Even though the initial dataset contained data from 40 patients, not all resulting datasets for each model had the same balancing for the label column. This means that when the process of balancing the datasets occurred, each dataset gained new amounts of data depending on how unbalanced it previously was. As we can see in figures 5, 6 and 7, the most unbalanced dataset was the dataset used for the UPDRS-III AI-based model, followed by the dataset used for the PDQ-39 AI-based model, which is followed by the H&Y AI-based model. Thus resulting in bigger amounts of data for the UPDRS-III dataset, followed by the PDQ-39 dataset, and lastly the H%Y dataset, after the balancing technique was performed.

Even though the results shown are promising, more testing needs to be done to improve the performance of these models. The dataset used for the development of this thesis contains data from 40 different patients. If larger amounts of data were accessible, the results obtained would most likely improve. Also, different types of models should be tested. As seen in [31] and [35], models like the SVM show promising results.

4.8 Conclusions

The dataset for each model was divided into two smaller datasets, one which contained 80% of the data and was used for training the model, and a smaller dataset which contained 20% of the data and was used for testing the model performance.

The results for each model showed slightly better performance during the training process than during the testing phase, which means that was a very slight and insignificant overfit. The UPDRS-III model achieved a test accuracy of 91.67%, a sensitivity of 90.44% and a specificity of 93.99%, while the PDQ-39 model achieved a test accuracy of 84.20%, a sensitivity of 82.14%, and a specificity of 90.24%, while the H&Y model achieved a test accuracy of 78.92%, a sensitivity of 74.89%, and a specificity of 86.96%.

The UPDRS model produced slightly worse results than the ones seen in [16], while the H&Y model produced better results than the ones seen in [32], and no results were found for PDQ-39. These slightly worse results in the UPDRS model can be justified by the lower amounts of data present in the dataset

used for this thesis compared to the dataset used by [16]. Thus, if the amounts of data present in each of the dataset used for these models were higher, the results obtained would most likely be better.

5

APP

5.1 Introduction

The goal of this thesis is to develop AI-based models capable of classifying patients' quality of life, motor conditions and disease level. These AI-based models are implemented into an intuitive APP. This APP is to be used by physicians and related researchers to help with PD management. This APP will be helpful to physicians because, by combining data of every day activities of the patients acquired by the +sense device, a wearable sensory system,an extra objective assessment and smart classification can be achieved. This chapter will explain how to use the user-friendly APP that was developed.

The APP is divided intro 5 different tabs, which are the Start, Activity, Gait Analysis, Metrics and AI Report tabs. In the next section, the first first tab will be introduced and there will be an explanation on how to use this tab and its objectives. The fifth and final tab will have its own section for an in-depth analysis of the development of this tab.

5.2 User-friendly APP for identification of digital biomarkers of PD based on recognized gait patterns

The developed APP starts with a log-in screen where the users can enter their *username* and respective *password*.

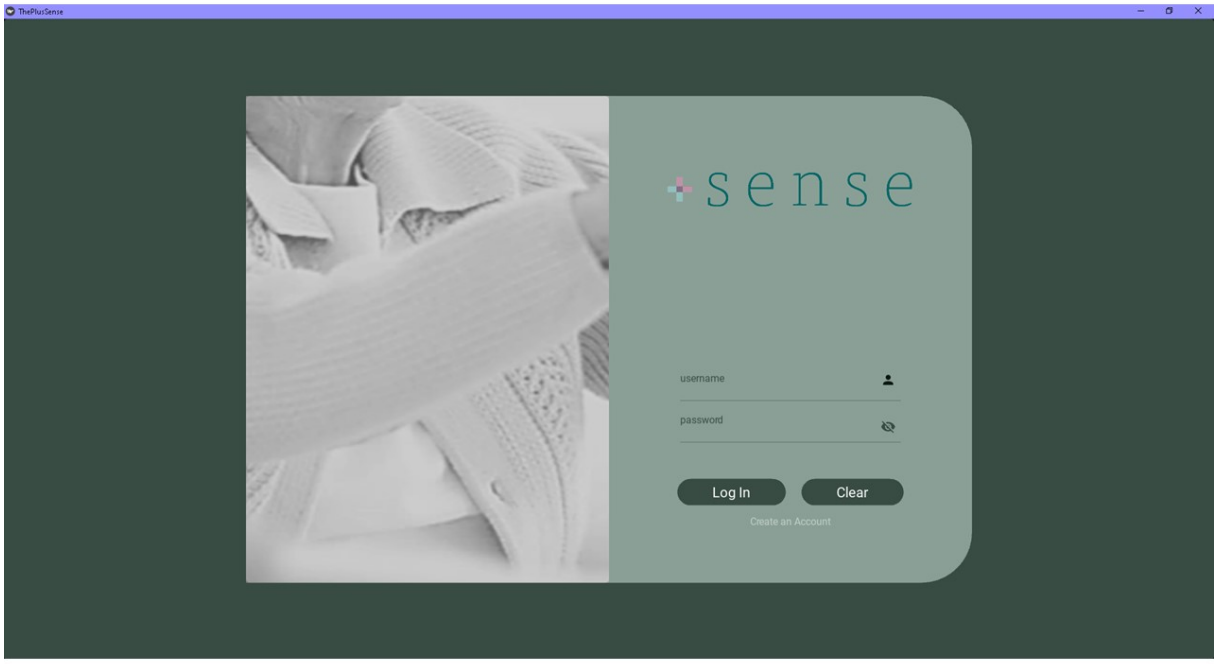


Figure 23: +Sense APP Log-in screen

After this process, the user will access the Start screen, as seen in Figure 24. This part of the APP is used to indicate the sociodemographic patients' data, such as name, age, gender, weight and height. Also, it is possible to add some observations. In this screen there is also a button that enables to load data files from the patients' motor acquisitions, as well as select the type of activity that was done, such as getting up from and sitting on a chair, laying on and getting up from a bed, walking, 180° turns, 90° right and lefts turns and finally the pull test. In this dissertation, it was used always the data from walking, but for future applications is it possible to select the acquired motor activity. The loaded data are then used in the remaining APP tabs.

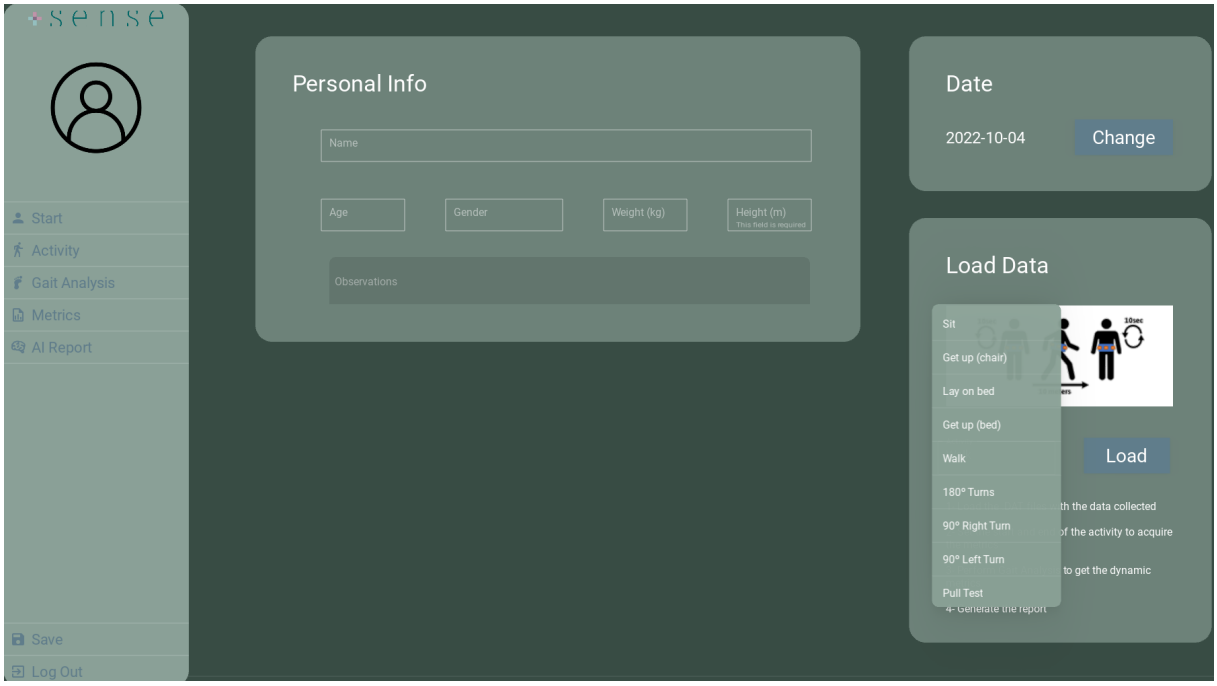


Figure 24: +Sense APP Start screen

The following APP window is the Activity window. In this screen, the users can use the loaded data to plot the acceleration and angular velocity of the patient's data on the x, y and z axis by using the "Plot"button present in this tab. Also, the user can change the start and stop points to analyze the signal in the following tabs.

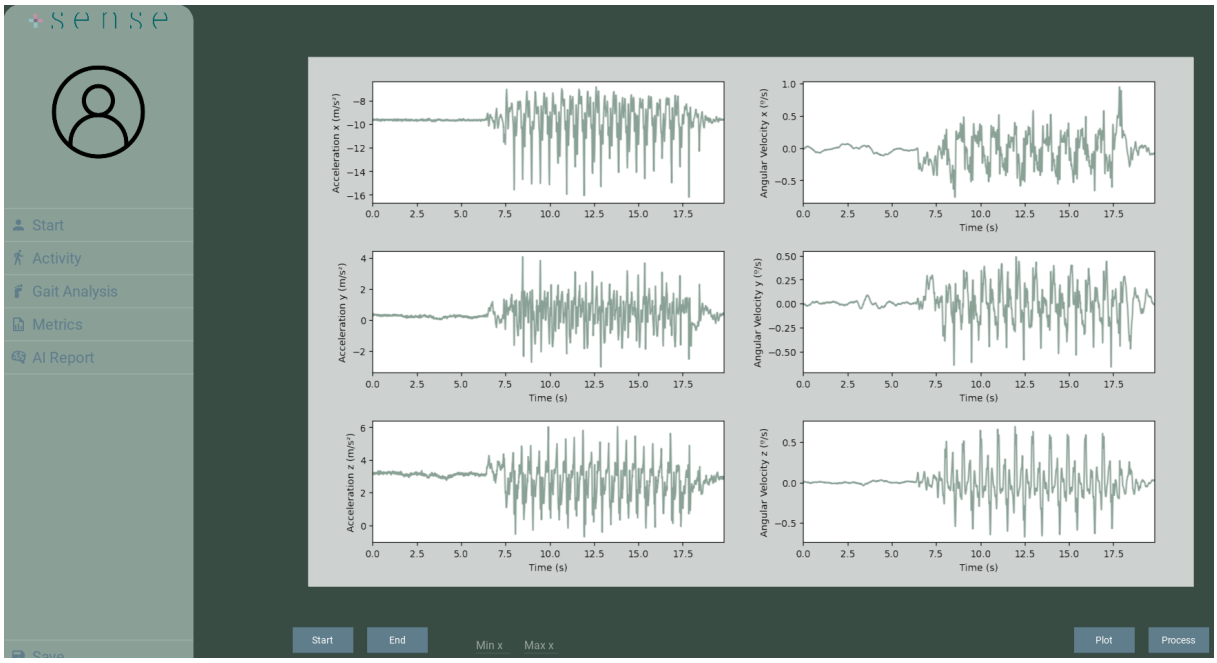


Figure 25: +Sense APP Activity Plot

If the user wishes to take a close look at a smaller sample of the patient's data, he or she can limit

the time window in which she wants to visualize the data by filling out the "Min x"and "Min y"spaces.

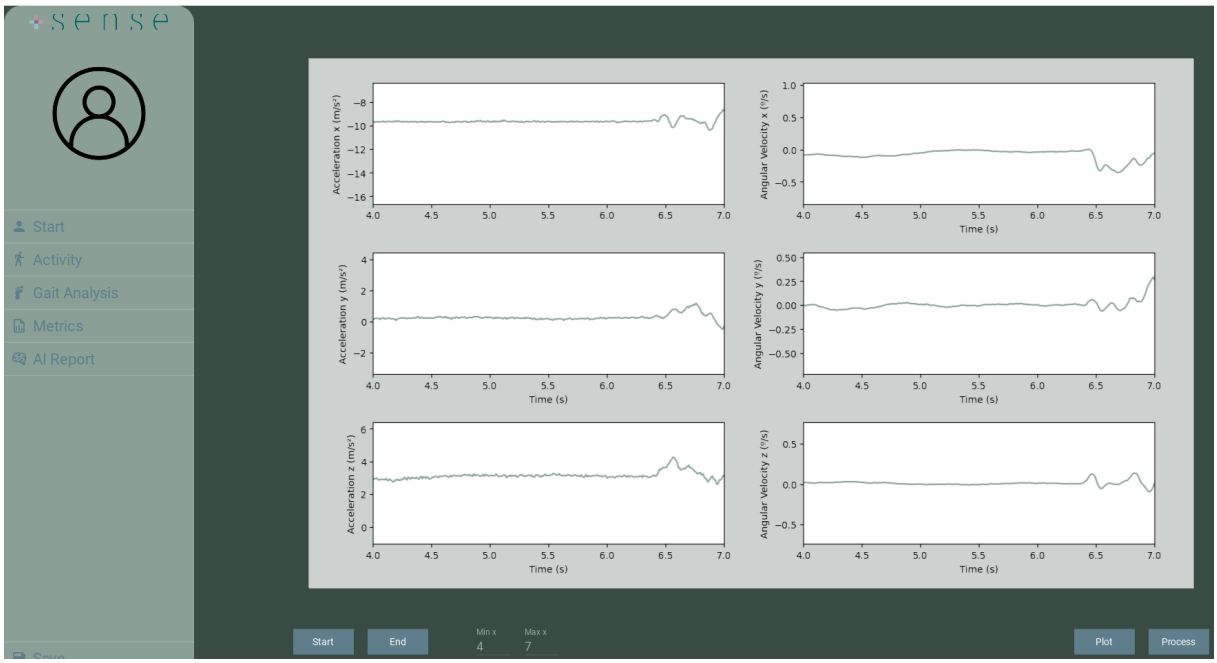


Figure 26: +Sense APP Activity time window limit

One more thing useful in this screen is the Start, End and Process buttons. These buttons allow the physician to choose a specific window of time, using the Start and End buttons and selecting the timestamps desired as seen in Figure 27, and process the data in said window.

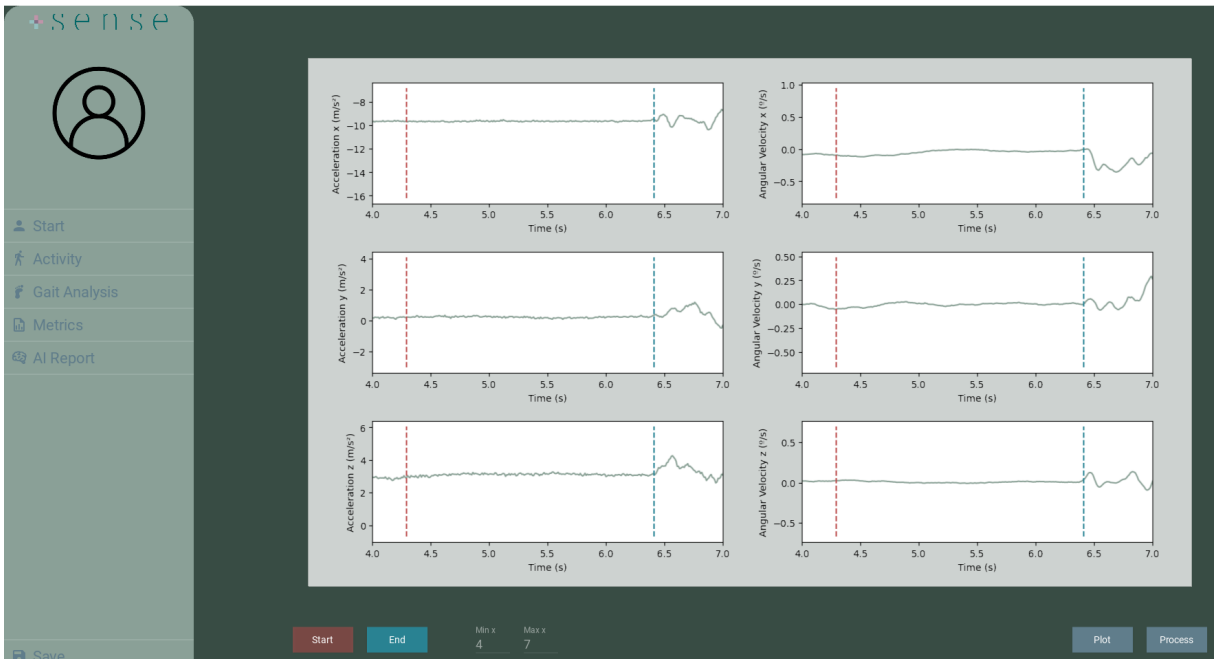


Figure 27: +Sense APP Activity Process

5.3 Integration of AI in APP

The data previously loaded in the Start tab are used for this part of the APP. The first step consisted in grouping the data in windows of 56 rows, followed by the Min-Max normalization. Once the data is loaded, the developed AI-based models need to be loaded as well in order to classify the patient in the aforementioned scales (UPDRS-III, PDQ-39, H&Y).

Once everything needed is loaded, the APP proceeds to feed the loaded and prepared data from the patient to the 3 different AI-based models. All windows from the input data were classified, and the mode is used to retrieve the most frequent classification. These classifications are mapped to their respective values in each scale, as indicated in Table 11.

Table 11: Classification Mapping

Value	UPDRS-III	H&Y	PDQ-39
0	Mild	Mild	Low
1	Moderate	Moderate	Middle
2	Severe	Severe	High

These classifications are then updated in the APP as seen in Figure 28.

Figure 28: +Sense APP AI Report Tab

5.4 Conclusions

An APP was developed to aid physicians with the diagnosis, monitoring and management of PD.

First, the data is loaded and can afterwards be graphically visualized and analyzed. Then, the developed and implemented AI-based models are loaded and, using the loaded data, a classification on the UPDRS-III, H&Y and PDQ-39 scales is made. These results are automatically updated and displayed on screen.

6

Conclusions and Future Directions

PD is a degenerative disease of the central nervous system, characterized by causing several disabling motor symptoms associated with the mobility of patients. It is often associated with a vast list of gait-associated disabilities, for which there is still a limited pharmacological/surgical treatment efficacy. These disabilities considerably increase the risk of fall, limit the quality of life and autonomy of the patients, who become dependent on third parties for the most trivial and daily activities. Bradykinesia, shortened stride length, shuffling gait and freezing of gait are some of these prototypical gait-associated signs in PD.

With the aim of monitoring these gait associated symptoms, physicians typically use clinical assessment scales based on observational data and surveys on the well-being of patients. However, these tools provide non-objective data. During consultations, physicians are limited to the data indicated by the patient, so objective and continuous metrics of the patient's mobility cannot be gathered.

With the help of IMUs, objective and continuous data can be obtained. These sensors are typically small in size, easily portable, low-cost and have low power consumptions. These traits allow their usage in everyday activities without interfering with the patient's mobility. Thus, these sensors can capture long-term data about the true stage of thei patient's condition. The use of IMUs also eliminates the patient's bias, resulting in the acquisition of more objective data.

Taking all this in mind, the goal of this thesis is to develop an AI-based model which is able to automatically classify Parkinsonian gait patterns, using data acquired by wearable sensory systems.

This thesis was separated into four major parts: The first phase of this thesis aimed to complete an intensive and extensive research on PD, with great emphasis on its typical gait patterns. During this phase, a critical review of the technologies based on AI for automatically classifying patients' disease stage will also be carried out. The goal of this review was to identify the achievements and limitations of current studies in order to contribute in an innovative way to the scientific panorama. The second phase of this thesis covers the preparation of a dataset based on acquisitions of PD patients' gait with wearable sensors to feed the expected AI models. The third phase of this thesis consists on the development of AI-based models capable of classifying PD patients in terms of their motor diasbility, ilness degree and quality of life leves using the UPDRS-III, H&Y and PDQ-39 scales to identify the classes. To this end, the training of several AI models with the aim of selecting the one with the best performance was performed. For this,

a model training pipeline was created. Each model was fed either the UPDRS-III, H&Y or PDQ-39 dataset and then GridSearchCV was used for hyperparameter tuning. The fourth phase of this thesis consisted in the development of a user-friendly APP that integrates the models with the best performance, aiming to be used by physician and researchers in the area.

A review about the use of AI models to classify PD gait patterns was carried out. With it came in-depth knowledge about PD gait patterns and gait impairments in other diseases such as ALS and Huntington's Disease. These impairments are usually identified using either an IMU or a camera motion analysis system. The most used IMUs are accelerometers and gyroscopes and some other common sensors used are foot pressure sensors. Various reviewed papers used SVM, Support Vector Regression, Naïve-Bayes, Logistic Regression and ANN algorithms to develop AI models. These models are usually implemented using either MATLAB or Python. Taking this into account, some limitations were found which drive the following steps of the next dissertation tasks. This segment of the thesis is explained in Chapter 2.

This dissertation is part of a project research titled by +SENSE: Sensory biofeedback devices for patients with Parkinson's Disease, which objective is to develop high-tech solutions to reduce motor symptoms in PD. The +Sense project aims to improve patients' quality of life. For this, the main goal of this dissertation is to contribute to the project with the development of a clinical decision support tool, that aims to help physicians examinations about patients' motor conditions with more objective and reliable data. In this tool, there will be various AI models that can accomplish PD management. The data fed to the AI-based models was gathered using a wearable motion +SENSE device, which is an instrumented waisband that has integrated an IMU. The used dataset is made of data from 40 patients with PD who were asked to walk at comfortable speeds three times for a distance of 10 meters. The patients were also assessed considering their (i) disease level, using H&Y scale; (ii) motor condition assessed by UPDRS-III; and (iii) QoL level using PDQ-39. Everything related to the +sense project is shown in Chapter 3.

With the help of the programming language Python, the aforementioned gathered data was processed so it could be fed to the AI models. The main dataset was divided into 3 different datasets, one to be used for the model able to classify a patient on the UPDRS-III scale, another to be used for the model able to classify a patient on the H%Y scale, and the last to be used for the mode able to classify a patient on PDQ-39 questionnaire. It was necessary to apply the synthetic minority over-sampling technique (SMOTE) to balance the input data, and Min-Max normalization was also applied. This processed data served then as input to the 3 different LSTM models. The models were then trained, and GridSearchCV was used for hyperparameter tuning. The UPDRS-III model results were slightly worse than the ones found in the state-of-the-art research. It achieved a test accuracy of 91.67%, a sensitivity of 90.44% and a specificity of 93.99%, which can be improved by including more data. The H&Y model achieved a test accuracy of 88.98%, a sensitivity of 88.71%, and a specificity of 92.79%. The H&Y model produced much better results than the related state-of-the-art outcomes. The PDQ-39 model achieved a test accuracy of 84.20%, a sensitivity of 82.14%, and a specificity of 90.24%. To the best knowledge, no studies were found on the development of an AI-based models to classify patients on the PDQ-39 questionnaire, which reveals the innovative character of this model. These results can be seen in Chapter 4.

The final step of this thesis, consisted on the development of an intuitive clinical APP, in which the previously developed models were integrated. In this APP it is possible to load data gathered from patients' motor acquisitions. This data can then be graphically visualized and analyzed. But more importantly, using the loaded data and LSTM models, a classification, of the patients' data, on the UPDRS-III, H&Y and PDQ-39 scales is made. The obtained results are afterwards updated automatically on screen. This part of the thesis can be seen in Chapter 5

The work herein presented enables to answer the Research Questions outlined in Chapter 1:

- **RQ 1:** Which scales allow a more complete assessment of the patient?

  From a motion condition standpoint, the UPDRS is the gold standard for motor measurement in PD and has been used worldwide for clinical management and research [61], and the most frequently used global assessment for PD is the H%Y scale [62]. Thus, the main focus of this thesis was to classify automatically patients in these scales.

- **RQ 2:** Which and how many sensors are used, and where are these sensors placed on the patient's body?

  As seen in Section 2.3.3, the number of sensors varies from article to article, between 2 [32] and 22 [31]. However, the smallest number of sensors possible is more advantageous, so that their use is not intrusive in the daily activities of the patients. These sensors are commonly placed in the feet [15] and legs [35], but should as well be placed in places where it won't affect the daily activies of the patients.

- **RQ 3:** Which AI-based models produce the best results?

  The model with the best performance for classifying PD patients on the UPDRS-III scale was a Decision Tree, with a 99.4% accuracy, 99.6& sensitivity and 99.8% specificity [16], while the best model for classifying PD patients on the H&Y scale was an ANN with a 66.16% accuracy, a 66% sensitivity, and a specificity of 85%.

- **RQ 4:** How to integrate the previously developed models in the APP?

  These models (architecture and weights) can be saved after being trained and obtaining the pretended results. Succeeding, these models can also be loaded using the saved architecture and weights. After the models are loaded, they can be used to classify the patients on the UPDRS-III, H&Y and PDQ-39 scales.

Hereupon, it is concluded that the delineated goals and RQs raised in the introduction of this thesis were addressed in Chapter 2. It was developed 3 different AI-based models, using an LSTM, capable of automatically classifying a patient on the UPDRS-III, H&Y, and PDQ-39 scales. These models were integrated into an intuitive clinical APP capable of helping physicians on the PD assessment of patients.

6.1 Future Work

Some suggestions for future research and improvements were raised during the development of this dissertation:

- Perform acceptability and usability tests. For this, a testing phase of the developed APP should take place in order to collect opinions from physicians about its usability and acceptability.
- Utilize larger amounts of data to train the AI-based models with the aim of improving their performance. To achieve this, more data using the +sMotion module should be acquired from patiens with varying types of PD stages.
- Optimize the hyperparameter tuning phase by widening the range of values for each hyperparameter.
- Experiment with different types of models, and analyze which obtains the best performance.

Bibliography

[1] S. M. Kim et al. “Gait Patterns in Parkinson’s Disease with or without Cognitive Impairment”. In: *Dementia and Neurocognitive Disorders* 17.2 (July 2018), pp. 57–65. doi: 10.12779/DND.2018.17.2.57. url: https://synapse.koreamed.org/articles/1120132 (cit. on p. 1).

[2] H. R. Gonçalves, A. M. Rodrigues, and C. P. Santos. “Vibrotactile biofeedback devices in Parkinson’s disease: a narrative review”. In: *Medical and Biological Engineering and Computing* 59.6 (2021), pp. 1185–1199. issn: 17410444. doi: 10.1007/s11517-021-02365-3 (cit. on p. 1).

[3] D. De Bartolo et al. “Effect of different music genres on gait patterns in Parkinson’s disease”. In: *Neurological Sciences* 41.3 (2020), pp. 575–582. issn: 15903478. doi: 10.1007/s10072-019-04127-4 (cit. on p. 1).

[4] A. Mirelman et al. “Gait impairments in Parkinson’s disease”. In: *The Lancet Neurology* 18.7 (2019), pp. 697–708. issn: 14744465. doi: 10.1016/S1474-4422(19)30044-4. url: http://dx.doi.org/10.1016/S1474-4422(19)30044-4 (cit. on pp. 1, 3).

[5] E. Rovini, C. Maremmani, and F. Cavallo. “How wearable sensors can support parkinson’s disease diagnosis and treatment: A systematic review”. In: *Frontiers in Neuroscience* 11.OCT (2017). issn: 1662453X. doi: 10.3389/fnins.2017.00555 (cit. on p. 1).

[6] P. D. W. Poewe et al. *Nature Reviews Disease Primers* (cit. on p. 1).

[7] A. E. Guttmacher et al. *genomic medicine*. 2003. url: www.nejm.org (cit. on p. 1).

[8] *Epidemiology of Parkinson’s disease*. Aug. 2017. doi: 10.1007/s00702-017-1686-y (cit. on p. 1).

[9] M. J. Armstrong and M. S. Okun. *Diagnosis and Treatment of Parkinson Disease: A Review*. Feb. 2020. doi: 10.1001/jama.2019.22360 (cit. on pp. 2, 6).

[10] R. Balestrino and A. H. Schapira. *Parkinson disease*. Jan. 2020. doi: 10.1111/ene.14108 (cit. on p. 2).

[11] M. Belić et al. “Artificial intelligence for assisting diagnostics and assessment of Parkinson’s disease—A review”. In: *Clinical Neurology and Neurosurgery* 184.May (2019). issn: 18726968. doi: 10.1016/j.clineuro.2019.105442 (cit. on pp. 2, 6, 10, 11, 14, 15).

[12] S. R. Hundza et al. “Accurate and reliable gait cycle detection in parkinson’s disease”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 22.1 (2014), pp. 127–137. issn: 15344320. doi: 10.1109/TNSRE.2013.2282080 (cit. on p. 2).

[13] A. Family, J. D. Gazewood, and D. R. Richards. *Parkinson Disease: An Update*. 2013. url: www.aafp.org/afp (cit. on p. 2).

[14] O. Dehzangi, M. Taherisadr, and R. ChangalVala. “IMU-based gait recognition using convolutional neural networks and multi-sensor fusion”. In: *Sensors (Switzerland)* 17.12 (2017). issn: 14248220. doi: 10.3390/s17122735 (cit. on pp. 2, 7).

[15] H. H. Manap, N. Md Tahir, and A. I. M. Yassin. “Statistical analysis of Parkinson disease gait classification using artificial neural network”. In: *IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2011* (2011), pp. 060–065. doi: 10.1109/ISSPIT.2011.6151536 (cit. on pp. 2, 7, 10, 12, 14, 15, 33, 57).

[16] “Supervised machine learning based gait classification system for early detection and stage classification of Parkinson’s disease”. In: *Applied Soft Computing Journal* 94 (2020), p. 106494. issn: 15684946. doi: 10.1016/j.asoc.2020.106494. url: https://doi.org/10.1016/j.asoc.2020.106494 (cit. on pp. 4, 6, 7, 10, 12, 15, 33, 46–48, 57).

[17] J. Tang et al. “Artificial Neural Network–Based Prediction of Outcome in Parkinson’s Disease Patients Using DaTscan SPECT Imaging Features”. In: *Molecular Imaging and Biology* 21 (6 Dec. 2019), pp. 1165–1173. issn: 18602002. doi: 10.1007/s11307-019-01334-5 (cit. on p. 4).

[18] Y. Moon et al. “Gait variability in people with neurological disorders: A systematic review and meta-analysis”. In: *Human Movement Science* 47 (June 2016), pp. 197–208. issn: 18727646. doi: 10.1016/j.humov.2016.03.010 (cit. on p. 6).

[19] Y. Wu and S. Krishnan. “Computer-aided analysis of gait rhythm fluctuations in amyotrophic lateral sclerosis”. In: *Medical and Biological Engineering and Computing* 47.11 (2009), pp. 1165–1171. issn: 01400118. doi: 10.1007/s11517-009-0527-z (cit. on p. 6).

[20] Q. Ye, Y. Xia, and Z. Yao. “Classification of Gait Patterns in Patients with Neurodegenerative Disease Using Adaptive Neuro-Fuzzy Inference System”. In: *Computational and Mathematical Methods in Medicine* 2018 (2018). issn: 17486718. doi: 10.1155/2018/9831252 (cit. on p. 6).

[21] M. H. Thaut et al. “Velocity modulation and rhythmic synchronization of gait in Huntington’s disease”. In: *Movement Disorders* 14.5 (1999), pp. 808–819. issn: 08853185. doi: 10.1002/1531-8257(199909)14:5<808::AID-MDS1014>3.0.CO;2-J (cit. on p. 6).

[22] J. Massano and K. P. Bhatia. "Clinical approach to Parkinson's disease: Features, diagnosis, and principles of management". In: *Cold Spring Harbor Perspectives in Medicine* 2 (6 2012). issn: 21571422. doi: 10.1101/cshperspect.a008870 (cit. on p. 6).

[23] J. Sringean et al. "How well do Parkinson's disease patients turn in bed? Quantitative analysis of nocturnal hypokinesia using multisite wearable inertial sensors". In: *Parkinsonism and Related Disorders* 23 (2016), pp. 10–16. issn: 18735126. doi: 10.1016/j.parkreldis.2015.11.003 (cit. on p. 6).

[24] N. Hasegawa et al. "How to select balance measures sensitive to parkinson's disease from body-worn inertial sensors—separating the trees from the forest". In: *Sensors (Switzerland)* 19 (15 Aug. 2019). issn: 14248220. doi: 10.3390/s19153320 (cit. on p. 6).

[25] H. Nguyen et al. "Using Inertial Sensors to Automatically Detect and Segment Activities of Daily Living in People with Parkinson's Disease". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26 (1 Jan. 2018), pp. 197–204. issn: 15344320. doi: 10.1109/TNSRE.2017.2745418 (cit. on p. 6).

[26] A. J. Espay et al. "Technology in Parkinson's disease: Challenges and opportunities". In: *Movement Disorders* 31 (9 Sept. 2016), pp. 1272–1282. issn: 15318257. doi: 10.1002/mds.26642 (cit. on p. 6).

[27] A. Marcante et al. "Foot pressure wearable sensors for freezing of gait detection in parkinson's disease". In: *Sensors (Switzerland)* 21 (1 Jan. 2021), pp. 1–12. issn: 14248220. doi: 10.3390/s21010128 (cit. on p. 6).

[28] F. Aydın and Z. Aslan. "Recognizing Parkinson's disease gait patterns by vibes algorithm and Hilbert-Huang transform". In: *Engineering Science and Technology, an International Journal* 24.1 (Feb. 2021), pp. 112–125. issn: 2215-0986. doi: 10.1016/J.JESTCH.2020.12.005 (cit. on pp. 6, 7, 10, 11, 14, 15, 33).

[29] W. Zeng et al. "Parkinson's disease classification using gait analysis via deterministic learning". In: *Neuroscience Letters* 633 (Oct. 2016), pp. 268–278. issn: 18727972. doi: 10.1016/j.neulet.2016.09.043 (cit. on pp. 6, 7, 10, 11, 14, 15).

[30] F. Horst et al. "Gutenberg Gait Database, a ground reaction force database of level overground walking in healthy individuals". In: *Scientific Data* 8.1 (2021), pp. 1–9. issn: 20524463. doi: 10.1038/s41597-021-01014-6 (cit. on p. 7).

[31] S. Aich et al. "A validation study of freezing of gait (FoG) detection and machine-learning-based FoG prediction using estimated gait characteristics with a wearable accelerometer". In: *Sensors (Switzerland)* 18.10 (2018). issn: 14248220. doi: 10.3390/s18103287 (cit. on pp. 7, 10, 12, 15, 16, 33, 47, 57).

[32] T. Varrecchia et al. “An artificial neural network approach to detect presence and severity of Parkinson’s disease via gait parameters”. In: *PLoS ONE* 16.2 February (2021), pp. 1–16. issn: 19326203. doi: 10.1371/journal.pone.0244396. url: http://dx.doi.org/10.1371/journal.pone.0244396 (cit. on pp. 7, 10, 12, 14, 15, 47, 57).

[33] A. Holzinger, ed. *Machine Learning for Health Informatics*. Vol. 9605. Springer International Publishing, 2016. isbn: 978-3-319-50477-3. doi: 10.1007/978-3-319-50478-0. url: http://link.springer.com/10.1007/978-3-319-50478-0 (cit. on p. 7).

[34] V. G. Torvi, A. Bhattacharya, and S. Chakraborty. “Deep Domain Adaptation to Predict Freezing of Gait in Patients with Parkinson’s Disease”. In: *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018* (2019), pp. 1001–1006. doi: 10.1109/ICMLA.2018.00163 (cit. on pp. 7, 10, 11, 14, 15, 33).

[35] L. Borzì et al. “Wearables and Machine Learning”. In: (2021), pp. 1–19 (cit. on pp. 10, 13, 15, 16, 33, 47, 57).

[36] Y. J. Zhao et al. “Progression of Parkinson’s disease as evaluated by Hoehn and Yahr stage transition times”. In: *Movement Disorders* 25 (6 Apr. 2010), pp. 710–716. issn: 15318257. doi: 10.1002/mds.22875 (cit. on p. 10).

[37] A. Tsanas et al. *Statistical Analysis and Mapping of the Unified Parkinson’s Disease Rating Scale to Hoehn and Yahr Staging* (cit. on p. 10).

[38] “Parkinson’s Disease: Hoehn and Yahr Scale”. In: *Current Clinical Neurology* 36 (2012), pp. 4–5. issn: 15590585. doi: 10.1007/978-1-60327-426-5_2 (cit. on p. 10).

[39] “Validation of the Italian version of the Movement Disorder Society - Unified Parkinson’s Disease Rating Scale”. In: *Neurological Sciences* 34 (5 May 2013), pp. 683–687. issn: 15901874. doi: 10.1007/s10072-012-1112-z (cit. on p. 10).

[40] M. Castelli, L. Vanneschi, and S. Silva. “Prediction of the Unified Parkinson’s Disease Rating Scale assessment using a genetic programming system with geometric semantic genetic operators”. In: *Expert Systems with Applications* 41 (10 Aug. 2014), pp. 4608–4616. issn: 09574174. doi: 10.1016/j.eswa.2014.01.018 (cit. on p. 10).

[41] H. R. Gonçalves, A. Rodrigues, and C. P. Santos. “Gait monitoring system for patients with Parkinson’s disease”. In: *Expert Systems with Applications* 185 (Dec. 2021). issn: 09574174. doi: 10.1016/j.eswa.2021.115653 (cit. on pp. 17, 30).

[42] C. Janiesch, P. Zschech, and K. Heinrich. “Machine learning and deep learning”. In: (). doi: 10.1007/s12525-021-00475-2/Published. url: https://doi.org/10.1007/s12525-021-00475-2 (cit. on p. 18).

[43] *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*. isbn: 9781538652572 (cit. on p. 18).

[44] *Continuous home monitoring of Parkinson's disease using inertial sensors: A systematic review*. Feb. 2021. doi: 10.1371/journal.pone.0246528 (cit. on p. 24).

[45] A. Fernández et al. *SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary*. 2018, pp. 863–905 (cit. on p. 27).

[46] S. I. Ao et al. "World Congress on Engineering : WCE 2014 : 2-4 July, 2014, Imperial College London, London, U.K." In: (), p. 1538 (cit. on p. 27).

[47] N. V. Chawla et al. *SMOTE: Synthetic Minority Over-sampling Technique*. 2002, pp. 321–357 (cit. on p. 27).

[48] L. Torgo et al. "SMOTE for regression". In: vol. 8154 LNAI. 2013, pp. 378–389. isbn: 9783642406683. doi: 10.1007/978-3-642-40669-0_33 (cit. on p. 27).

[49] W. Zaremba, I. Sutskever, and O. Vinyals. "Recurrent Neural Network Regularization". In: (Sept. 2014). url: http://arxiv.org/abs/1409.2329 (cit. on p. 31).

[50] "Reservoir computing approaches to recurrent neural network training". In: *Computer Science Review* 3 (3 Aug. 2009), pp. 127–149. issn: 15740137. doi: 10.1016/j.cosrev.2009.03.005 (cit. on p. 31).

[51] R. C. Staudemeyer and E. R. Morris. "Understanding LSTM – a tutorial into Long Short-Term Memory Recurrent Neural Networks". In: (Sept. 2019). url: http://arxiv.org/abs/1909.09586 (cit. on p. 31).

[52] F. A. Gers and J. Schmidhuber. "LSTM recurrent networks learn simple context-free and context-sensitive languages". In: *IEEE Transactions on Neural Networks* 12 (6 Nov. 2001), pp. 1333–1340. issn: 10459227. doi: 10.1109/72.963769 (cit. on p. 31).

[53] K. Smagulova and A. P. James. *A survey on LSTM memristive neural network architectures and applications*. Oct. 2019. doi: 10.1140/epjst/e2019-900046-x (cit. on p. 31).

[54] B. Nakisa et al. "Long short term memory hyperparameter optimization for a neural network based emotion recognition framework". In: *IEEE Access* 6 (Aug. 2018), pp. 49325–49338. issn: 21693536. doi: 10.1109/ACCESS.2018.2868361 (cit. on p. 31).

[55] N. Reimers and I. Gurevych. "Optimal Hyperparameters for Deep LSTM-Networks for Sequence Labeling Tasks". In: (July 2017). url: http://arxiv.org/abs/1707.06799 (cit. on p. 31).

[56] Y. Zhang et al. "Prediction of Freezing of Gait in Patients with Parkinson's Disease by Identifying Impaired Gait Patterns". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28 (3 Mar. 2020), pp. 591–600. issn: 15580210. doi: 10.1109/TNSRE.2020.2969649 (cit. on p. 33).

[57] A. Korotcov et al. “Comparison of Deep Learning with Multiple Machine Learning Methods and Metrics Using Diverse Drug Discovery Data Sets”. In: *Molecular Pharmaceutics* 14 (12 Dec. 2017), pp. 4462–4475. issn: 15438392. doi: 10.1021/acs.molpharmaceut.7b00578 (cit. on p. 33).

[58] “Comparing Multiple Machine Learning Algorithms and Metrics for Estrogen Receptor Binding Prediction”. In: *Molecular Pharmaceutics* 15 (10 Oct. 2018), pp. 4361–4370. issn: 15438392. doi: 10.1021/acs.molpharmaceut.8b00546 (cit. on p. 33).

[59] B. J. Stetter et al. “Estimation of knee joint forces in sport movements using wearable sensors and machine learning”. In: *Sensors (Switzerland)* 19 (17 Sept. 2019). issn: 14248220. doi: 10.3390/s19173690 (cit. on p. 33).

[60] R. Susmaga. *Confusion Matrix Visualization* (cit. on p. 43).

[61] K. R. Chaudhuri et al. *Parkinson’s disease: The non-motor issues*. Dec. 2011. doi: 10.1016/j.parkreldis.2011.02.018 (cit. on p. 57).

[62] P. Martínez-Martín et al. “Parkinson’s disease severity levels and MDS-Unified Parkinson’s Disease Rating Scale”. In: *Parkinsonism and Related Disorders* 21 (1 Jan. 2015), pp. 50–54. issn: 18735126. doi: 10.1016/j.parkreldis.2014.10.026 (cit. on p. 57).