UNIVERSIDADE D
COIMBRA

André Luís Ferreira Teixeira

AUTOMATIC DIAGNOSIS OF GAIT PATHOLOGIES USING GAITREC DATASET

Dissertação no âmbito do Mestrado em Engenharia Eletrotécnica e de Computadores, ramo de Computadores orientada pelos Professores Doutores Manuel Marques Crisóstomo e João Paulo Morais Ferreira, apresentada à Faculdade de Ciências e Tecnologia/Departamento de Engenharia Eletrotécnica e de Computadores.

Setembro de 2022

UNIVERSIDADE D
COIMBRA

Automatic diagnosis of gait pathologies using GaitRec dataset

André Teixeira

Dissertation for obtaining the Master's Degree in

Electrical and Computer Engineering

Advisor:	Manuel Marques Crisóstomo	FCTUC
Co-Advisor:	João Paulo Morais Ferreira	ISEC

Coimbra, September 2022

Resumo

A marcha humana tem sido estudada por anos, devido às várias aplicações possíveis, como prevenção de diferentes patologias, análise de rendimento em atletas, entre outros. Recentemente, aplicações para usar a marcha como identificação de um indivíduo têm sido estudadas.

Das várias formas de caracterizar a marcha, vamos usar as forças de reação do solo. As forças de reação do solo são compostas em três forças, vertical, médio-lateral e anterior-posterior. Foi usada uma grande base de dados, GaitRec, que contém bastante informação descritiva e usa estas forças para caracterizar a marcha.

O objetivo passa por dar uma referência de como a marcha de um indivíduo específico deve ser para se considerar saudável. A identificação do indivíduo vai passar pelas variáveis: género, velocidade da marcha, idade, índice de massa corporal. Sendo dada uma referência para cada lado, esquerdo e direito. Outro objetivo é dar um diagnóstico através de índices de marcha humana.

Para implementar estas referências saudáveis de marcha, algoritmos de aprendizagem computacional foram usados, como *linear regression*, *extreme machine learning*, *ridge extreme machine learning*, *convolutional neural networks* e *multiple output support vector machine*.

Aqui são apresentados dois critérios visando selecionar o melhor algoritmo. Os dois critérios são: capacidade de se comportar igual à verdadeira resposta e o menor valor de distância entre a resposta verdadeira e a resposta criada pelos algoritmos de aprendizagem computacional. Para métrica de distância será usada a *dynamic time warping*.

O algoritmo *extreme machine learning* provou ter o melhor resultado. Dois índices de marcha humana, índice de simetria e erro global, provam que, juntos, conseguem diagnosticar 100% de forma correta.

Para apresentar resultados, duas aplicações foram elaboradas para poder apresentar resultados de forma mais simples e estão disponíveis em https://github.com/adenrteixeira/Gait-Analysis-App.

Palavras-chave: Aprendizagem computacional; Marcha humana; Forças de reação de solo (FRS); GaitRec; Indexes de marcha;

Abstract

The human gait has been analyzed for centuries since it helps to prevent and diagnose different pathologies. Lately was also studied the possibility to be used in identification of a person due to the uniqueness of one's gait.

There are many ways to characterize the human gait, but the late trend and becoming standard are the ground reactions forces. They are composed by three forces, vertical, medio-lateral and anterior-posterior. This what we will use to study the human gait. It's going to be used a dataset, GaitRec, that is a large, fully annotated dataset that uses ground reaction forces to represent gait data.

The objective passes through giving a specific individual their healthy gait reference (we will give two references because we have two limbs, left and right), based on body characterization such as age, gait speed, sex, body mass index. This way we can specifically give a healthy reference instead of a general approach of what it should be looking like. Also, we will be giving a diagnosis through human gait indexes.

To implement this healthy gait references, machine learning algorithms were used, like, convolutional neural networks, extreme machine learning, ridge extreme machine learning, linear regression and multiple output support vector machine.

Here were presented two criteria in order to select the best algorithm. Capability of following the true response and the least distance between the true response and the created one. Dynamic time warping was used as a metric to give the distance value.

The best algorithm proves to be extreme machine learning. Two human gait indexes, symmetric index and global error, prove, that together, are enough to have 100% accurate classification.

To present results, two apps were built and are available at https://github.com/adenrteixeira/Gait-Analysis-App.

Keywords: Machine learning; Human Gait; Ground reaction forces (GRF); GaitRec; Gait Indexes;

Acknowledgments

" As people, we argue all day over whether someone is or isn't a success based off of whether they have or haven't achieved what we value and want. " - Teal Swan

I would like to thank my professors, Dr. Manuel Crisóstomo, Dr. João Ferreira and Dr. Paulo Coimbra, for the guidance and knowledge to complete this project. Thank you for making me develop as a student.

To my family, who played a vital role in this project. Thank you for the assistance. For the love. For your patience. For being present. Thank you for never giving up and understanding my needs and wants. Grateful for your existence.

I'm grateful for finding the work of Teal swan, because it made me get my power back towards my life, and seeing what I really wanted to do.

This experience made me grow, and I would like to thank all that made this experience possible. Thank you universe. Thank you, collective consciousness.

Contents

List of Tables

List of Figures

Acronyms

AP Anterior-posterior.

BMI Body mass index.

CNN Convolutional neural network.

COP Center of pressure.

DTW Dynamic time warping.

ELM Extreme machine learning.

Fl Female left.

Fr Female right.

GDI Gait deviation index.

GE Gait error.

GGI Gillette gait index.

GPS Gait profile score.

GRF Ground reaction forces.

GVS Gait variable score.

LR Linear Regression.

LSTM Long term-short memory recurrent neural network.

MBW Multiple of body weight.

ML Medio-lateral.

Ml Male left.

Mr Male right.

MSVR Multiple output support vector regression.

RELM Ridge extreme machine learning.

RMS Root-mean-square.

SI Symmetric index.

SVM Support vector machine.

SVR Support vector regressor.

V Vertical.

VGRF Vertical ground reaction force.

Chapter 1

Introduction

For years, human gait has been study. It has been a field of interest due to the fact that it can prevent and diagnose pathologies. There is a necessity of creating a targeted healthy gait reference for a specific person, and we will be taking this path (Matsushita *et al.*, 2021a; Horsak *et al.*, 2020).

With machine learning algorithms, like linear regression, extreme machine learning and many more, by using a training dataset, we can train the algorithms to learn the information on that data and then on a test dataset show how it performs. Based on defined criteria and milestones, if it shows a good performance, we have then, a fully capable algorithm, that, by using information that describes that person, creates the targeted healthy gait reference. Machine learning algorithms are only going to be used only for the purpose of creating targeted healthy gait references.

Sobral *et al.*, 2018, takes a similar path and also creates healthy gait references. Yet, the data captured and used are small in quantity. With the recent development of the GaitRec dataset (Horsak *et al.*, 2020), exists now the possibility to use a large amount and verified data by physiatrists.

Also, by applying human gait indexes, we will give a diagnosis. Human gait indexes are a measurement of the difference between the person's gait and a reference for that person (Schutte *et al.*, 2000; Sobral *et al.*, 2018; Schwartz *et al.*, 2008).

1.1 Objectives

The apps created are available at https://github.com/adenrteixeira/Gait-Analysis-App.

- Build, with the help of machine learning algorithms, healthy gait references for a targeted person, using three features, age, body mass index and gait speed.

- Give a diagnosis by applying human gait indexes that classifies the gait as healthy or pathological.
- Create two apps. One specifically to see the healthy gait reference with a studied algorithm and another for the gait analysis with human gait indexes.

1.2 Dissertation outline

The dissertation outline will pass through essentially 5 chapters.

Chapter 2, Background information, describes and presents all the information necessary to be possible to understand the results here presented.

Chapter 3, Materials and methods, particularize the procedures and shows studies about the tools that are going to be used.

Chapter 4, Results and discussion, shows all the results from studies done and described in chapter 3.

Chapter 5, Conclusion. After reading the chapter 4, conclusion resumes it all to submit the final conclusion about the objective.

Chapter 2

Background information

Because we need background information to understand the objectives defined, here is presented, the information necessary to understand them. First, human gait is analyzed. An explanation about what it is, how to measure and what to use to measure. In order to have results, data is necessary. Here we explain details and what the dataset GaitRec is. Machine learning algorithms are going to be used to produce the healthy gait references, so an explanation about the algorithms used is provided. Human gait indexes will be used to give a diagnosis, understanding what they do is shown here. Dynamic time warping concept, and its application, is explained.

2.1 Human gait

2.1.1 What is human gait

Human gait is the manner of walking or moving on foot. Research on gait analysis began in the 19th century, and for many years it was developed a lot of quantitative measurements for the human gait. Gait is affected by many factors like age, sex, weight, height, organ systems, possible injuries, and others (Matsushita *et al.*, 2021a).

The **figure 2.1** shows what a full cycle of gait is and that is divided in many parts. We have a full gait cycle by starting with the striking of either foot with the ground and ends with the successive strike of the same foot (Matsushita *et al.*, 2021a). These parts are simple to understand and intuitive if we look at the **figure 2.1**.

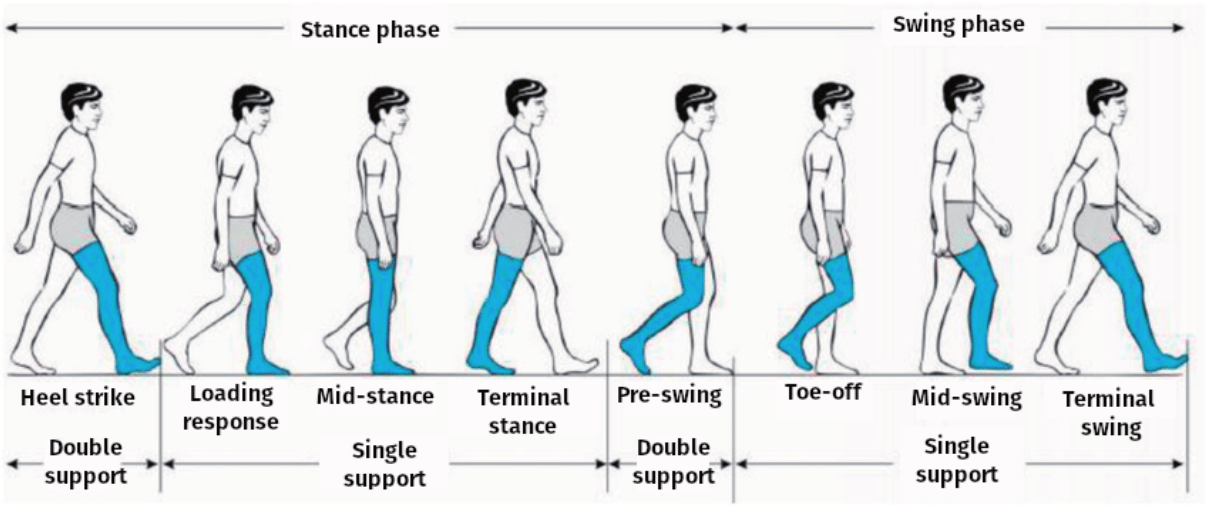


Figure 2.1: Full Gait cycle of a human. Adaptation from: (Nordin *et al.*, 2018).

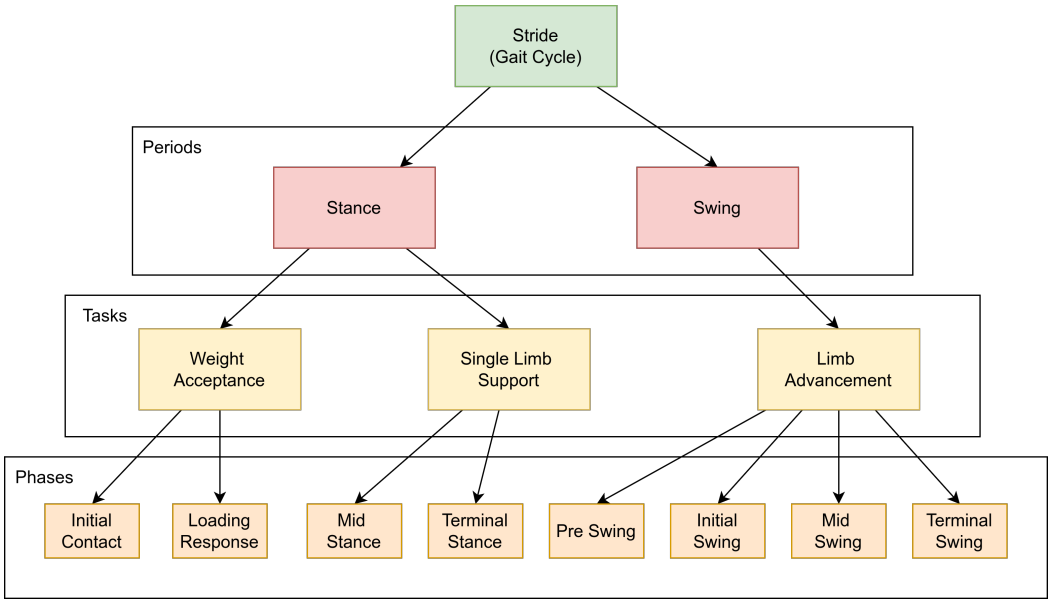


Figure 2.2: Division of the gait cycle. An advanced classification of human gait cycle. Divides it in periods, tasks, and phases. Adaptation from: (Kharb *et al.*, 2011).

A very similar classification and division of the gait is used in (Kharb *et al.*, 2011), displayed in **figure 2.2**, yet, it furthers divides it and presents some different names for each part of the gait cycle.

Using machine learning to understand patterns on human gait can help clinicians to get a better understanding of it and give quicker results. Also, can give future insights (Alaqtash *et al.*, 2011) and preventive diagnoses (Matsushita *et al.*, 2021a).

One of the many applications of the studies about gait is sports. It can be used to prevent and detect fatigue, faulty performance and detect early imbalances that might cause an injury. Another

application can be in security. A person's gait is always different and a genuine characterization of the unit, therefore, can serve for identification (Matsushita *et al.*, 2021a).

2.1.2 Human gait measurements and characterization

There are many variables to quantify the human gait. We are follow the taxonomy of the (Matsushita *et al.*, 2021a) to have an organization and understanding. Spatio-temporal parameters are related to time and distance in a gait cycle, examples are: steps, stride length, walking speed, events, like foot strike and toe-off, etc. Then we have kinematic parameters. They measure the entire body mass, angular rotation of the joints (ankle, hip, many more), etc. The physiological parameters are indicators of energy expenditure, like, energy cost, heart rate, energy consumption, etc. Then we have the anthropometric parameters, which is information directly taken from the patient profile like, age, gender, weight, height, body mass index, etc. The last one is kinetic parameters which refers to ground reaction forces (GRF), foot plantar center of pressure (COP), joint torque, etc.

The quantitative measurements that will be used to describe the gait are the ground reaction forces (due to the fact that we will be using GaitRec dataset, and GaitRec uses this variable to quantify the gait). Ground reaction forces are the force of equal magnitude applied in the opposite direction from the ground up to the foot. There are 3 GRF, vertical (V), anterior-posterior (AP) and medio-lateral (ML). **Figure 2.3** has a representation of the direction and orientation of the 3 GRF. This kinetic parameters, GRF, are becoming a standard tool for clinicians to measure, describe and analyze human locomotion (Horsak *et al.*, 2020).

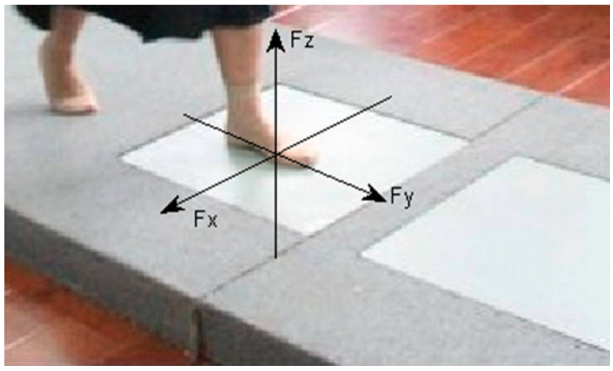


Figure 2.3: Representation of the ground reaction forces. Fz is the vertical GRF; Fx is the medio-lateral GRF; Fy is the anterior-posterior GRF. Source: (Liang *et al.*, 2015).

2.1.3 Sensores

In pursuance of getting information that can characterize the human gait, it is imperative to capture information from this reality into digital information. Sensors, by definition, are capable of it. The following **figure 2.4**, gives a possible taxonomy for sensors that are used for gait study.

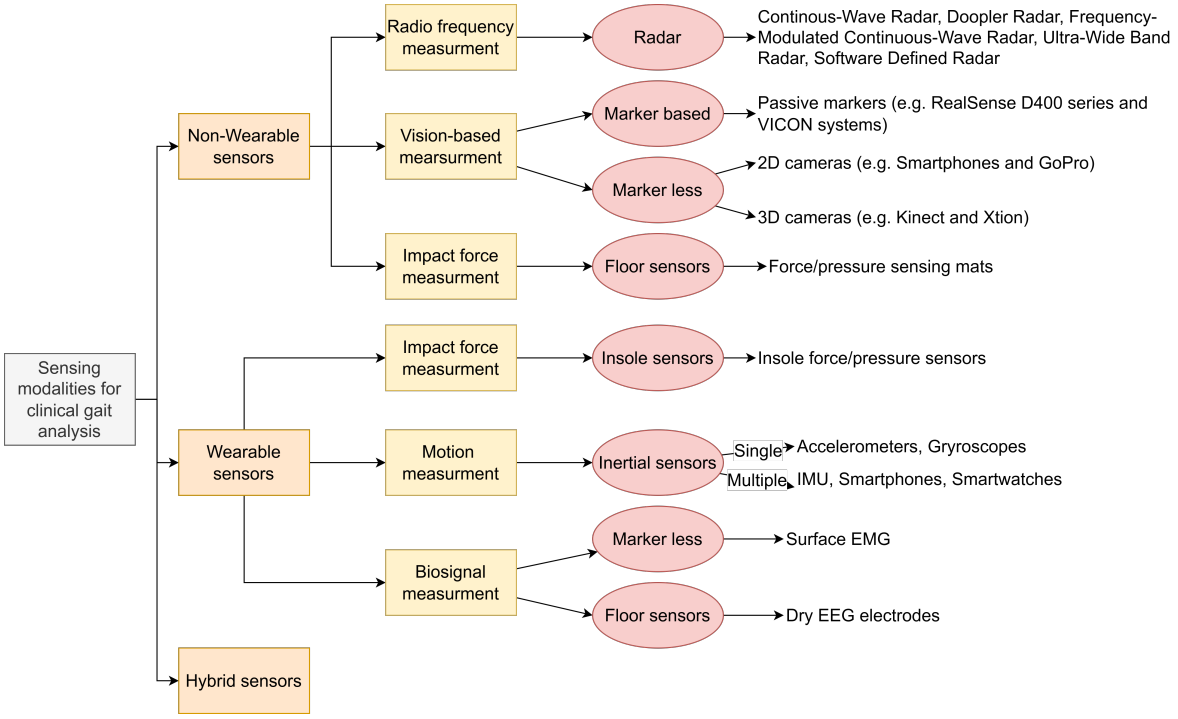


Figure 2.4: Gait acquisition modalities for clinical gait analysis. A possible taxonomy to describe gait acquisition modalities for clinical gait analysis. Adaptation from: (Matsushita *et al.*, 2021a).

The nonwearable sensors are very typical in controlled facilities, but some can be used installed at unprofessional facilities. One example is the radar which is non-invasive, non-contact, and insensitive to lights, and can, for example, help to inform if a fall has happened in an elderly house. This kind of sensors can be considered as powerful since there are no power constrains and low physical limitations, but can be very expensive to install.

The wearable sensors on the other hand are preferred for gait analysis. They are simpler and can be used in daily life. This can be considered the opposite of the nonwearable sensors in terms of advantages and disadvantages. One does not have portability, the other has. One has no power constrains, the other has. Likewise, one can be inexpensive, but nonwearable are quite expensive.

The hybrid category is the use of various sensors to get more information about gait, possible taking both kinetic and kinematic information. It doesn't necessary means that are a mix of nonwearable sensors and wearable (Matsushita *et al.*, 2021a).

Because we are using the GaitRec dataset, it is important to explain the type of sensor they adopt. GaitRec uses nonwearable sensors in order to measure the GRF, to measure the impact force. Usually, they are called floor/pressures platforms sensors. They measure the static and dynamic pressure or force under the foot and are placed along the floor (Matsushita *et al.*, 2021a; Horsak *et al.*, 2020).

2.2 GaitRec

For the sake of building results and studying, data is necessary. This is where GaitRec comes in. GaitRec is a large scale dataset built by a research group that the main objective is to fulfill the need of large datasets for state-of-the-art machine learning algorithms. Another necessity that was denoted is the fact that the availability of a completely annotated large-scale datasets is very scarce. This is where the GaitRec excels, it is a large scale dataset, completely annotated and classified by certified physiatrists, ensuring the quality of the results (Horsak *et al.*, 2020).

GaitRec has 211 healthy units and, 2084 non-healthy units. It has 104 healthy males and 107 healthy females; 1636 unhealthy males and 448 unhealthy females. In terms of balanced dataset, the healthy sub-dataset is quite balanced, but for the pathologic that doesn't happen. There are much more pathological males than females. For our case of studying this is good since we only need the healthy dataset to implement the algorithms. This is because we want to give a healthy gait reference for a specific unit (non-healthy or healthy) based on their personal information, so the algorithms can only learn the healthy cases. The unhealthy units will be used to take results, and the unbalanced fact is not relevant to the veracity of the results.

Healthy units had the possibility to do the sessions barefoot and with their regular shoes. But the pathological units only did with their shoes, and if they were using orthopedic insoles it was referenced. Each session, each measurement, had at least 10 trials.

The only true classification given is for the first session of that person. This first session is classified as *initial measurement*. Then, in the case of the unhealthy persons, we also have *control measurement* and *initial measurement after readmission*. One relates to the sessions that the person did while on rehabilitation to resolve the pathology, and the other to the case if the person comes back for any reason, respectively. We only have a physiatrist classification on the *first measurement*. So trying to give a classification on the last session of the person, assuming it is healed, could fall on potential error. The person could have given up on treatment, actually healed, or not at all and needs to come back to rehabilitation again and be on the session type *initial measurement after readmission*. This is important because giving a diagnosis is possible,

but having the opportunity to contrast between our classification and the true classification on the last sessions will not be possible. We can only assume that evolution towards a healthy state can be seen in most cases.

Healthy units did the sessions with at maximum 3 gait speeds. 1-slow (mean 0.98 m/s, 0.14 for standard deviation); 2-self-selected (mean 1.27 m/s, 0.13 for standard deviation); 3-fast (mean 1.55 m/s, 0.15 for standard deviation); Unhealthy units, understandably, only at speed 2.

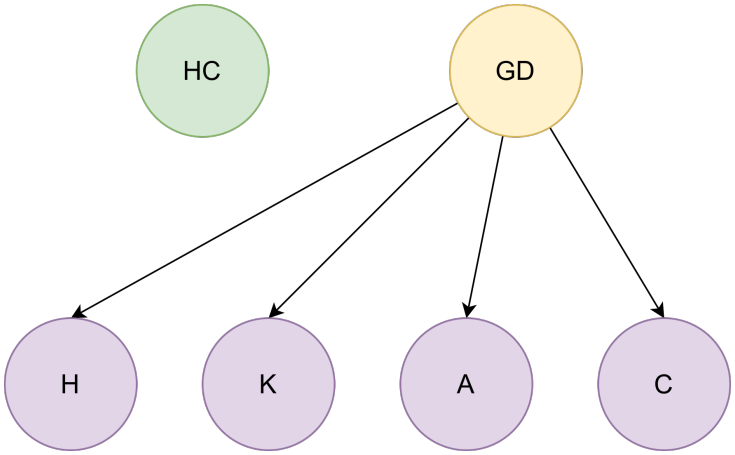


Figure 2.5: GaitRec taxonomy. Taxonomy used for classifying the patients. Adaptation from: (Horsak *et al.*, 2020).

The last thing that needs to be examined is the classification taxonomy the authors of the GaitRec use. And by analyzing the **figure 2.5** it is easy to understand it. Patients can be classified as: Healthy Controls (HC) or Gait Disorders (GD). For the patients that have gait disorders, they can be further classified as, Hip (H), Knee (K), Ankle (A), and Calcaneus (C). So, for the unhealthy patients, we have 4 major problems and their names comes from the location it happened, hip, ankle, knee, and calcaneus. Further classification is given within the 4 major problems (Horsak *et al.*, 2020), but they are irrelevant to our objective.

2.3 Machine learning

2.3.1 Linear regression

Linear regression is a supervised learning algorithm that tries to find a representative curve, or in higher dimensions, a hyperplane, to a set of input/output data. This will give the opportunity to build a model who has the capacity, by the information summarized, to give an output that follows the same information. To find the parameters of the hyperplane that best represents the

dataset used, we need a cost function to measure how well the model is representing. Least squares is the one used in this work and is one of the many popular choices (Watt *et al.*, 2016).

2.3.2 Extreme machine learning

Extreme machine learning (ELM) is a single-hidden layer feed forward network, and its weights and bias are randomly assigned. The creators of ELM prove that if the activation functions are infinitely differentiable, the weights and bias can be randomly assigned. The output weights then are determined by a simple inverse operation of the hidden layer output matrix (when calculating computationally, pseudo-inverse is used) (Huang *et al.*, 2006; Cao *et al.*, 2016).

The algorithm itself, it's simple to understand. The only hyperparameter that needs to be defined is the number of the neurons on the hidden layer ($\tilde{N}$).

Defining a set of training samples $\{(\mathbf{x}_j, \mathbf{t}_j)\}_{j=1}^{N}$ with the size of the training vectors (N), (m) for the number of outputs and (n) the number of features, the single-hidden layer feed foward network is expressed as:

$$\sum_{i=1}^{\tilde{N}} \beta_i g_i(\mathbf{x}_j) = \sum_{i=1}^{\tilde{N}} \beta_i g_i(\mathbf{w}_i \mathbf{x}_j + b_i) = \mathbf{o}_j, \quad j = 1, ..., N \tag{1}$$

Where $g(\cdot)$, in **equation** (1), is the activation function, $\mathbf{x}_j = [x_{j1}, x_{j2}, ..., x_{jn}]^T$ the input, $\mathbf{t}_j = [t_{j1}, t_{j2}, ..., t_{jm}]^T$ the desired output, $\mathbf{w}_i = [w_{i1}, w_{i2}, ..., w_{in}]^T$ the connecting weights of the *ith* hidden neuron to input neurons and b_i the bias of the *ith* hidden node, $\beta_i = [\beta_{i1}, \beta_{i2}, ..., \beta_{im}]^T$ is the connecting weights of the *ith* hidden neuron to the output neurons and $\mathbf{o}_j$ is the network output with respect to input $\mathbf{x}_j$, where $i = 1, ..., \tilde{N}$. Then, the weights ($\mathbf{w}_i$) and bias (b_i) are randomly generated. The objective is, minimize the following error $\min_{\beta} ||\mathbf{H}\beta - \mathbf{T}||$, the difference between $\mathbf{o}_j$ and $\mathbf{t}_j$ where **H**, called the hidden layer output matrix, defined in **equation** (2).

$$\mathbf{H}(w_1, ...w_{\tilde{N}}, b_1, ...b_{\tilde{N}}, x_1, ...x_{\tilde{N}}) = \begin{bmatrix} g(w_1 * x_1 + b_1) & \cdots & g(w_{\tilde{N}} * x_1 + b_{\tilde{N}}) \\ \vdots & \ddots & \vdots \\ g(w_1 * x_N + b_1) & \cdots & g(w_{\tilde{N}} * x_N + b_{\tilde{N}}) \end{bmatrix} \tag{2}$$

β is calculated with $\beta = \mathbf{H}^{\dagger}\mathbf{T}$ († is the pseudo-inverse operation) since, $\min_{\beta} ||\mathbf{H}\beta - \mathbf{T}||$, is a least square problem (Huang *et al.*, 2006; Cao *et al.*, 2016).

2.3.3 Ridge extreme machine learning

This algorithm is equal to ELM, but the fact that the numerical instability of the pseudo-inverse exists, a version of the ridge regression is used to optimize $\underset{\beta}{min}||\mathbf{H}\beta - \mathbf{T}||$. To do this another hyperparameter, λ, is inserted on the equation $\underset{\beta}{min}||\mathbf{H}\beta - \mathbf{T}|| + \frac{1}{\lambda}||\beta||$ (Cao *et al.*, 2016).

2.3.4 Support vector machine

Support vector machine, SVM, uses kernels functions in order to separate the classes within a margin and give a classification for non-linear cases (when it's not possible to separate the data with a simple line). The problem in hands (create the targeted healthy gait reference) have 101 outputs (full gait cycle by GaitRec), so, SVM for regression problems (SVR), with multiple output variables it's going to be called multiple output support vector regression (MSVR) and instead of giving a categorical output like SVM, gives a numerical output (Marsland, 2014; Sanchez-Fernandez *et al.*, 2004; Watt *et al.*, 2016). The following explanation is about SVR and not MSVR, but it can be deducted to the MSVR case.

If we consider the training set as $\{(\mathbf{x}_i, t_i)\}_{i=1}^{N}$, where $j = i = 1, ..., N$, $\mathbf{x}_i \in \mathbb{R}^n$, $t_i \in \mathbb{R}^1$, SVR solves the following problem (with, N the size of the training vectors, n the number of features, $\mathbf{x}_i$ being the input vector and t_i being the desired output):

$$\underset{w,b,\zeta,\zeta^*}{min} \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{N}(\zeta_i + \zeta_i^*) \tag{3}$$

Equation (3) is subject to these conditions:

$$t_i - \mathbf{w}^T\phi(\mathbf{x}_i) - b \leq \varepsilon + \zeta_i$$
$$\mathbf{w}^T\phi(\mathbf{x}_i) + b - t_i \leq \varepsilon + \zeta^*$$
$$\zeta_i, \zeta_i^* \geq 0$$

With these settings, we are penalizing samples whose prediction is at least ε away from their true target. Depending on whether their predictions lie above or below the ε tube, they are penalized by ζ_i or ζ_i^*. Where C is a regularization parameter.

$$\underset{\boldsymbol{\alpha},\boldsymbol{\alpha}^*}{min} \frac{1}{2}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)^T Q(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*) + \varepsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{N} t_i(\alpha_i - \alpha_i^*) \tag{4}$$

Equation (4) is subject to these conditions:

$$e^T(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*) = 0$$
$$0 \leq \alpha_i, \alpha_i^* \leq C$$

Where $Q_ij \equiv K(x_i, x_j) = \phi(x_i)^T\phi(x_j)$. Q is an N by N matrix and because of equivalance is equal to K which is the kernel. $\phi(\cdot)$ is a function that maps the training vectors into a higher dimensional space. α_i, α_i^* are called the dual coefficients.

Then we can predict with **equation** (5) (Pedregosa *et al.*, 2011; C.-C. Chang *et al.*, 2011):

$$\sum_{i=1}^{N}(-\alpha_i + \alpha_i^*)K(\mathbf{x}_i, \mathbf{x}) + b \tag{5}$$

2.3.5 Convolutional neural network

Convolutional neural networks (CNN) are one of the most famous neural networks algorithms that exist. It is known for its good results for a wide range of problems. It is possible to work with them in many dimensions, for example, images can be considered an 2-D grid, and time-series as 1-D.

In order to be considered as a convolutional neural network, it needs to have at least one convolutional layer. Convolutional, pooling (max pooling for example) and normalization operations are very used and typical. In the interest of constructing an architecture, a model, we have to use many layers with these operations to have a model that is capable of giving a good result. Activation functions are also super important because they are functions who change the values of the features maps (sigmoid, ReLu, tahn are a few of many).

Why use convolutional layers? Because they help learning from sparse interactions, parameter sharing and equivalent representations. For the convolutional layers, the following taxonomy needs to be taken in account. Kernel, filters, input, and feature map. Kernel is a matrix of weights that is used for the convolutional operation with the input. Filters represent the channels of the convolutional layer, the depth of it. And the results of the operations is called the feature maps (Goodfellow *et al.*, 2016).

A general convolutional neural network has also dense layers, which is a layer that is deeply connected with its previous layer, so, the neurons of the layer are connected to every neuron of its preceding layer. **Figure 2.6** has an example of what a CNN architecture can look like. We see convolution layers, pooling layers, and at the end, deeply connected layers.

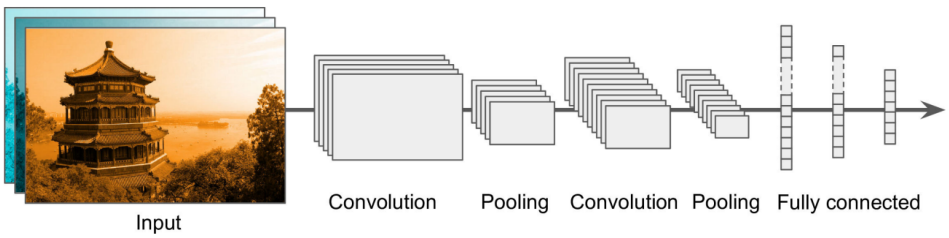


Figure 2.6: CNN model. A typical CNN architecture. Source: (Géron, 2019).

2.4 Human gait indexes

There is some investigation around indexes that can provide a diagnosis based on parameters that are directly connected to the person's Gait.

2.4.1 Gillette gait index (GGI)

Gillette Gait Index (GGI) is widespread in terms of usage and popularity. Although with the advantage of resuming all the gait analysis in to one index, it has some major disadvantages, lack of normal score distribution, and arbitrary and incomplete nature of the 16 constituent variables (although not restricted to 16, on their investigation they used 16), not easy to interpret, etc (McMulkin *et al.*, 2015).

GGI is calculated by setting the values of the variables used and the mean of them. Then, in favor of getting more information we have to uncorrelate these two results, the variables values and the mean value of all the variables. To do so, they apply numerous technics from multivariate statics to accomplish this. Then, the Euclidean distance between the variables and mean, now uncorrelated, is calculated and summed to give a final index value (Schutte *et al.*, 2000).

2.4.2 Gait deviation index (GDI)

Gait Deviation Index (GDI) solved many problems that the GGI had. It includes data from the entire gait cycle, it's easy to interpret, has an inherent filter and is normally distributed, allowing parametric statistical testing. But although building a better response than GGI it comes with a cost that can be considered as a major disadvantage. It has the need for a large sample size, and this limits the application to other studies (McMulkin *et al.*, 2015).

GDI is a little more complex than GGI. First, it passes through preprocessing the data and then subjecting it to principal component analysis. After, select the data that can represent the information about the used data and represent it linearly to create a reduced order approximation. Then, a distance metric is defined to measure the proximity between the selected data and the reduced order approximation (Schwartz *et al.*, 2008).

2.4.3 Gait profile score (GPS)

Another alternative to GGI and GDI is gait profile Score. It's based on the GDI, but aimed for easier calculations and with the objective of resolving some problems the other two had. One major advantage is that it can adapt to new models without the need for large control database

like GDI needed. One disadvantage the GPS has is the lack of normal distribution (McMulkin *et al.*, 2015).

GPS uses all kinematic data, instead of the approach of the GDI that only uses the kinematics that represents the data. Using the root-mean-square (RMS) between the mean and the all kinematic itself, proves that, if using all the kinematics, no principal component analysis is necessary. This way, GPS avoids intense calculations (Baker *et al.*, 2009).

One thing that is important to talk about is the gait variable score (GVS). GVS is calculated just like GPS, but is called GVS when we are analyzing one gait variable (Baker *et al.*, 2009).

2.4.4 Symmetric index (SI)

The symmetric index analyzes the behavior of both limbs, left and right, and identify if they both behave identically. The symmetric index is built by analyzing the same person's gait, for left (on the y-axis) and right side (on the x-axis), and trying to find a curve that fits them. With the curve defined, we can find ψ, that is the angle between the x-axis and the curve.

There are many articles that assumed that gait was perfectly symmetrical with a view to simplification. Although, sometimes, pure symmetry is found (like analyzing the foot-floor contact), gait asymmetry is present and considered normal on healthy units. This difference relates to the contribution of each limb to propulsion and control tasks by preferring one side to do one task, and the other limb to do the other task. Other reasons might exist, but need more studies to actually be plausible to be considered (Sadeghi *et al.*, 2000).

So, to have perfect symmetric index, we would need that ψ was equal to 45°. To calculate the symmetric index, it will be followed the following **equation** (6) (Sobral *et al.*, 2018).

$$SI = \frac{\psi - 45^{\circ}}{45^{\circ}} * 100 \tag{6}$$

2.4.5 Gait error (GE)

Gait error was also analyzed. It takes sixteen variables based on the gait parameter, vertical ground reaction force. Only the vertical GRF is going to be applied because it's the one used on (Sobral *et al.*, 2018).

Gait error, **equation** (7), is calculated using the difference between the variables from the patient's and the respective reference variables (Sobral *et al.*, 2018).

$$GE = \sqrt{\frac{1}{U}\sum_{i=U}^{U}(\frac{x_i - x_{iref}}{x_{iref}})^2} \tag{7}$$

Where x_i represents each variables calculated using the patient's profile and x_{iref} represents each variables using the respective reference profile. U is the number of parameters used.

The **table 2.1** has all sixteen variables that were used by (Sobral *et al.*, 2018).

Table 2.1: List of the variables analyzed to construct the GE and its description

Variable	Brief Description
F1	First peak force
F2	Minimum force
F3	Second peak force
T1	Time of the first peak
T2	Time of the Minimum force
T3	Time of the second peak force
T_{st}	Stance phase duration
T_{isng}	Beginning of the single stance phase
T_{esng}	End of the single stance phase
T_{sgst}	Single stance phase duration
LR	Load rate
ULR	Unload rate
A_{spec}	Area of the spectral density
A_{deriv}	Area of the VGRF derivative vs the VGRF
GVS	Gait Variable Score
DTW	Dynamic Time Warping

These variables are a compilation that tries to give as much information as possible in favor of creating an index that is capable of give a diagnosis (Sobral *et al.*, 2018).

GVS and DTW are already a comparison between the true response and reference. So when producing the reference value for the GE, a comparison between the created reference and the literature reference is employed ("Appendix A: Kinematic, Kinetic, and Energy Data", 2009).

2.5 Dynamic time warping (DTW)

To compare two time series, sometimes it is advantageous to understand their trajectory. Gait can be very influenced by changes on gait speed (Matsushita *et al.*, 2021a) or even the randomness variable on how a person may walk on a measurement. This can introduce some error that is not derivative from possible pathologies, but from the randomness and typical

variations of the way a person walks, specially time related. A person might have the same gait between measurements but changes the gait speed or the time it starts, or gets into a phase of the gait cycle a little earlier or later. Because of this, using time warping, comparison of trajectories, or comparison of sequences, is necessary and is done by time-sampling. The objective is, by selecting some points and comparing the two sequences, to find the minimum distance between them. This way we are analyzing the pattern of the sequences.

The analysis of the gait patterns is vital for the objective, time warping is a very special tool to achieve that (Kruskall *et al.*, 2022) taking out the time dependency of the Gait. A representation of what DTW does is presented in **figure 2.7**. DTW is applied in this work as a metric to give a value of distance between two gait curves.

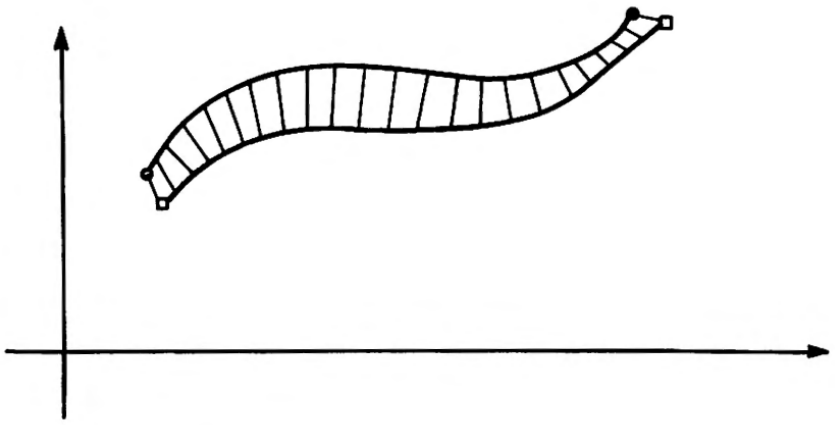


Figure 2.7: DTW representation. Representation of two continuous sequences and the minimum distance found for some points of the sequence. Source: (Kruskall *et al.*, 2022).

Chapter 3

Materials and methods

In this chapter, a description of what was defined to get results, the methods, is presented. The tools used are here discriminated. Besides that, a deep study of the subdataset from the GaitRec used to train the machine learning algorithms, is also shown.

3.1 GaitRec

The GaitRec dataset (Horsak *et al.*, 2020), is shared in 2 different formats, *RAW* and *PRO*. For this work, it was used the *PRO* one. This decision was based on the fact that it has much more processing, which is necessary in order to avoid noise that could destroy the learning of the algorithms implemented. The *RAW* dataset was also pre-processed, although not as much as *PRO*. **Figure 3.1** describes the processing made in both formats, *RAW* and *PRO* suffered.

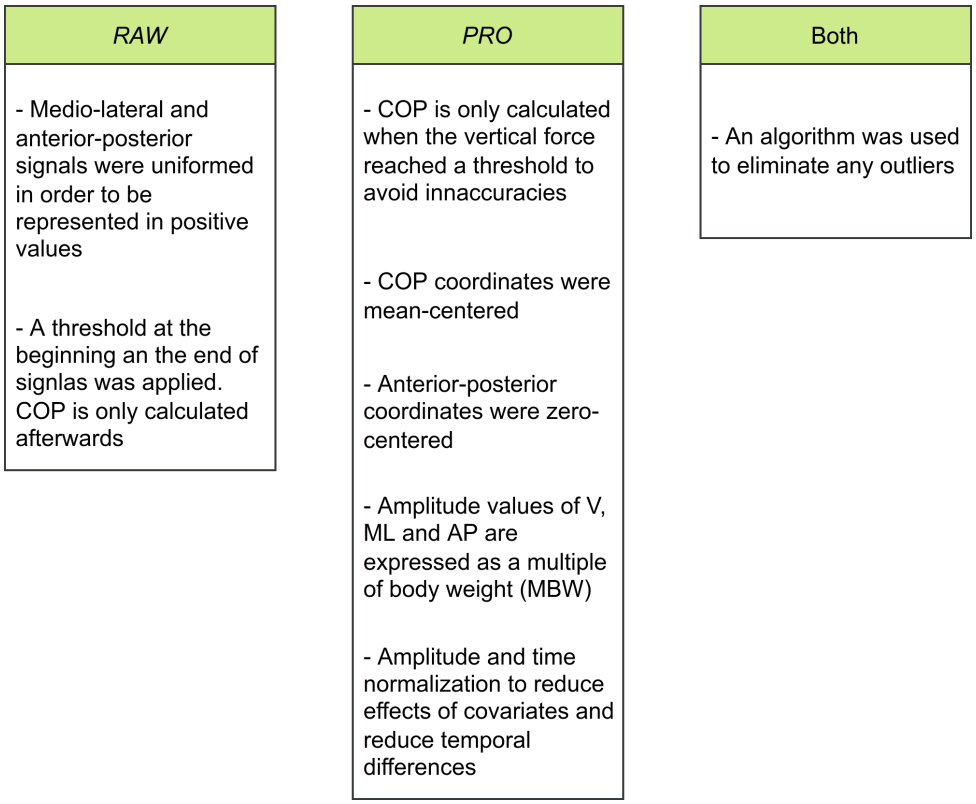


Figure 3.1: GaitRec processing. List of processing done that were done to the GaitRec dataset in *RAW* and *PRO* formart

3.1.1 A deep study of healthy units in GaitRec dataset

In the direction of building the response to the first objective here defined (creating a targeted healthy gait reference), we need to build a subset of the dataset GaitRec. Since the objective is to build algorithms that are capable of giving a healthy gait reference, we need to choose only healthy units.

GaitRec, although severely well analyzed and with correct medical diagnoses, has some data missing. For healthy units, it has every information about the patient. Though, sometimes people didn't always walk all the three gait speeds or didn't walk barefoot.

The structure in GaitRec dataset for healthy units should be exactly like the structure described in the **figure 3.2**. Because we have three gait speeds, must exist 3 sessions for each healthy unit. For each session, it was done 10 trials. Also, we have a repetition of these 3 sessions, but this time, people walked barefoot, instead of the previous one, where they used their regular shoes. So in total, it is expected 6 sessions for all healthy units. But there is a range of values that go from 1 to 6 and not always 6. Sometimes, and this is something that was expressed in their article (Horsak *et al.*, 2020), they didn't save all the 10 trials.

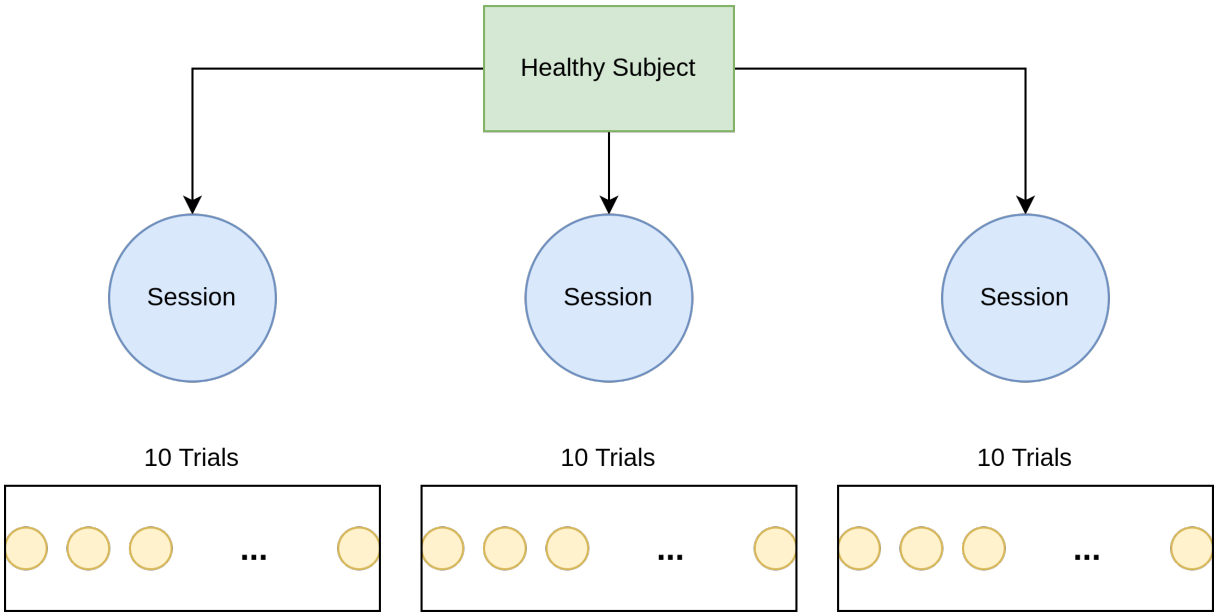
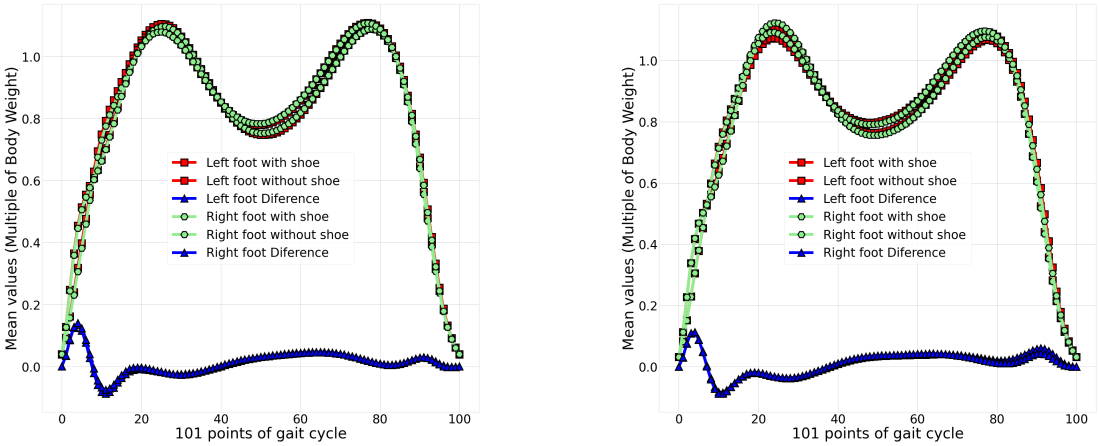


Figure 3.2: Healthy GaitRec structure diagram. Diagram that represents how the data structure should be looking, with sessions and trials, in the healthy units of GaitRec.

3.1.2 Walking shoes

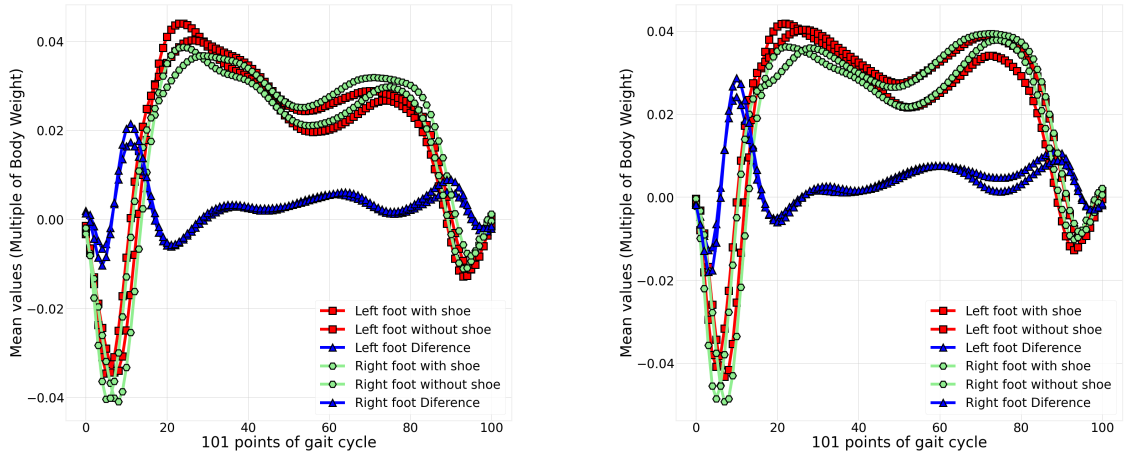
The sessions realized by healthy units were done with regular shoes and some barefoot. In order to understand how much different the response would be we built the mean of all the trials of walking barefoot, and the mean of all the trials of walking with regular shoes. This was done to all the 3 forces, for male and female. Then we calculated the difference between the both means and present it on graphics in the **figures 3.3, 3.4 and 3.5**.



(a) Female case.

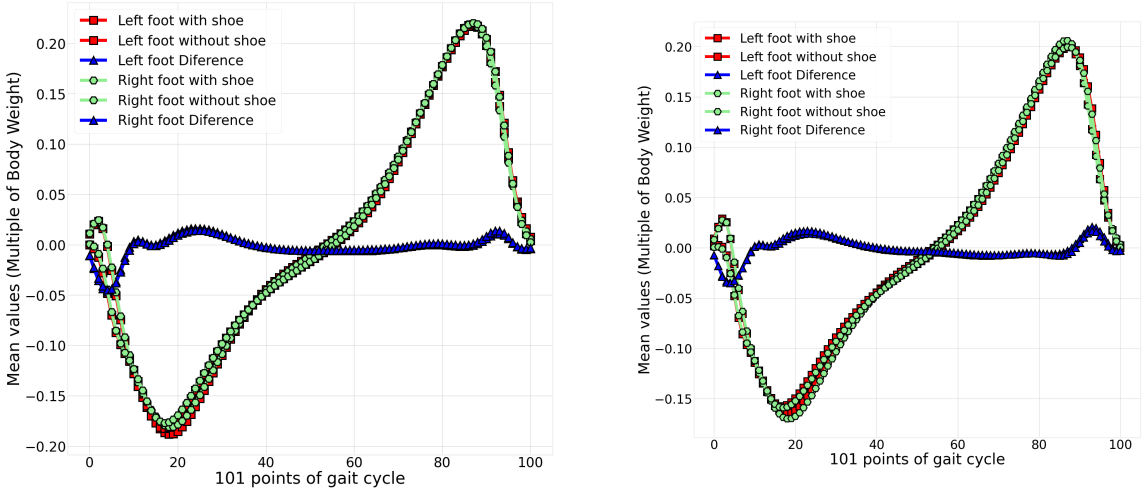
(b) Male case.

Figure 3.3: Mean value of the difference between vertical GRF with shoes and without, for left and right foot.



(a) Female case. **(b)** Male case.

Figure 3.4: Mean value of the difference between medio-lateral GRF with shoes and without, for left and right foot.



(a) Female case. **(b)** Male case.

Figure 3.5: Mean value of the difference between anterior-posterior GRF with shoes and without, for left and right foot.

In the figures 3.3, 3.4 and 3.5, the line that has the blue triangles represents the difference between the means. We have two lines for the difference, left and right limbs, but distinguishing these two blue lines is unnecessary in this case because of their similarity. The line that has red squares represent the left side (with and without shoes) and the green hexagons represent the right side (with and without shoes).

Analyzing the vertical GRF, both for man and female, the maximum value of the difference doesn't get close to the value 0.2 (MBW). Seeing the maximum value of the mean of walking

with regular shoes and walking barefoot, around 1.2 (MBW), 0.3 (MBW) can represent a close value to the quarter of the mean values. This way, because it doesn't get close to the 0.3 (MBW), staying around 0.2 (MBW), the value of the difference it's not high.

Now, analyzing the medio-lateral GRF, the maximum value of the difference is around 0.02 (MBW), and in the male case we have even higher values. Seeing that 0.02 (MBW) represents a close value of the half of the maximum value of the means (around 0.04 (MBW)), we can say that we have an error close to 50 %.

At last, the anterior-posterior GRF, it's really similar to the vertical GRF, and the values don't even get close to the quarter of the maximum value of the means, each in this case is around 0.2 (MBW). So the difference here is also not high.

In general, for both left and right side, exists its differences but an insignificant and expected difference. Asymmetry is the natural human response, even for healthy people, but symmetry is and was assumed in many studies because of simplification (Sadeghi *et al.*, 2000).

So we can conclude, by analyzing the graphics, that between male and female, we have much difference on medio-lateral GRF. Between forces, although for the vertical and anterior-posterior the differences don't show to be high enough, on medio-lateral, ther is a huge difference. And because there is a propagation of difference on all the difference curves, and by this we mean that the difference mean values don't are close to zero for long periods of time of the gait cycle, we conclude that there is difference on walking barefoot and with regular shoes.

3.1.3 Person distributions

So, although the conclusion that walking barefoot and with regular shoes has its differences, we still included all trials and sessions, because this data still is correct. We are considering it as a different person, that has the same age and body mass index for each different trial and session.

To fully analyze the created dataset of only healthy units, it is present here the distribution of the 3 features that we choose and influence the gait. Age, body mass index and gait speed. Sex and foot also influence, but, in the methods' section will be better explained the strategy that was deployed to deal with these features.

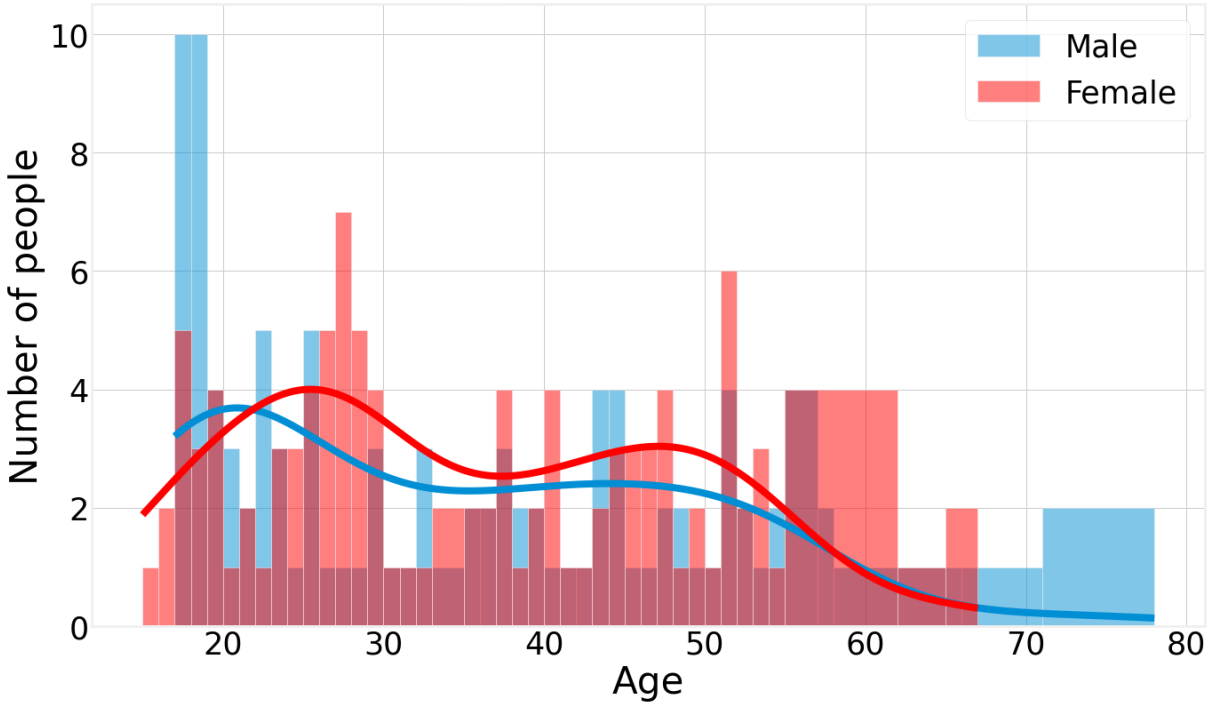


Figure 3.6: Histogram of the age of healthy people.

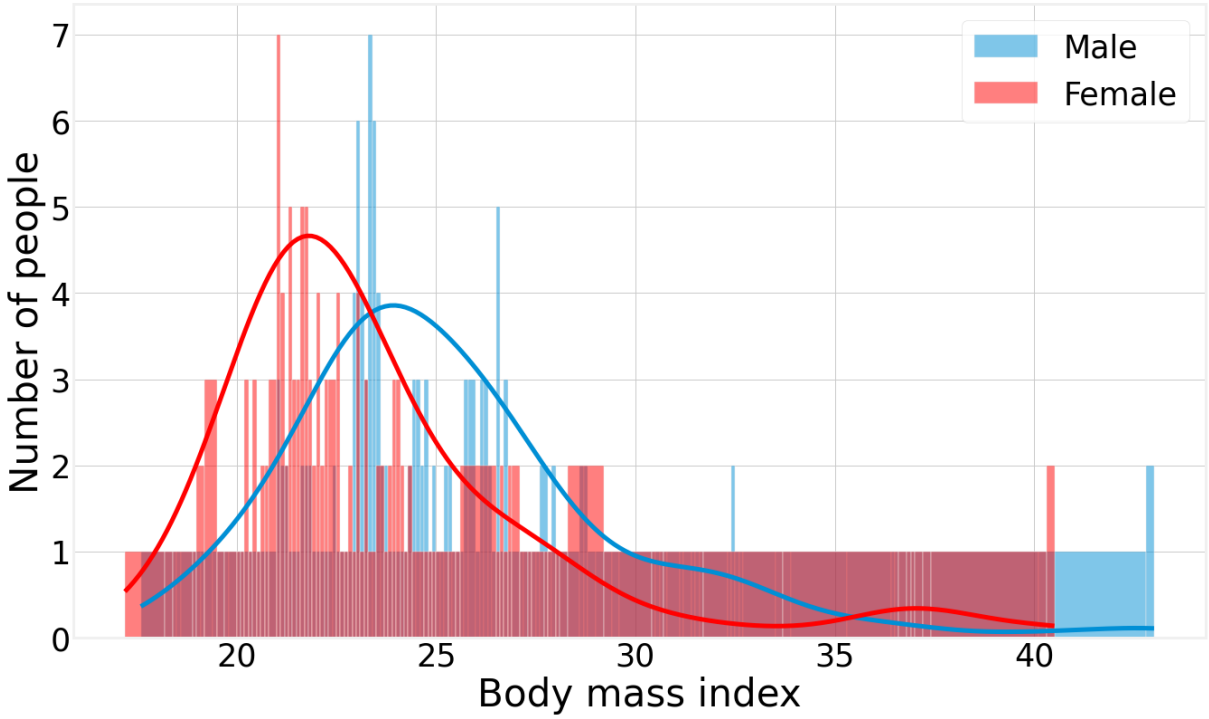


Figure 3.7: Histogram of the body mass index (BMI) of healthy people.

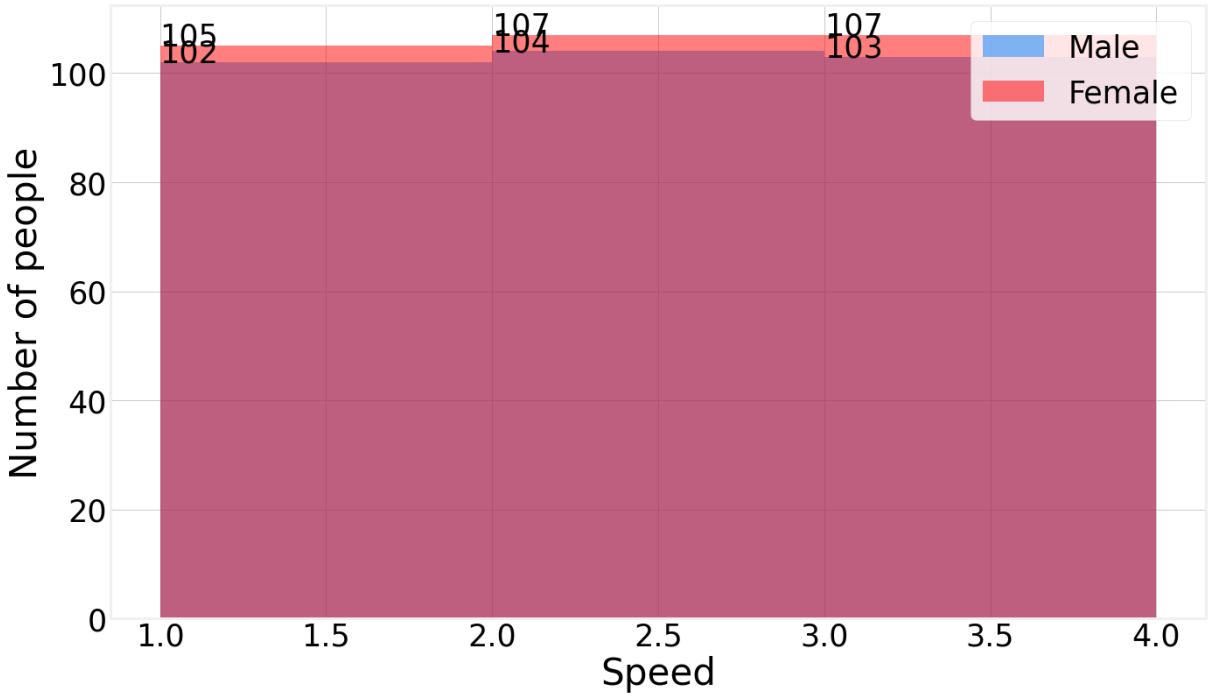


Figure 3.8: Histogram of the gait speed of healthy people.

Analyzing the **figure 3.6**, the age histogram of healthy people, we can see that we have a wide range of values. Between males and females, the distribution is not that different. But we can see that males have a wider range of values. For females, we have values that go from 15-97 and for males 17-78. We can see higher number of persons in the interval 20-30 years old and around 40-50 years old. But in general, the data is well distributed between all age values, both for male and female.

In the body mass index histogram graph, **figure 3.7**, we can see that we have a huge concentration in values around 20-25 for females and 23-27 for males. We can consider that this kind of distribution is well similar to a Gaussian one and this behavior is not new and already expected (Nuttall, 2015).

When we analyze the gait speed, **figure 3.8**, we can see what already was commented. Some people did not walk all 3 gait speeds.

3.2 Development

In order to development the tests, that will be described in future sections, to gain some results, it was used the following tools. Matlab was only used for the objective of constructing the healthy gait references. Python, on the other hand, was used for all the objectives denoted.

Programming languages

- **Python** (Van Rossum *et al.*, 2009).

  – **TensorFlow** (Martín Abadi *et al.*, 2015), **Sklearn** (Pedregosa *et al.*, 2011), **PyQt5**, **Pandas** (team, 2020; McKinney, 2010), **NumPy** (Harris *et al.*, 2020), **Seaborn** (Waskom, 2021) and **Matplotlib** (Hunter, 2007) were used and were vital to develop the tests. They are commonly used in problems related to machine learning.

  – **DTW** (dynamic time warping) package (Giorgino, 2009) was also used to implement the metric.

- **Matlab** (MATLAB, 2021).

  – **Deep Learning Toolbox** (MATLAB, 2021), **Signal Processing Toolbox** (MATLAB, 2021), **ELM Toolbox** (Zhang, 2022), were used to create and run the ELM and RELM algorithm with the proper changes to adapt to the objective.

3.2.1 Optuna

Optuna (Akiba *et al.*, 2019), offers a *define-by-run API*, an efficient implementation for searching and pruning strategies and an easy-to-setup environment. Searching and pruning hyperparameters are required to construct a machine learning algorithm. From many frameworks available, optuna was used because it is recent and easy-to-setup.

Optuna is a tool that was used to improve the speed of implementation and the quality of the tests, because the growth of the complexity of the models implies more time to implement the tests. Because of its importance to the tests, a section was given to it, in the interest of understanding its application. The only algorithms that were not studied with optuna were LR, ELM and RELM because they only have one simple hyperparameter or none (LR case).

Basically, this framework along with the other tools creates the possibility to study the hyperparameters of each algorithm in order to find the best case scenario based on the training dataset. So the objective of using optuna is allowing automatization of tests.

Analyzing the results in optuna's article (Akiba *et al.*, 2019) we can see that optuna choices for algorithms and strategies produce better results than random search.

3.3 A comment on unhealthy people

Unhealthy people weren't deeply studied because it is not necessary that kind of analysis in order to use them. It was presented some understanding in previous sections, but they are important to just present results.

Unhealthy people only did self-selected gait speed. Their first session has the session type as *initial measurement* and it's classified as a gait disorder, or to simplify, pathological. The last session, has the session type as *control measurement* or even *initial measurement after readmission* (Horsak *et al.*, 2020).

3.4 Developed methods

Reading the objective, creating the healthy gait references and then building the diagnosis, creates a necessity of dividing the problem in two different methods. The creation of two apps, also part of the objective, does not need a method. In this section, we will explain the developed methods.

3.4.1 Work on frequency domain

Because of previous works with GRF (Wu *et al.*, 2014; Stergiou *et al.*, 2002; Giakas *et al.*, 1997) worked or studied the possibility of working in the frequency domain, an analysis was planned to see if it is worth it.

So, by using fast fourier transform (Cooley *et al.*, 1965), we are trying to describe the curves with minimum coefficients possible, to have a similarity down to 95 %, obtaining this milestone using the metric DTW. If we use all the coefficients we will have the perfect curve, the best case scenario. If we use none, we have our worst case scenario. This way its possible to normalize and obtain an interval to show the results varying from 0 to 100 %.

Five males and five females will be randomly selected and will be the base to study how many coefficients will be necessary. Since the maximum coefficients is 101, values down to 65 will not be considered worthy of the additional complexity.

3.4.2 Implementing the healthy gait references

Because sex, age, gait speed, height, weight, and foot changes the gait itself (Matsushita *et al.*, 2021b), 4 cases for each algorithm were created. Male left (Ml), Male right (Mr), Female

right (Fr), Female left (Fl). Body mass index, age, and gait speed will be the features that will be used on the algorithms. Body mass index is the person's weight divided by the square of height.

The following **figure 3.9** explains what the problem is. We are going to use three features, age, body max index and gait speed. The values of these three features will represent a session, they are the input, and with the measured gait cycle (called real gait cycle in the **figure 3.9**), the desired output, that have 101 points, train the algorithm in order to find the best representative weights and bias so the algorithm is capable of generalizing and create an output of 101 gait cycle points for an input given.

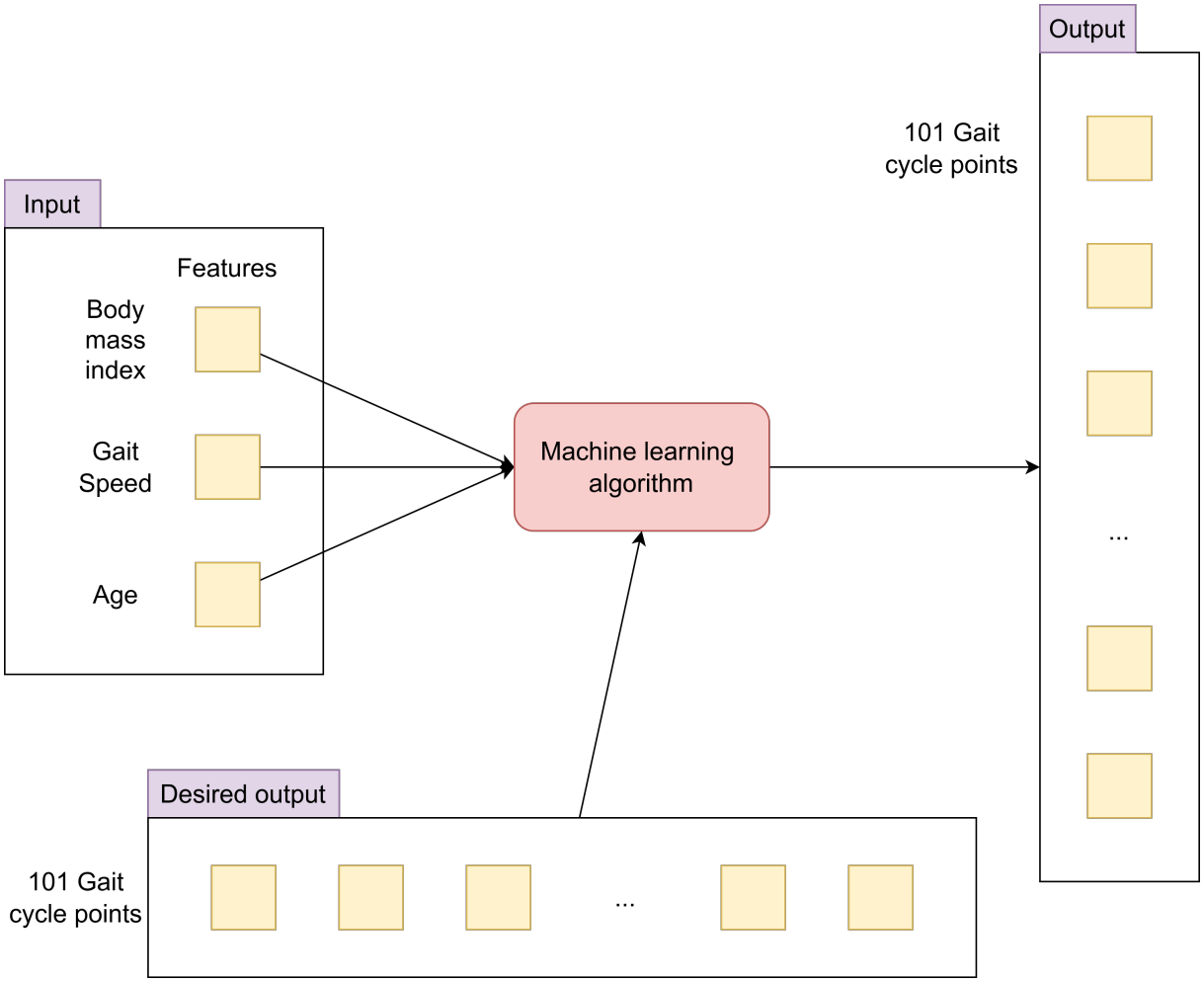


Figure 3.9: Machine learning diagram for the objective. Diagram that describes the structure of the problem, implementing the healthy gait references.

The method created for the objective, implementing the healthy gait references, follows the coming steps:

- **Preprocessing the dataset**: Although the authors of the dataset already did vast steps of preprocessing, its necessary to select what is in favor of the objective. In this case, selecting all the healthy units. Then split the dataset to train and test. We chose 80 % for training and 20 % for test.
- **Implement algorithms**: Based on previous works, Multiple output support vector regression and extreme learning machine were studied (Vieira *et al.*, 2018; Figueiredo *et al.*, 2018; Y. Chang *et al.*, 2021). Linear regression (Stansfield *et al.*, 2006) and Convolution neural network were also studied. In addiction a variation of the extreme machine learning was used, RELM.
- **Analyze the results**:
  - **First Criteria**: Capability of following the true response of a healthy person. A visual comparision between the true response and the result of the created healthy gait reference with the algorithms.
  - **Second Criteria**: With help of the metric dynamic time warping (Kruskall *et al.*, 2022), and with the implementation of dynamic time warping (Giorgino, 2009), give a value that is a distance. Less distance to the real case, better result it is. If they are not good enough, changes in the algorithms needs to be done.
- **Create and present the graphics of the three GRF**: This is a self-explanatory step, when exhausting tests and satisfaction with results happens, presenting the results is a natural step.

3.4.3 Giving a diagnosis

To complete the second part of the objective, this is, building a diagnosis, we need to define some variables that give a measurement that can be understood has a classification. In this case of study, it will be used human gait indexes.

- **Calculate the variables**: Since they were already described on chapter background information, **table 2.1**, this step will be only to calculate the set of variables. But a compromise was done. Simplification. Not all the parameters described in the second chapter are going to be analyzed. This is due to the simplification cause, where we are using the ones that empirically we find relevant and verify if they are enough for building a diagnosis.

  - Forces: F1, F2 and F3;
  - Rates: Load rate and Unload rate;
  - Area: Spectral density and Derivate of the vertical ground reaction force (VGRF) in representation of the VGRF;
  - GVS;
  - DTW;

- **Intervals for classification**:

  First, calculate the GE and SI indexes.

  Then, apply the global error and symmetry index to all healthy units. After build an interval to say, if the result of any session is between this defined interval, consider it healthy or if it doesn't, not healthy.

  This interval will be created by using the mean and standard deviation of the values obtained from the healthy units. So in order to be classified as healthy, both the values of the GE and SI have to be in the intervals of the healthy units.

  Every single value will be normalized between, 0, as the minimum and, maximum being the mean of that corresponding value. For example: The value of GE of an unhealthy male for the left side will be normalized, with maximum being the mean of the GE for the healthy male's left side.

- **Get results and present them**: Presenting the results in favor of getting some conclusions.

Chapter 4

Results and discussion

4.1 Work on the frequency domain

Some works (Wu *et al.*, 2014; Stergiou *et al.*, 2002; Giakas *et al.*, 1997) are working in the frequency domain instead of working on the time domain, or at least studying the possibility of it. Like it was described in the materials and methods chapter, we tried to represent the curves of the ground reaction forces with at least 95 % of similarity with minimum coefficients possible. The similarity is analyzed with the DTW.

This was not possible because the results we got were not in favor (**table 4.1**) since the maximum number of coefficients is 101 and the results are really close to this number.

Table 4.1: Results of the quantity of coefficients necessary for each GRF to have at least 95 % similarity.

	Vertical GRF	**Medio-lateral GRF**	**Anterior-Posterior GRF**
Male left	100	101	101
Male right	100	101	101
Female left	99	101	101
Female right	99	101	101

Clearly, these results don't accomplished the defined rule that was decided in the previous chapter, that the number of coefficients needs to be from 0 to 65 in order to justify the additional complexity. Obviously, the possible advantage of reducing the time by working with smaller representations is not possible in this case.

4.2 The healthy gait references

After the implementation of the algorithms using the materials described, we constructed the code necessary to apply the tool Optuna, so we could automatize the tests in the pursuance of finding the best case scenario for each algorithm.

Since we have 5 algorithms, LR (linear regression), ELM (extreme machine learning), RELM (ridge extreme learning machine), MSVR (multiple output support vector regression) and CNN (convolution neural network); 4 cases, Mr, Ml, Fr, Fl; And three GRF, V, ML, AP. We have 60 objects to analyze. Towards simplification, only the best combination of hyperparameters will be here described.

4.2.1 Linear regression

Linear regression is a simple algorithm in comparison with others that we use here. It doesn't have any hyperparameters to be here listed for any of the 12 cases (Mr, Ml, Fr, Fl and the 3 GRF).

4.2.2 Extreme machine learning

In order to compile ELM, we have only to define 1 hyperparameter. The size of the hidden layer ($\tilde{N}$). To study this hyperparameter, empiric numeric values were used. Bigger values than 500 did not show improvement on the algorithm. So for each 4 cases and 3 GRF (12 cases in total), we studied the interval of 1 to 500, and selected the best instance. The selection is shown on **table 4.2**.

Table 4.2: The best value of the hyperparameter, hidden layer ($\tilde{N}$), from the ELM algorithm for each 12 cases.

GRF	**V**				**ML**				**AP**			
	ML	**MR**	**FL**	**FR**	**ML**	**MR**	**FL**	**FR**	**ML**	**MR**	**FL**	**FR**
$\tilde{N}$	254	113	137	49	283	156	34	38	93	37	119	150

4.2.3 Ridge extreme machine learning

To compile RELM, we have to define 2 hyperparameters. The size of the hidden layer $\tilde{N}$ and the lambda (λ). To have the best hidden layer size was applied the same process of the ELM algorithm. To the lambda hyperparameter a suggested interval, based on the implementation on

MATLAB (Zhang, 2022) by the author of the algorithm RELM and ELM was used. **Table 4.3** shows what values were selected.

Table 4.3: The best value of the hyperparameters, hidden layer ($\tilde{N}$) and lambda (λ), from the RELM algorithm for each 12 cases.

GRF	**V**				**ML**				**AP**			
	ML	**MR**	**FL**	**FR**	**ML**	**MR**	**FL**	**FR**	**ML**	**MR**	**FL**	**FR**
$\tilde{N}$	406	25	288	288	131	108	34	38	350	376	405	152
λ	$2e^{-3}$	$3e^{-4}$	$4e^{-2}$	$2e^{-3}$	$2e^{-3}$	$2e^{-2}$	$3e^{-4}$	$8e^{-4}$	$2e^{-3}$	$6e^{-3}$	$1e^{-2}$	$4e^{-3}$

4.2.4 Multiple output support vector regression

This algorithm has a few hyperparameters, and to come with the best combination, optuna was used. The regularization parameter (C), the kernel function (K) and epsilon (ϵ) were the ones focused. The best values for each hyperparameter studied are presented in the **table 4.4**.

Table 4.4: The best value of the hyperparameters, regularization parameter (C), kernel function (K) and epsilon (ϵ), from the MSVR algorithm for each 12 cases. lin stands for linear, sig stands for sigmoid and rbf stands for radial based function.

GRF	**V**				**ML**				**AP**			
	ML	**MR**	**FL**	**FR**	**ML**	**MR**	**FL**	**FR**	**ML**	**MR**	**FL**	**FR**
C	1925	154	2000	6	20	9712	5621	8995	1241	50	317	180
ϵ	$4e^{-4}$	0.74	0.05	0.06	0.0	0.0	$6e^{-3}$	0.0	0.0	$3e^{-4}$	0.0	0.0
K	sig	rbf	sig	rbf	lin	lin	rbf	lin	lin	sig	lin	lin

4.2.5 Convolution neural network

On CNN, several architectures were tested. It was founded that the simplest architecture possible is the best case scenario. This is, a dense layer and a convolutional layer (figure 4.1). The hyperparameters studied were, filters, activation function and learning rate and the best values are shown in **table 4.5**. Batch size was set to 12 and the epochs were set as 200. This algorithm was also implemented with optuna for the tests. Kernel size was set to 3 and stride was set to 1.

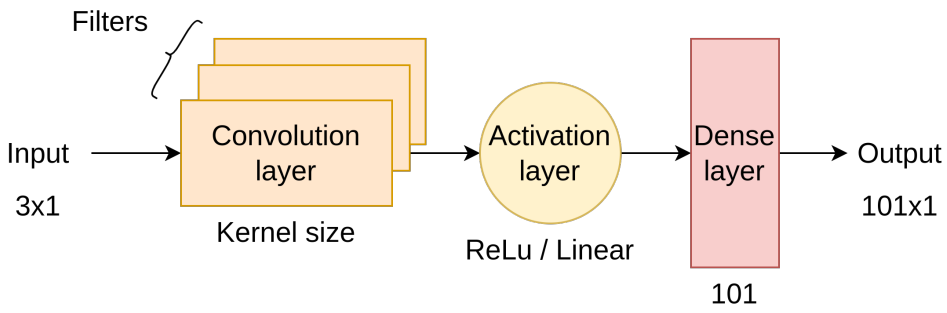


Figure 4.1: Architecture used for the CNN.

Table 4.5: The best value of the hyperparameters, filters, activation function and learning rate, from the CNN algorithm for each 12 cases. R stands for ReLu, Lin for linear, F stands for filter, ActF - activation function and LRat - learning rate.

GRF	**V**				**ML**				**AP**			
	ML	**MR**	**FL**	**FR**	**ML**	**MR**	**FL**	**FR**	**ML**	**MR**	**FL**	**FR**
F	16	16	8	16	4	16	4	8	8	8	8	16
ActF	R	R	Lin	Lin	R	Lin	R	Lin	R	R	R	Lin
LRat	0.07	0.03	0.09	0.06	0.02	0.03	0.09	0.07	0.07	0.01	0.009	0.09

4.3 Criteria of decision for the best machine learning algorithm

4.3.1 The first criteria - Capability of following the true response of a healthy person

The first criteria that is going to be applied is the capability to follow the true response of that specific person based on visual prove. The following **figures, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 4.10, 4.11, 4.12 and 4.13** represent the 4 cases, and are 4 different healthy units, one for each case (4 cases and 3 GRF). Red is LR; Yellow is MSVR; ELM is green; Blue is RELM; Pink is CNN; And black is the true response. To distinguish the 3 ground reaction forces, markers with different forms were used. White triangles for AP; Green hexagons for ML; And black pentagons for V.

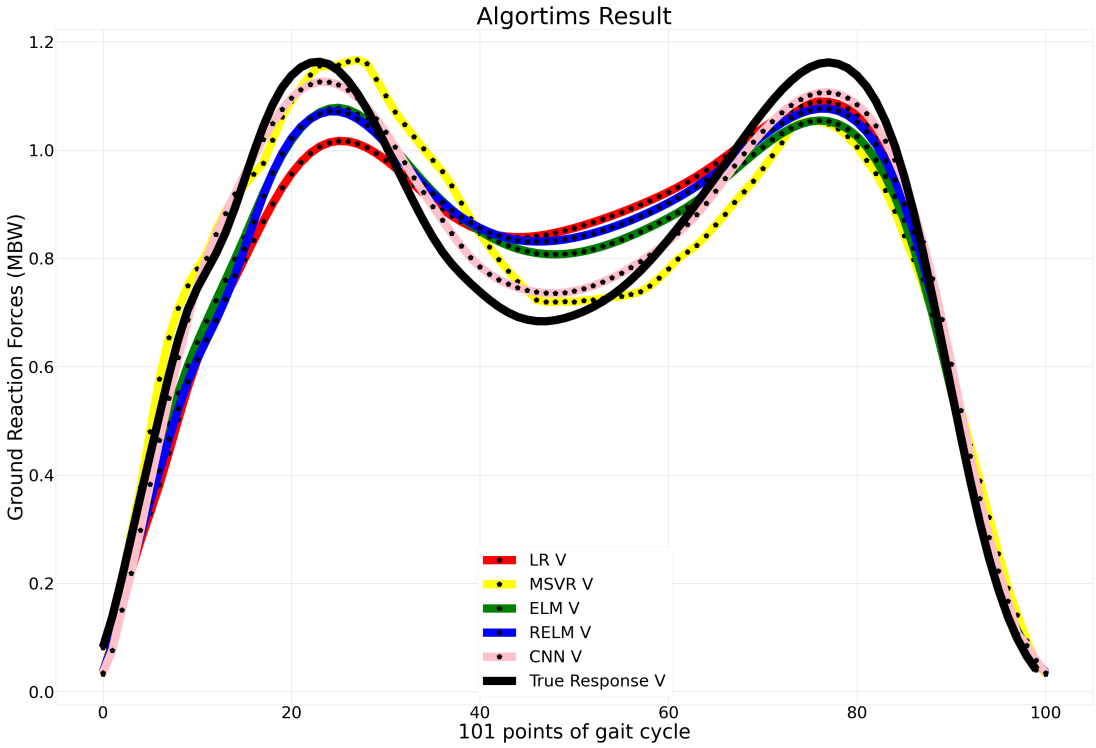


Figure 4.2: Results for creating the healthy gait reference for the case male right, vertical GRF, from all 5 algorithms.

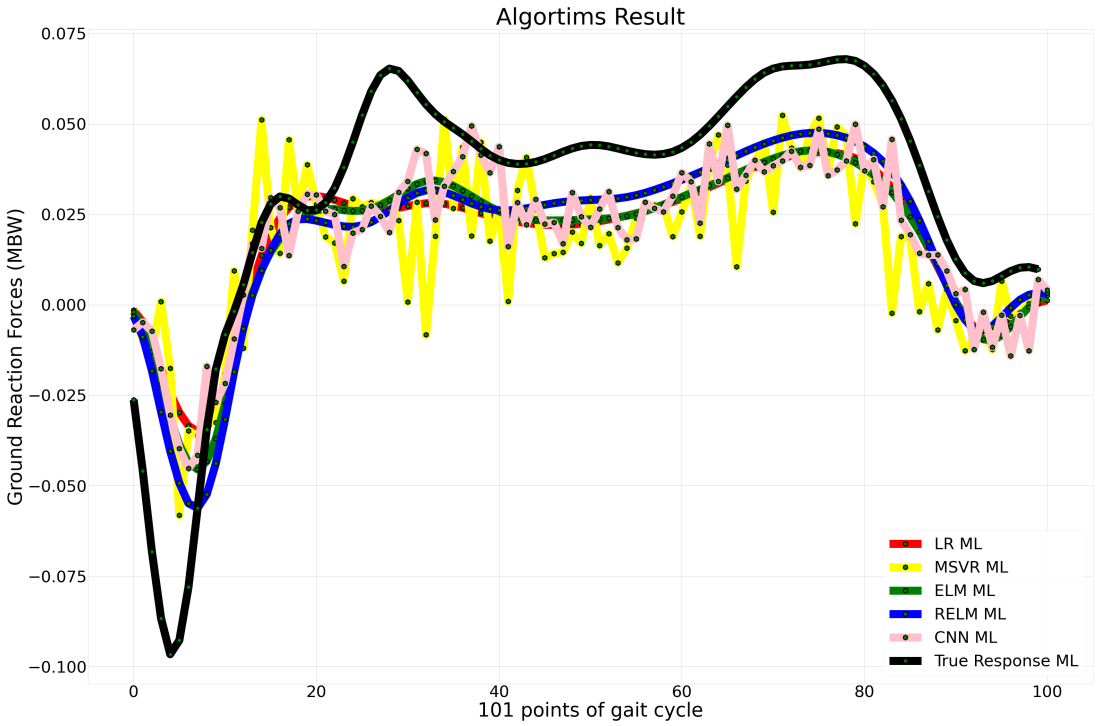


Figure 4.3: Results for creating the healthy gait reference for the case male right, medio-lateral GRF, from all 5 algorithms.

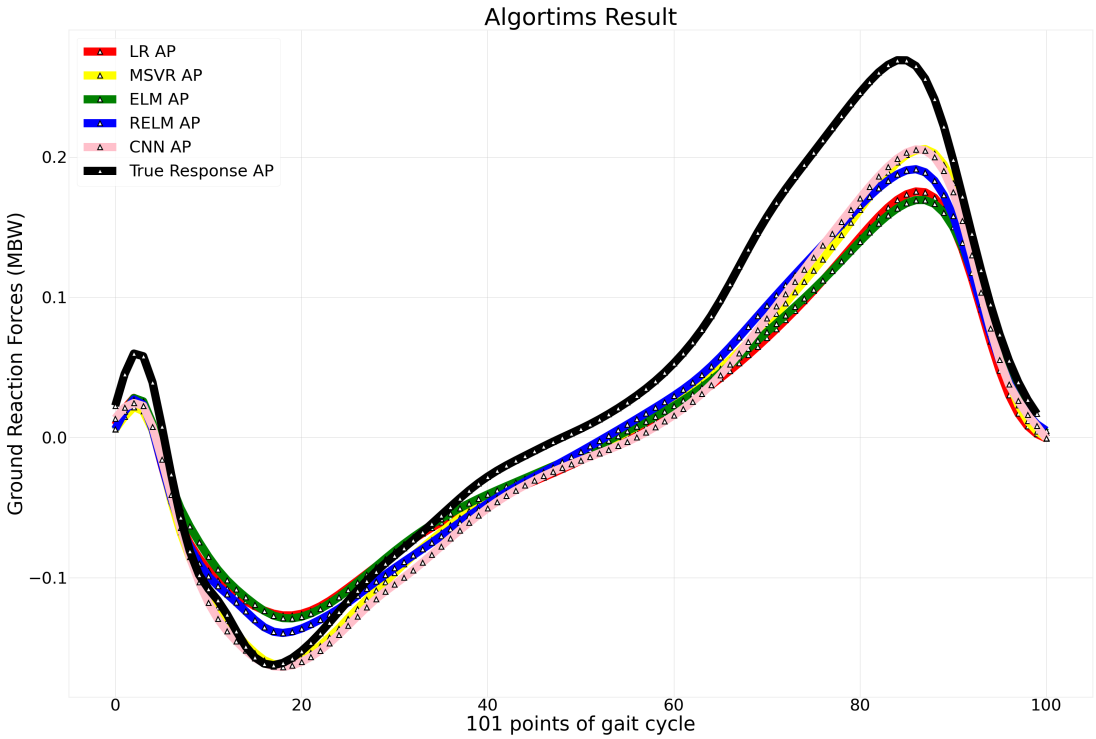


Figure 4.4: Results for creating the healthy gait reference for the case male right, anterior-posterior GRF, from all 5 algorithms.

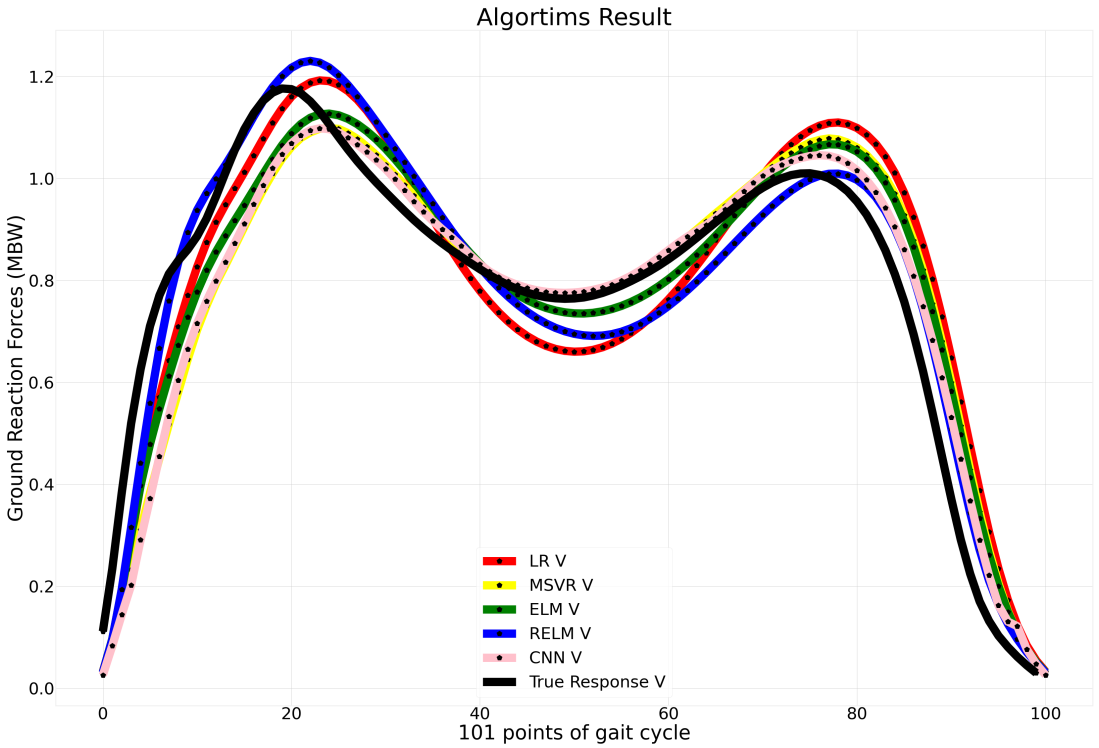


Figure 4.5: Results for creating the healthy gait reference for the case male left, vertical GRF, from all 5 algorithms.

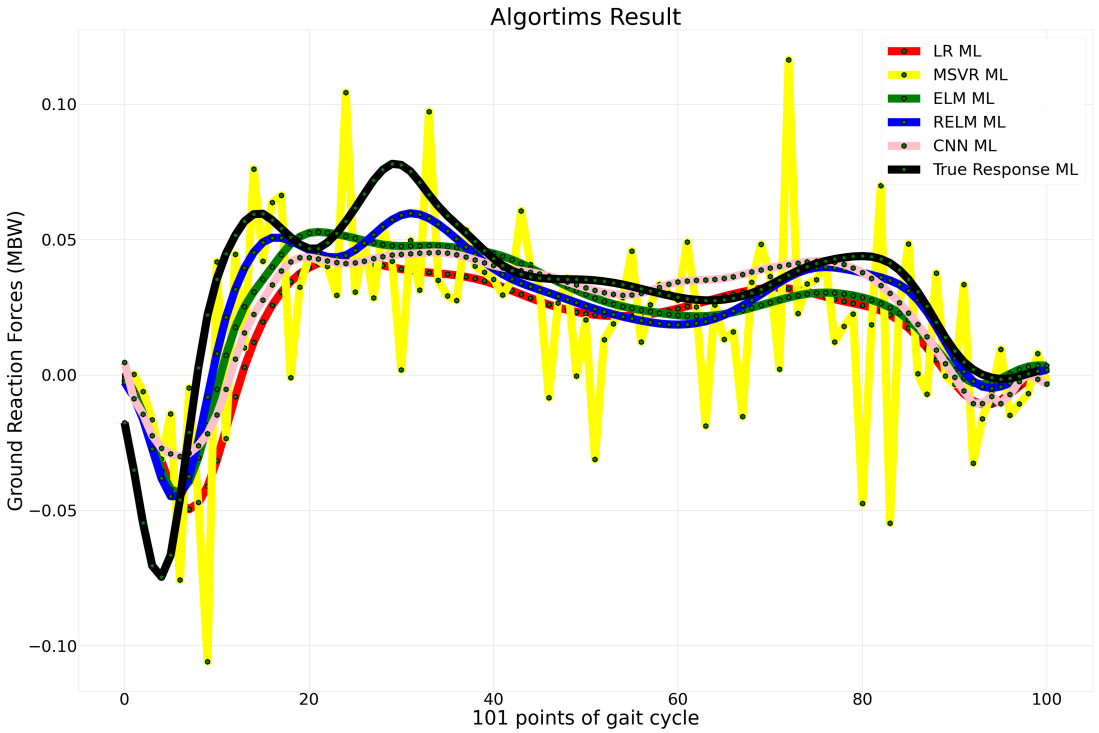


Figure 4.6: Results for creating the healthy gait reference for the case male left, medio-lateral GRF, from all 5 algorithms.

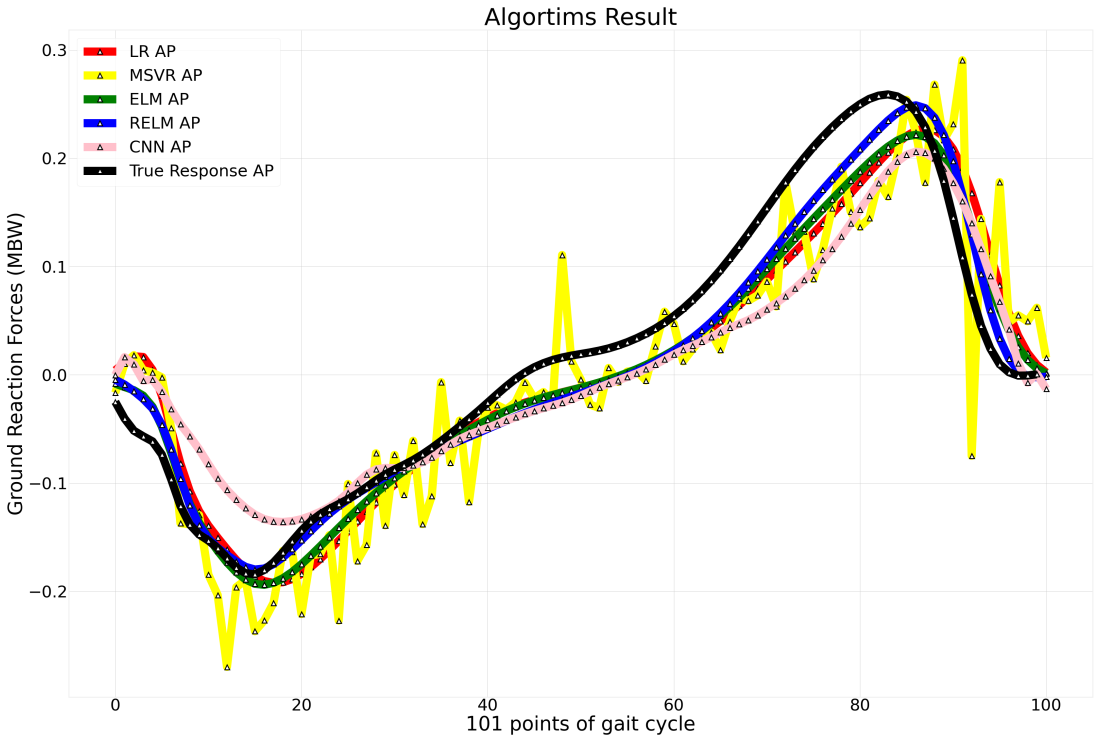


Figure 4.7: Results for creating the healthy gait reference for the case male left, anterior-posterior GRF, from all 5 algorithms.

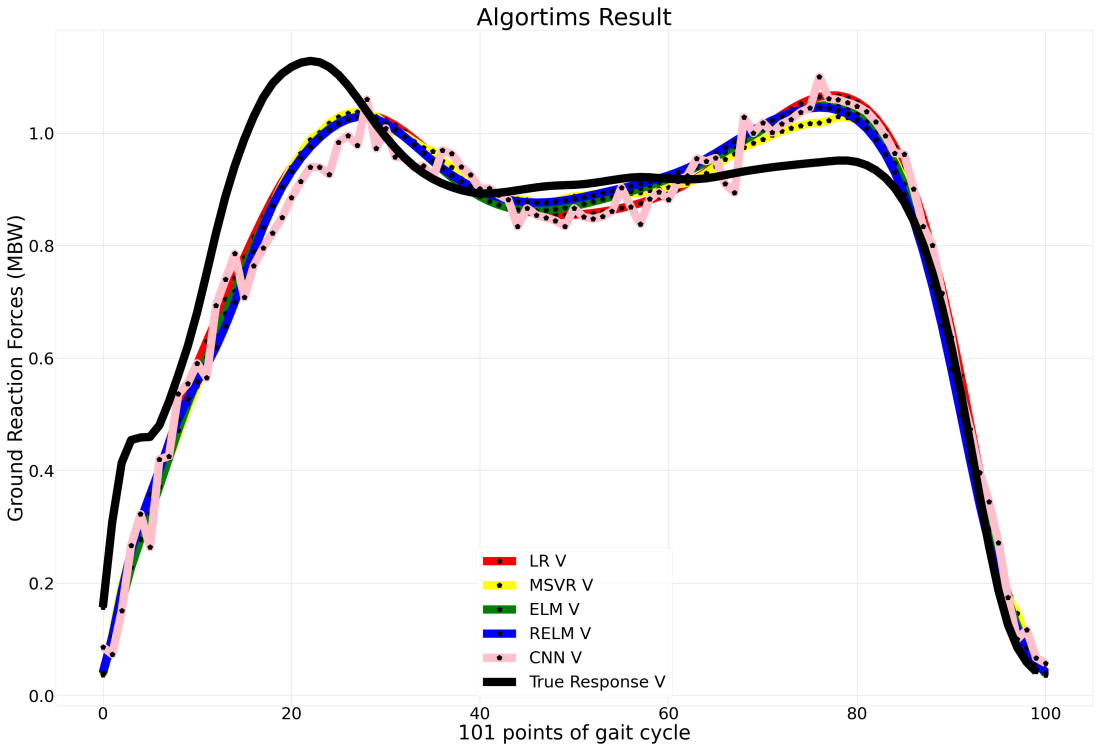


Figure 4.8: Results for creating the healthy gait reference for the case female right, vertical GRF, from all 5 algorithms.

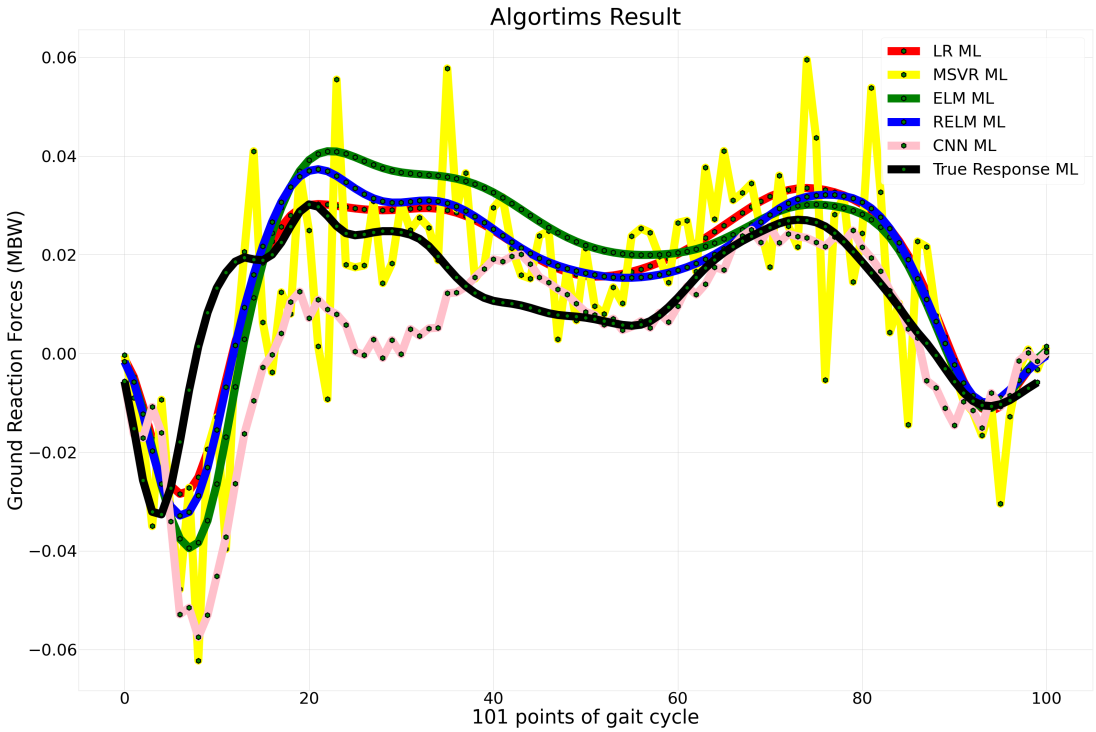


Figure 4.9: Results for creating the healthy gait reference for the case female right, medio-lateral GRF, from all 5 algorithms.

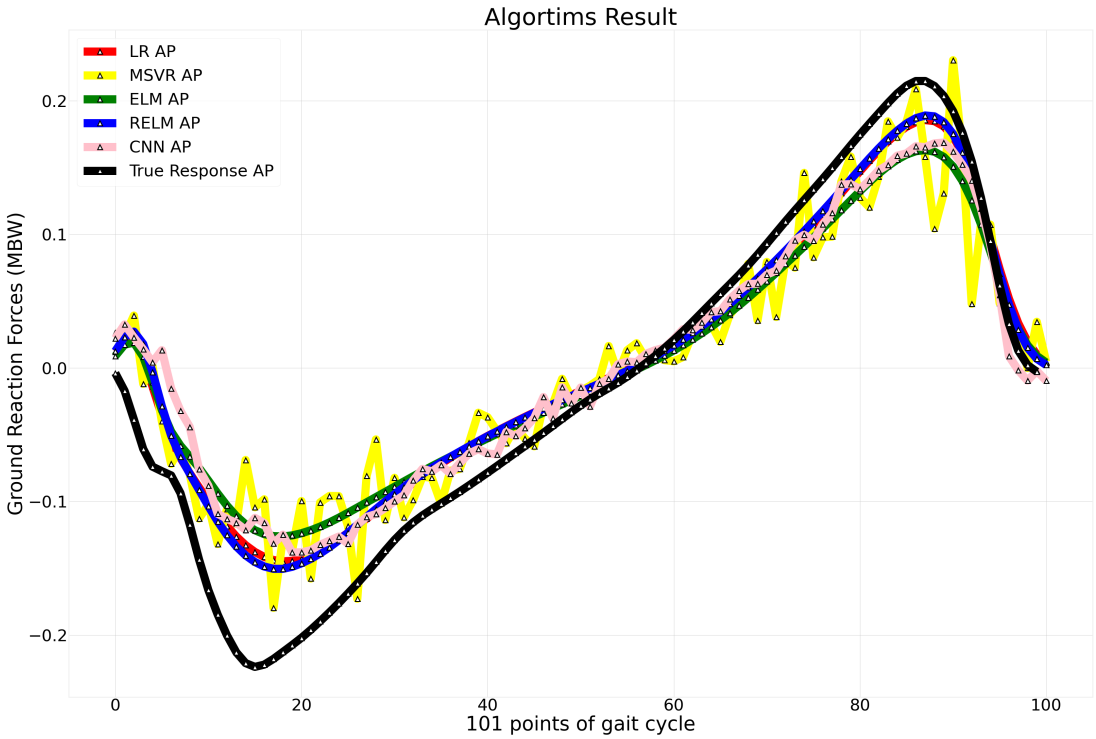


Figure 4.10: Results for creating the healthy gait reference for the case female right, anterior-posterior GRF, from all 5 algorithms.

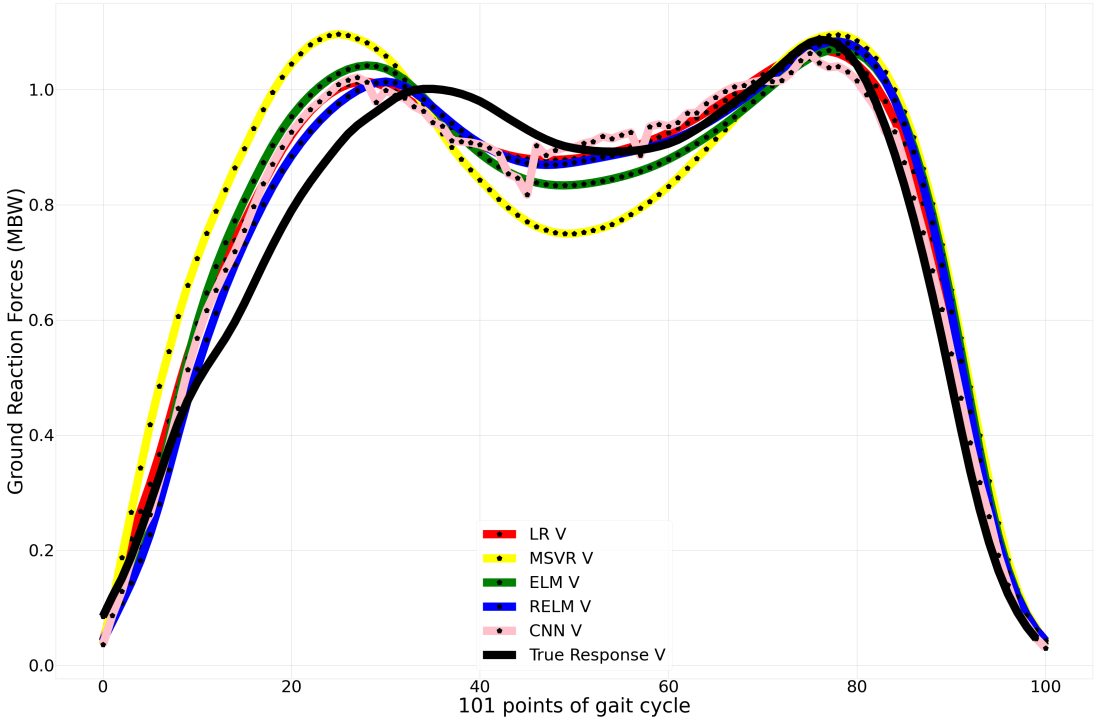


Figure 4.11: Results for creating the healthy gait reference for the case female left, vertical GRF, from all 5 algorithms.

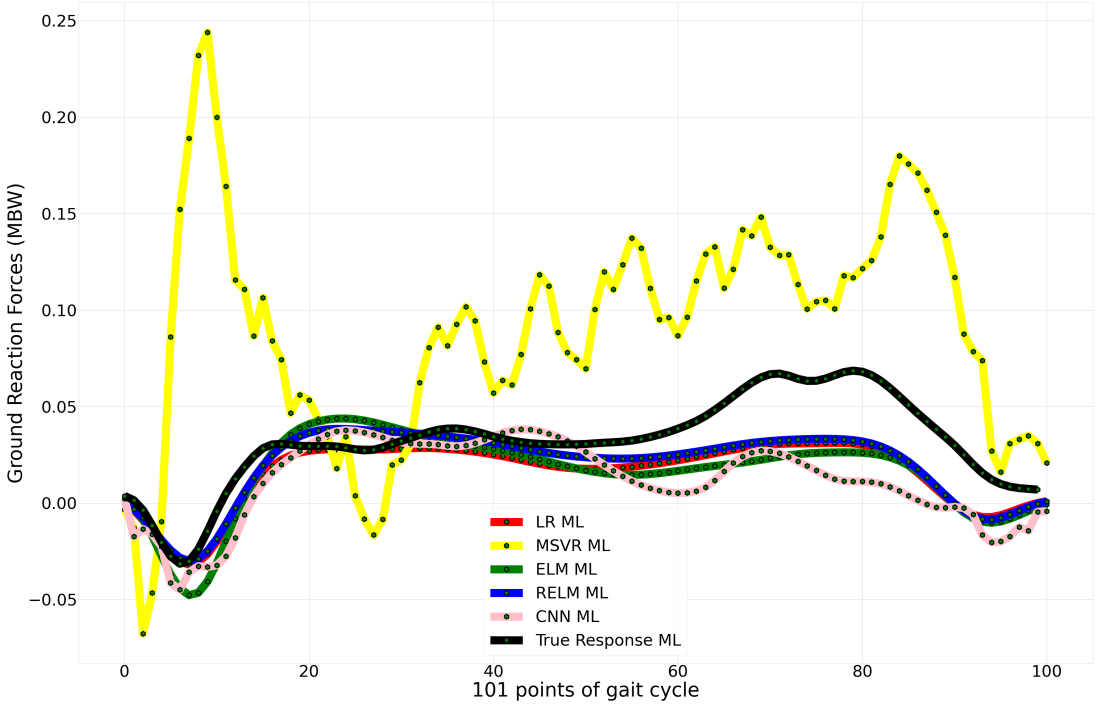


Figure 4.12: Results for creating the healthy gait reference for the case female left, medio-lateral GRF, from all 5 algorithms.

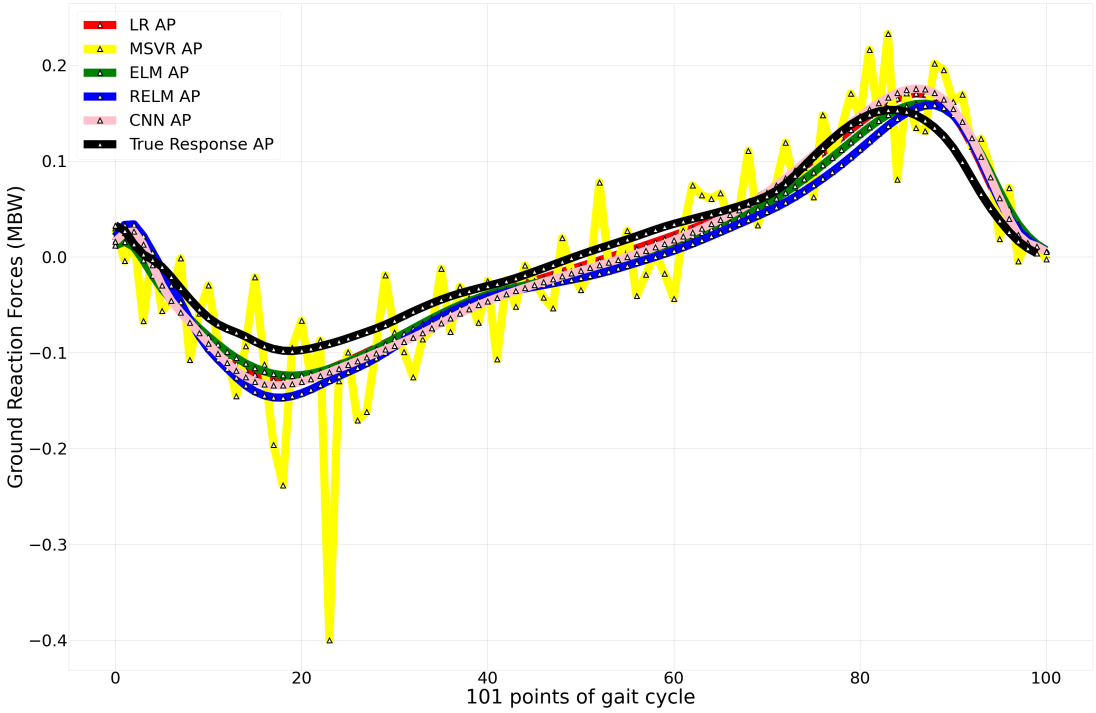


Figure 4.13: Results for creating the healthy gait reference for the case female left, anterior-posterior GRF, from all 5 algorithms.

ELM, RELM and LR were capable of following the true response of the healthy person in consideration for all the three GRF. On the other hand, it's clear to see that MSVR has the worst response in this criteria for all three GRF. We can see spikes along the gait cycle on MSVR, in the **figure 4.3, 4.6, 4.7, 4.9, 4.10, 4.12 and 4.13**, which in the true response, spikes, don't exist. CNN also doesn't behave correctly. Spikes, in CNN response, are present on the **figures 4.3, 4.8, 4.9, 4.10 and 4.11**. In general, CNN and MSVR, fail to learn the true/desired response.

Since we can observe that we have other algorithms that have better response than CNN and MSVR for all 3 GRF and these two, CNN and MSVR, fail at the mission to be similar to the real response these two will be off consideration.

4.3.2 The second criteria - The least distance between the true response and the created one

The second criteria passes by examining the values of the metric DTW for all 5 algorithms, 4 cases and 3 GRF, between the true response and the created healthy gait reference for all healthy units. The following tables, **table 4.6, 4.7, 4.8 and 4.9**, present the mean of the DTW on all healthy units and the training time of the algorithms.

Table 4.6: Results of the second criteria for the case Mr.

GRF	V		ML		AP	
	Distance (DTW)	Training Time (s)	Distance (DTW)	Training Time (s)	Distance (DTW)	Training Time (s)
LR	2.998	0.039	0.663	0.005	0.976	0.006
ELM	2.437	0.003	0.579	0.005	0.802	0.001
RELM	2.567	0.026	0.581	0.037	0.786	0.160
MSVR	4.765	0.066	1.623	6.122	1.320	14.888
CNN	4.432	28.287	1.209	26.918	1.245	27.740

Table 4.7: Results of the second criteria for the case Ml.

GRF	V		ML		AP	
	Distance (DTW)	Training Time (s)	Distance (DTW)	Training Time (s)	Distance (DTW)	Training Time (s)
LR	2.884	0.144	0.646	0.026	0.957	0.012
ELM	2.368	0.013	0.539	0.011	0.824	0.002
RELM	2.220	0.255	0.536	0.042	0.797	0.126
MSVR	4.475	16.222	1.810	5.815	2.589	6.431
CNN	4.341	34.826	0.758	26.044	1.286	26.467

After we have excluded the MSVR and CNN, we have left LR, ELM and RELM. Looking at the values from the tables for all 4 cases (**table 4.6, 4.7, 4.8 and 4.9**) we conclude that ELM and RELM are the best ones and don't really get distance from each other in terms of the DTW value. Linear regression presents a very good response, staying really close to the values of ELM and RELM in general. But we see a significant unit difference in all 4 cases. So because of this reason, ELM and RELM are considered better than LR, taking out of consideration the algorithm LR.

Table 4.8: Results of the second criteria for the case Fr.

GRF	V		ML		AP	
	Distance (DTW)	**Training Time (s)**	**Distance (DTW)**	**Training Time (s)**	**Distance (DTW)**	**Training Time (s)**
LR	3.536	0.030	0.677	0.011	1.113	0.007
ELM	2.490	0.001	0.783	0.001	0.919	0.004
RELM	2.308	0.127	0.725	0.022	1.102	0.058
MSVR	3.152	14.069	3.152	14.069	2.723	7.104
CNN	4.288	27.782	1.081	26.567	1.295	26.284

Table 4.9: Results of the second criteria for the case Fl.

GRF	V		ML		AP	
	Distance (DTW)	**Training Time (s)**	**Distance (DTW)**	**Training Time (s)**	**Distance (DTW)**	**Training Time (s)**
LR	3.633	0.019	0.650	0.011	1.119	0.004
ELM	2.565	0.004	0.709	0.001	0.973	0.012
RELM	2.403	0.121	0.709	0.029	0.957	0.212
MSVR	4.563	8.810	6.018	770.758	3.007	6.857
CNN	3.250	27.531	0.813	26.049	1.081	25.636

The only major difference between ELM and RELM is the training time. RELM takes more time to training, even though the difference only occurs from millisecond (ELM) to hundredth of a second (RELM). So these two criteria will not be enough to decide each one to choose.

One thing that is evident is that simplicity is the key here. Simpler algorithms do better results. Although ELM and RELM are more close to the complex side of the spectrum, there is a huge difference between CNN, MSVR and ELM and RELM in terms of complexity.

4.4 ELM vs RELM

ELM was chosen to be the algorithm that will create the healthy gait references. RELM has more complexity inherent to it. Although the time that both took to train are minimal, ELM has even smaller times. This is obviously due to RELM having more complexity to overcome the numerical instability of the pseudo-inverse, like it was already commented. But the difference on performance in terms of both criteria, here implemented, are almost null. So, avoiding complexity

is a plus in this case, since in this case complexity didn't show better performance.

4.5 The diagnosis

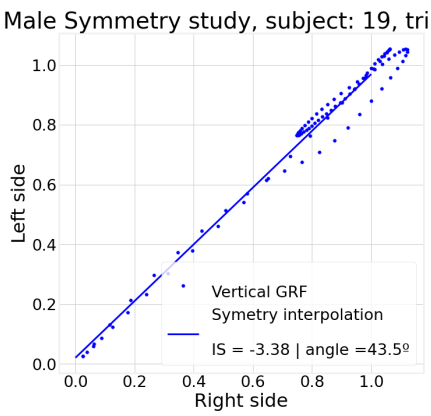
In favor of giving a diagnosis, human gait indexes were used, naming, Symmetric Index and Global Error. Also, using DTW as tool for classification was analyzed. Besides, analysis of the distribution of each index was also made, so decisions could have a starter ground.

Also on DTW and GE, results using the mean reference, besides machine learning reference, will be presented. The mean reference is the mean of all healthy unit for each 4 cases.

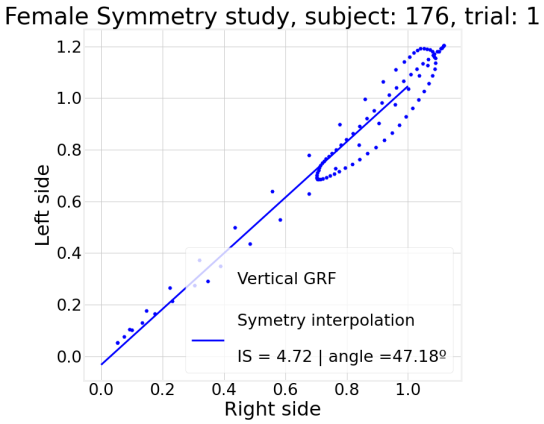
4.5.1 Symmetric index

The symmetric index is built by analyzing one person's gait, for left and right side, and trying to find a curve that fits them. On x-axis, we chose the right side and on the y-axis we chose the left so it's normal to have negative symmetric index values. In this case, it will mean that in that session, the dominant side (higher in terms of force) is the left side. Otherwise, if the value is positive, the right side is the dominant. There is no information on the dataset that informs if the person is left or right dominant, so no further conclusions can be taken. The only one is that on a specific session, the person was dominant in one side.

A healthy person will not have a perfect symmetric index, which in this case would necessarily mean that the person has a perfect $45°$ angle for the created curve. Studies already show that perfect symmetry does not exist (Sadeghi *et al.*, 2000).



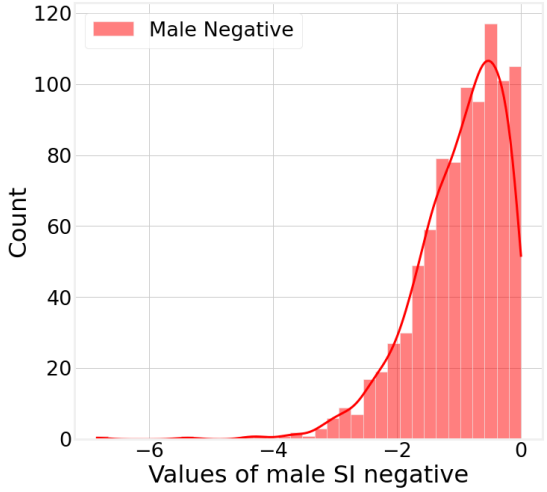
(a) Subject ID:19 and trial: 7. The result angle is: $43.5°$.



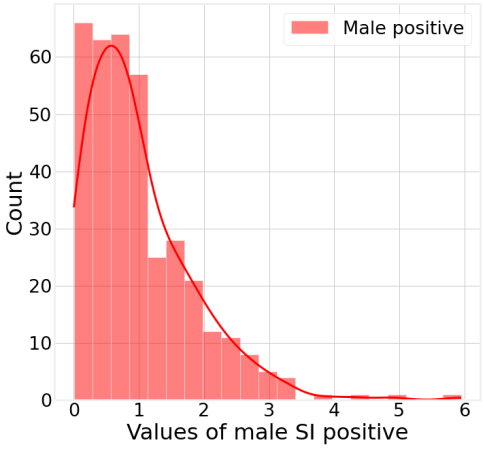
(b) Subject ID:176 and trial: 1. The result angle is: $47.18°$.

Figure 4.14: Symmetric index for a random healthy male **(a)** and for a random healthy female **(b)**.

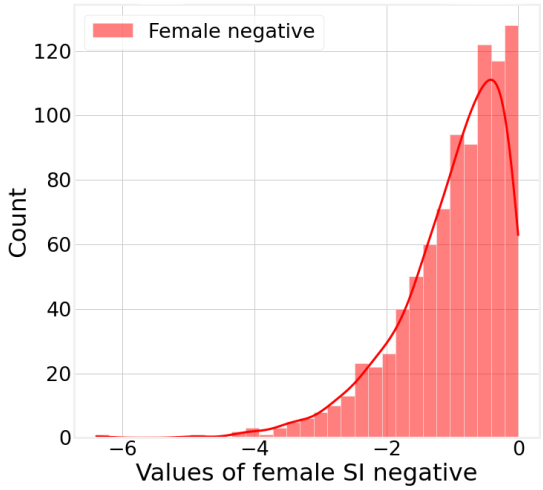
Examining the results for those two healthy units (**figure 4.14**), it is reinforced that not all healthy units will have $45°$ angle. It is expect a range of values for the healthy units. This way, we can construct the interval for the symmetric index that classifies a person as a healthy unit. Like it was said in the materials and methods chapter, it will be used the mean and the standard deviation to create this interval.



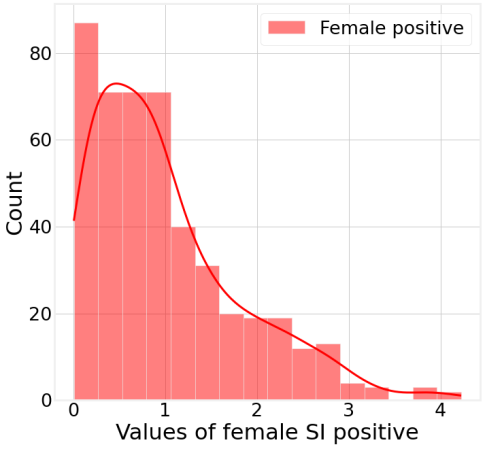
(a) Symmetric index distribution for males with negative values.



(b) Symmetric index distribution for males with positive values.



(c) Symmetric index distribution for females with negative values.



(d) Symmetric index distribution for females with positive values.

Figure 4.15: Symmetric index distribution for all healthy males (negative or positive SI) and all healthy females (negative or positive SI).

There was a necessity to separate the negative and positive values of SI. This is due to the fact that it is common sense that exists healthy units that are more dominant on the left or on the right. So the mean of the positives will be different from the mean of the negatives, both for

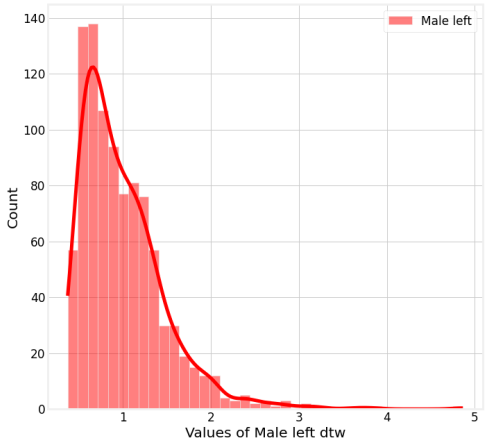
male and female.

These distributions (**figure 4.15**) will be considered similar to the Gaussian distribution, so the intervals that will be created will be with the mean and two times the standard deviation in favor of getting at least 95 % of all the cases.

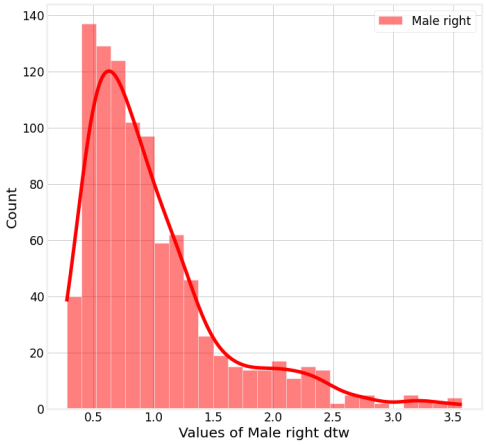
4.5.2 Dynamic time warping

Although DTW is a variable for the global error and not an index itself, research on how it would behave on classifying was done.

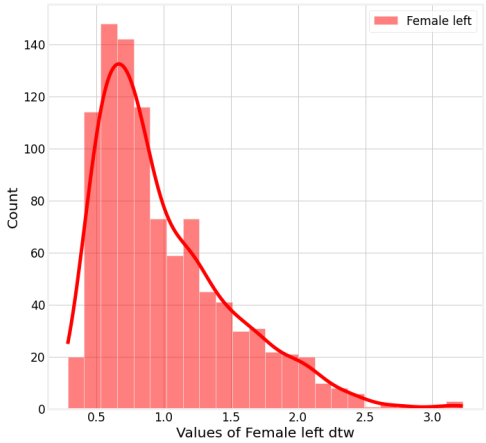
Again, the distributions (**figure 4.16**) are a close response to a Gaussian one. So to define the interval for the DTW, the mean and two times the standard deviation was used.



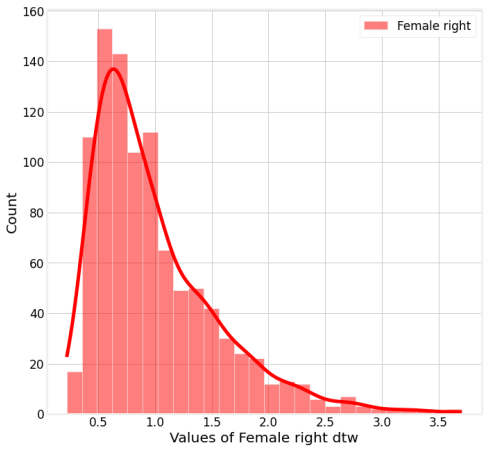
(a) DTW distribution for males, left side.



(b) DTW distribution for males, right side.

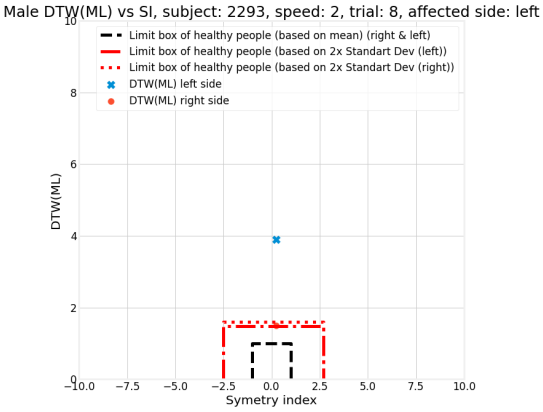


(c) DTW distribution for females, left side.

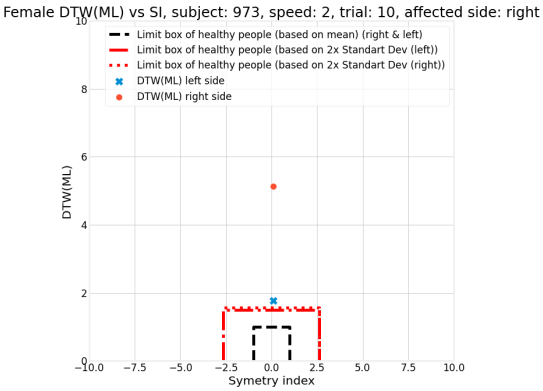


(d) DTW distribution for females, right side.

Figure 4.16: DTW distribution for healthy all males and females, left and right.

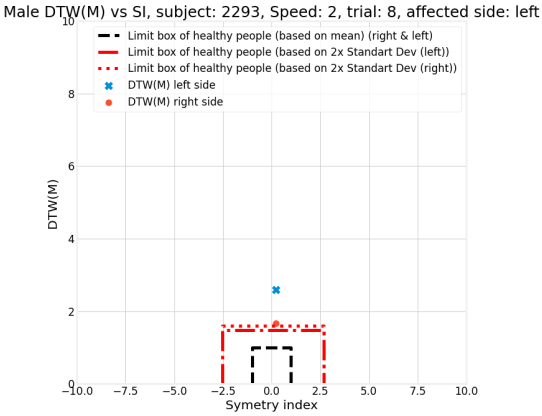


(a) DTW result of a random unhealthy male.

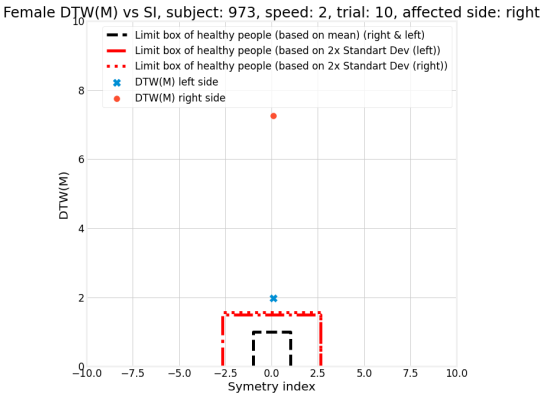


(b) DTW result of a random unhealthy female.

Figure 4.17: DTW results for random unhealthy male and female. Here, (ML) means machine learning. The DTW and SI values are normalized. Including the boxes.



(a) DTW result of a random unhealthy male.



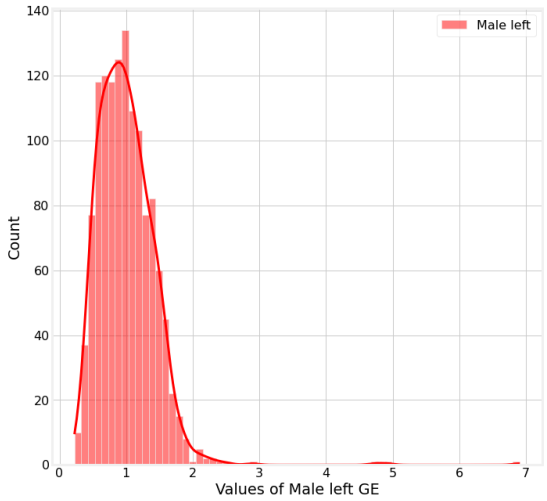
(b) DTW result of a random unhealthy female.

Figure 4.18: DTW results for random unhealthy male and female. Here, (M) means mean. The DTW and SI values are normalized. Including the boxes.

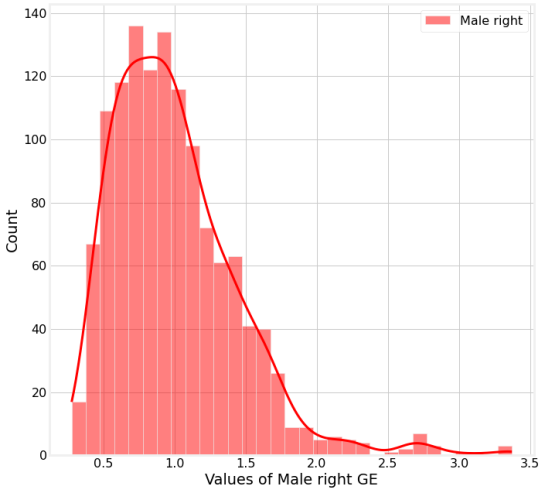
The application of the intervals is to create limits boxes on the **figures 4.17 and 4.18**. This way we can graphically see where the values are. And if it is inside those limit boxes, we can consider them as healthy. Because DTW is not going to be a tool to classification and only an interest to see how it would act, no further information about how it did to other patients is going to be presented. Yet, it was capable of performing correct diagnosis on this two random unhealthy male and female.

4.5.3 Global error

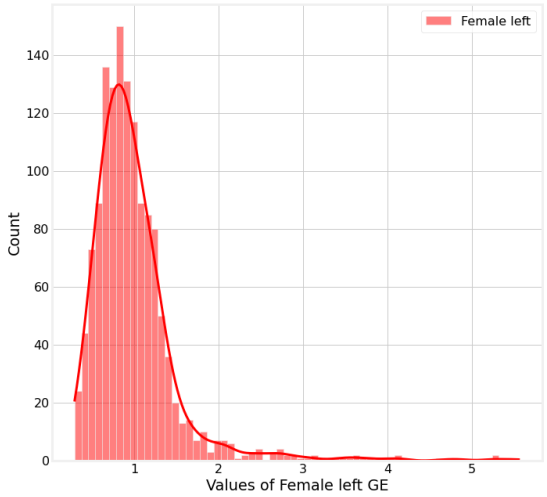
The distribution of the GE for healthy units (**figure 4.19**), also can be considered as a Gaussian distribution. Therefore, the intervals will be constructed the same way as the SI and DTW.



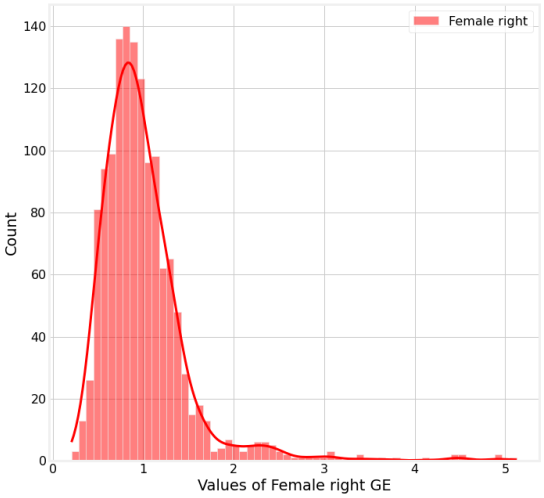
(a) GE distribution for males, left side.



(b) GE distribution for males, right side.

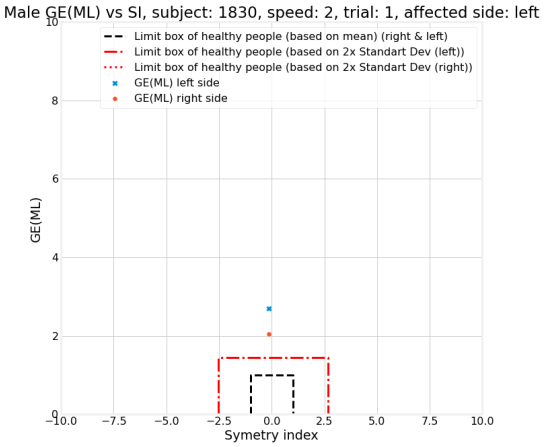


(c) GE distribution for females, left side.

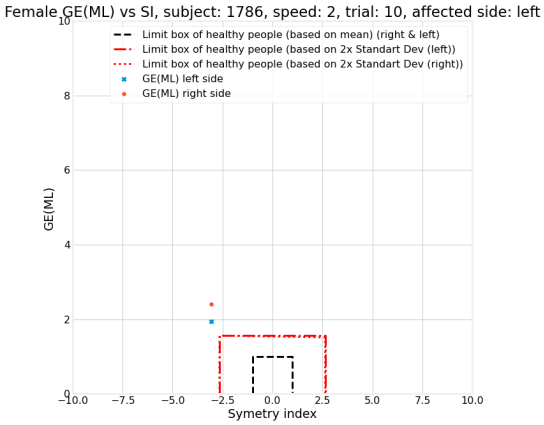


(d) GE distribution for females, right side.

Figure 4.19: GE distribution for healthy males and females, left and right.

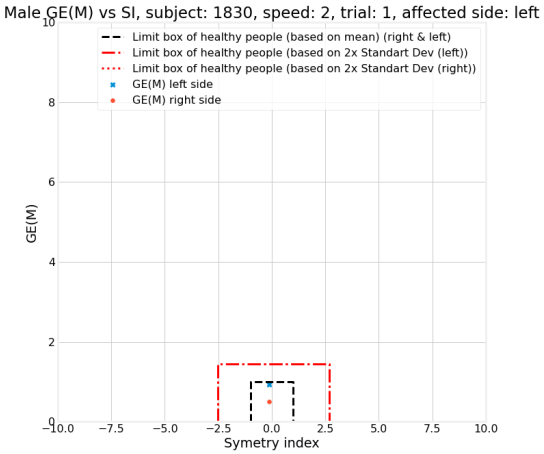


(a) GE result of a random unhealthy male.

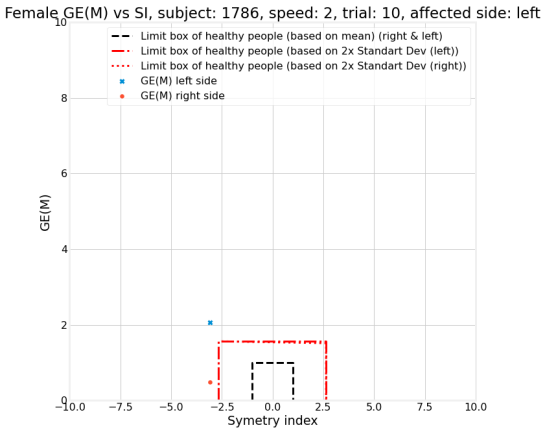


(b) GE result of a random unhealthy female.

Figure 4.20: GE results for random unhealthy male and female. Here, (ML) means machine learning. The GE and SI values are normalized. Including the boxes.



(a) GE result of a random unhealthy male.



(b) GE result of a random unhealthy female.

Figure 4.21: GE results for random unhealthy male and female. Here, (M) means mean. The GE and SI values are normalized. Including the boxes.

It's implicit that GE was capable of classifying this two random unhealthy male and female well based on the **figures 4.20 and 4.21**. It is important to clarify thatm T1, T2 and T3 value always changes, because every single human gait is different, and sometimes, the behavior of the minimum, T2, is different from what is expected (Lafuente *et al.*, 1998; Mezghani *et al.*, 2007). So what was done is use the mean of the T1, T2 and T3, for each 4 cases Ml, Mr, Fl, Fr. This way it takes out the dependency of the form/pattern of the vertical ground reaction force. And even though we do not use the value of T1, T2 and T3, they are necessary to give the value of F1, F2 and F3.

Considering only the first measurement to be sure that the classification as unhealthy is

accurate, applying GE to all unhealthy patients it was built the **table 4.10**.

Table 4.10: Results of the accuracy by classify unhealthy units as unhealthy. ML means the reference is built from machine learning and M means the reference is built from mean.

	GE(ML) and SI	GE(ML)	GE(M) and SI	GE(M)
Correctly classified Males	100 %	85.8 %	100 %	82.8 %
Correctly classified Females	100 %	62.1 %	100 %	63.9 %

The results demonstrated in the **table 4.10**, proves that the use of both indexes, SI and GE, has 100 % accuracy. We can classify all unhealthy units as unhealthy. So using both determines a necessity. The GE itself has an accuracy bigger than 80 % for males, but for females, has only an accuracy bigger than 60 %.

4.6 Why the mean and standard deviation for the intervals?

Although referenced by the authors of the dataset GaitRec (Horsak *et al.*, 2020) that, the dataset was passed through an algorithm to take out the outliers, we still found some responses that could be considered as an outlier. This is influencing on the machine learning response and when analyzing the mean and the maximums and the minimums of and index for the healthy dataset we can see a huge variance (DTW although not an index was also presented here).

Table 4.11: SI results for the healthy subdataset. Max is the maximum value found. Min is the maximum value found. PM is the positive SI case mean. Pstd positive SI case standard deviation. NM is the negative SI case mean. Nstd negative SI case standard deviation.

SI	Max	Min	PM	PStd	NM	NStd
Male	-19.4	9.4	1.6	1.3	-2.8	2.1
Female	-16.3	7.3	1.7	1.4	-2.5	2.1

The results show a high value for the standard deviation (**tables 4.11, 4.12 and 4.13**). This means that we have some outliers that passed. So it only made sense to create the intervals based on the mean and with two times the standard deviation, so we can consider at least 95 % of the healthy units.

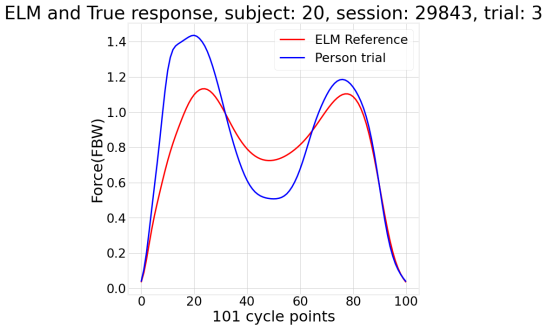
The **figure 4.22**, shows the worst cases for the DTW values (the maximum for the 4 cases Mr, Ml, Fr and Fl).

Table 4.12: DTW results for the healthy subdataset. Max is the maximum distance found. Min is the maximum distance found. Std means standard deviation.

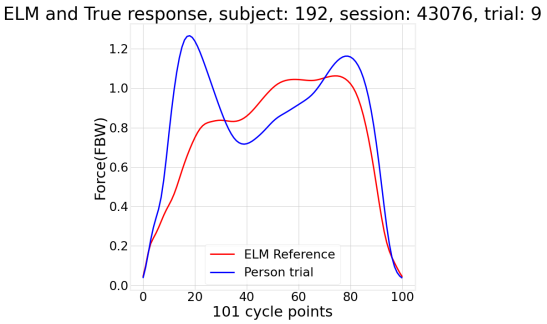
	Left				Right			
DTW	**Max**	**Min**	**Mean**	**Std**	**Max**	**Min**	**Mean**	**Std**
Male	10.4	0.8	2.1	1.0	8.8	0.7	2.5	1.5
Female	8.0	0.7	2.5	1.2	9.6	0.6	2.6	1.4

Table 4.13: GE results for the healthy subdataset. Max is the maximum value found. Min is the maximum value found. Std means standard deviation.

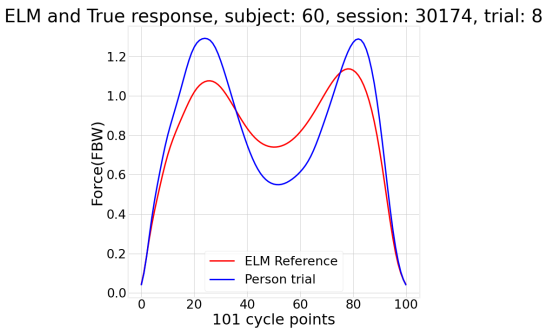
	Left				Right			
GE	**Max**	**Min**	**Mean**	**Std**	**Max**	**Min**	**Mean**	**Std**
Male	1.2	0.04	0.2	0.08	0.6	0.05	0.2	0.08
Female	1.1	0.06	0.2	0.1	1.0	0.04	0.2	0.1



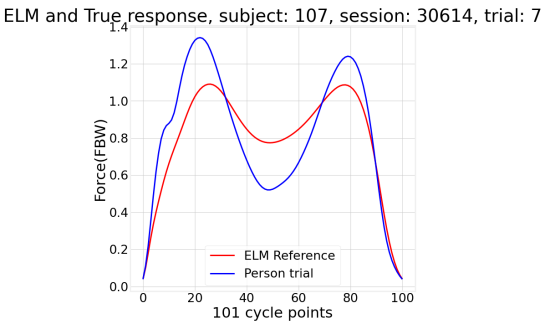
(a) Worst case scenario for male left based on DTW value.



(b) Worst case scenario for male right based on DTW value.



(c) Worst case scenario for female left based on DTW value.



(d) Worst case scenario for female right based on DTW value.

Figure 4.22: Worst case scenarios, based on DTW for all 4 cases. In blue, we have the true response and in red we have the machine learning reference.

4.7 APP

Two applications were built. One for showing all algorithms and their results for a random new person, called Reference Graphs app. And another that shows all unhealthy persons, their evolution, the created healthy gait reference with machine learning, the indexes, and the DTW metric, called Gait Analysis app. Both apps were built in python and the tools used are already described at the materials and methods chapter.

4.7.1 Reference graphs

This app was built, so every algorithm studied could be presented to a random person. The **figure 4.23** shows the created app. It has fields that need to be filled and selected in order to fulfill the features that are being studied, age, gait speed, body max index, female or male and left or right side. Then, when selected the algorithm that the user wants to see to produce the healthy gait reference and pressed the button generate, the graphic for all three forces is shown.

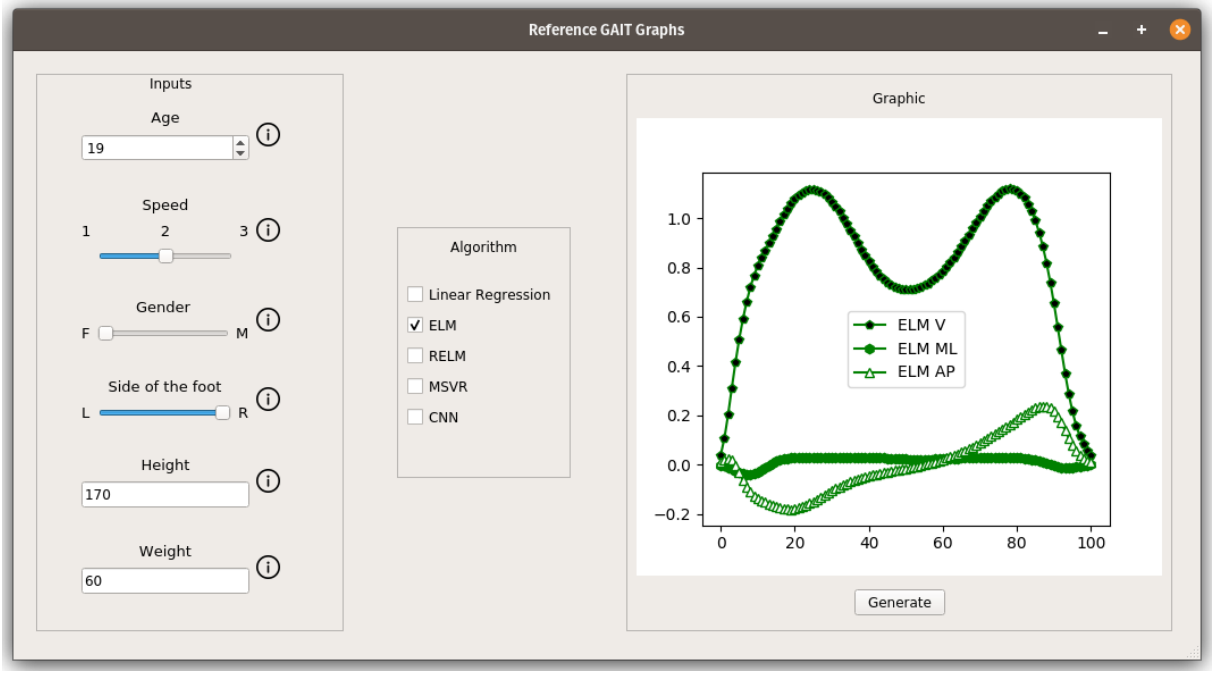


Figure 4.23: First app, the Reference Graphs app. Created to show the healthy gait references by all studied algorithms.

4.7.2 Gait analysis

Gait Analysis (**figure 4.24**), was created to demonstrate all the results and studies done here in one place. On this app, we see the healthy gait references created for a specific person to

all three forces. The evolution of the features, age, body max index (gait speed, because the subjects are unhealthy patients, is always 2). Left and right side and the metric DTW. Also, a global analysis is shown with the human gait indexes values and graphics (**figure 4.25**).

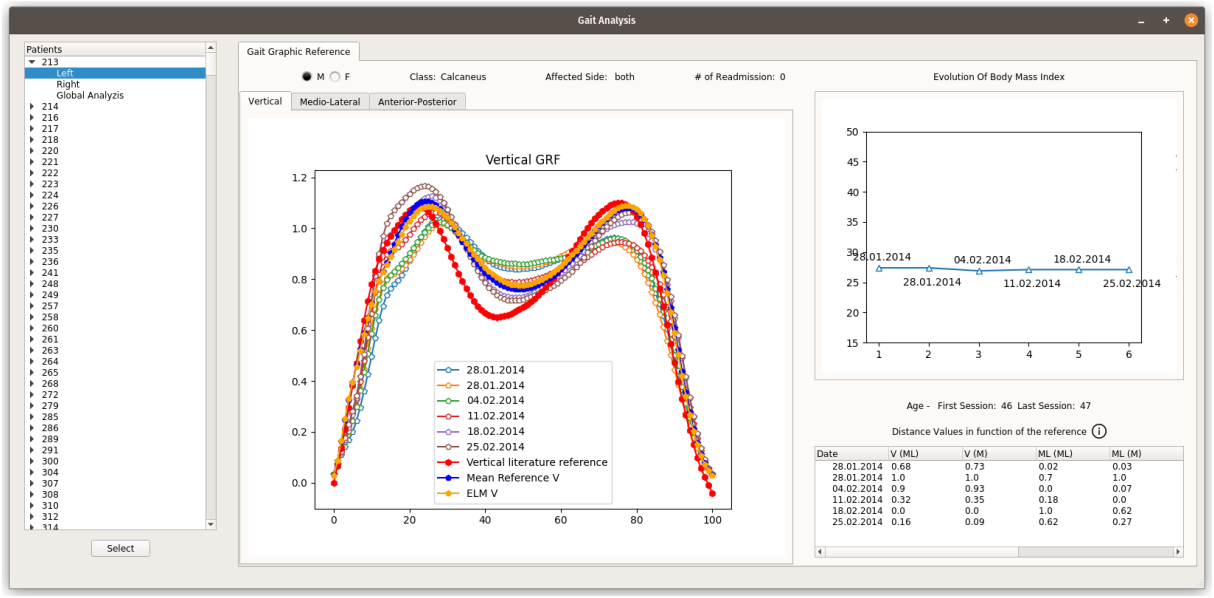


Figure 4.24: Second app, the Gait Analysis app. Here we can see that the subject 213 was selected, and we are watching the results for the left side.

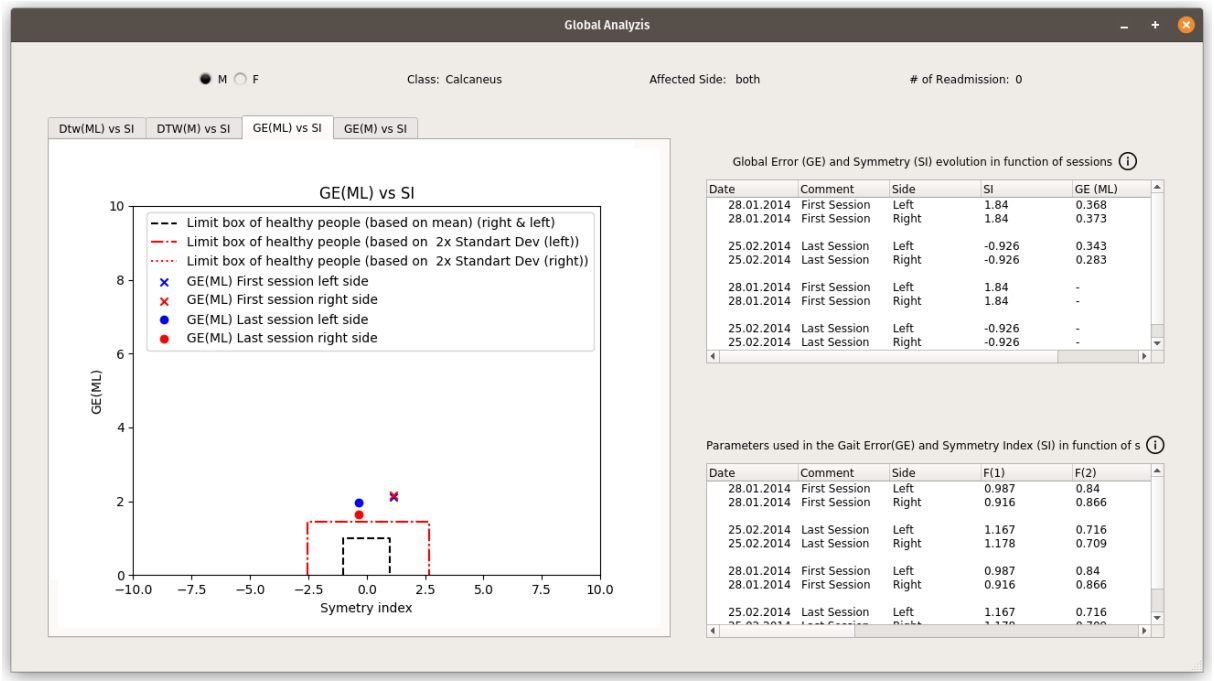


Figure 4.25: Second app, the Gait Analysis app. Here we can see that the subject 213 was selected, and we are watching the results for the global analysis.

Chapter 5

Conclusion

The study performed about the possibility of working on the frequency domain did not favor it. Working on the frequency domain could help get results faster since the size of the problem could be smaller (instead of 101 points, a smaller quantity of coefficients was plausible), but it didn't produce a smaller size.

ELM proves to be the best algorithm tested, like it was also shown on (Vieira *et al.*, 2018), even tough we have a better dataset, GaitRec. We also have seen that simplicity proves to be better, and ELM and RELM although in the boarder line of complexity still runs far away from the level of complexity inherent to CNN, for example. The two criteria here used and described should have been enough to decide what the best algorithm was, yet, because ELM and RELM are really similar, there was a necessity to another criteria in order to take the conclusion that ELM is the best. The criteria used was complexity (winning the less complex algorithm).

The necessity of using both human gait indexes, like we used, proves it self by the fact that it had 100 % of accuracy when using both. GE by himself does not have a good results for females, staying around 60 %, and, for males has values higher than 80 %, both for machine learning reference and mean reference. Because of this, GE might benefit all 16 variables for possible higher results. Because we need both the GE and SI to have 100 %, finding a way to combine both of these indexes into one is a possible path to be considered as future work.

The mean reference, used on human gait indexes, it's not a targeted reference, and the machine learning algorithms do that. This is a major voting cause that proves that machine learning is still the best case for this problem. Taking out the outliers discovered with the DTW could improve the response of the machine learning.

In culmination, it was possible to build healthy gait references to a target unit with machine learning and give a diagnosis with two indexes, GE and SI.

References

AKIBA, Takuya; SANO, Shotaro; YANASE, Toshihiko; OHTA, Takeru; KOYAMA, Masanori, 2019. Optuna: A Next-generation Hyperparameter Optimization Framework. In: *Proceedings of the 25rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

ALAQTASH, Murad; SARKODIE-GYAN, Thompson; YU, Huiying; FUENTES, Olac; BROWER, Richard; ABDELGAWAD, Amr, 2011. Automatic classification of pathological gait patterns using ground reaction forces and machine learning algorithms. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 453–457. Available from doi: 10.1109/IEMBS.2011.6090063.

Appendix A: Kinematic, Kinetic, and Energy Data, 2009. In: *Biomechanics and Motor Control of Human Movement*. John Wiley & Sons, Ltd, pp. 296–360. isbn 9780470549148. Available from doi: https://doi.org/10.1002/9780470549148.app1.

BAKER, Richard; MCGINLEY, Jennifer L.; SCHWARTZ, Michael H.; BEYNON, Sarah; ROZUMALSKI, Adam; GRAHAM, H. Kerr; TIROSH, Oren, 2009. The Gait Profile Score and Movement Analysis Profile. *Gait & Posture*. Vol. 30, no. 3, pp. 265–269. issn 0966-6362. Available from doi: https://doi.org/10.1016/j.gaitpost.2009.05.020.

CAO, Jiuwen; ZHANG, Kai; LUO, Minxia; YIN, Chun; LAI, Xiaoping, 2016. Extreme learning machine and adaptive sparse representation for image classification. *Neural Networks*. Vol. 81, pp. 91–102. issn 0893-6080. Available from doi: https://doi.org/10.1016/j.neunet.2016.06.001.

CHANG, Chih-Chung; LIN, Chih-Jen, 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*. Vol. 2, 27:1–27:27. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

CHANG, Ying; WANG, Lan; LIN, Lingjie; LIU, Ming, 2021. Research on SVM gait recognition based on single acceleration sensor. In: *2021 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE)*, pp. 147–150. Available from doi: 10.1109/MLISE54096.2021.00033.

COOLEY, James W.; TUKEY, John W., 1965. An Algorithm for the Machine Calculation of Complex Fourier Series. *Mathematics of Computation* [online]. Vol. 19, no. 90, pp. 297–301 [visited on 2022-10-01]. issn 00255718, issn 10886842. Available from: http://www.jstor.org/stable/2003354.

FIGUEIREDO, Joana; SANTOS, Cristina P.; MORENO, Juan C., 2018. Automatic recognition of gait patterns in human motor disorders using machine learning: A review. *Medical Engineering & Physics*. Vol. 53, pp. 1–12. issn 1350-4533. Available from doi: https://doi.org/10.1016/j.medengphy.2017.12.006.

GÉRON, Aurélien., 2019. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*. 2st. O'Reilly Media, Inc, Sebastopol. isbn 9781492032649.

GIAKAS, Giannis; BALTZOPOULOS, Vasilios, 1997. Time and frequency domain analysis of ground reaction forces during walking: an investigation of variability and symmetry. *Gait & Posture*. Vol. 5, no. 3, pp. 189–197. issn 0966-6362. Available from doi: https://doi.org/10.1016/S0966-6362(96)01083-1.

GIORGINO, Toni, 2009. Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package. *Journal of Statistical Software*. Vol. 31, no. 7, pp. 1–24. Available from doi: 10.18637/jss.v031.i07.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron, 2016. *Deep Learning*. MIT Press. http://www.deeplearningbook.org.

HARRIS, Charles R.; MILLMAN, K. Jarrod; WALT, Stéfan J. van der; GOMMERS, Ralf; VIRTANEN, Pauli; COURNAPEAU, David; WIESER, Eric; TAYLOR, Julian; BERG, Sebastian; SMITH, Nathaniel J.; KERN, Robert; PICUS, Matti; HOYER, Stephan; KERKWIJK, Marten H. van; BRETT, Matthew; HALDANE, Allan; RÍO, Jaime Fernández del; WIEBE, Mark; PETERSON, Pearu; GÉRARD-MARCHANT, Pierre; SHEPPARD, Kevin; REDDY, Tyler; WECKESSER, Warren; ABBASI, Hameer; GOHLKE, Christoph; OLIPHANT, Travis E., 2020. Array programming with NumPy. *Nature*. Vol. 585, no. 7825, pp. 357–362. Available from doi: 10.1038/s41586-020-2649-2.

HORSAK, Brian; SLIJEPCEVIC, Djordje; RABERGER, Anna-Maria; SCHWAB, Caterine; WORISCH, Marianne; ZEPPELZAUER, Matthias, 2020. GaitRec, a large-scale ground reaction force dataset of healthy and impaired gait. *Scientific Data*. Vol. 7, no. 1, p. 143. issn 2052-4463. Available from doi: 10.1038/s41597-020-0481-z.

HUANG, Guang-Bin; ZHU, Qin-Yu; SIEW, Chee-Kheong, 2006. Extreme learning machine: Theory and applications. *Neurocomputing*. Vol. 70, no. 1, pp. 489–501. issn 0925-2312. Available from doi: https://doi.org/10.1016/j.neucom.2005.12.126. Neural Networks.

HUNTER, J. D., 2007. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*. Vol. 9, no. 3, pp. 90–95. Available from doi: 10.1109/MCSE.2007.55.

KHARB, Ashutosh; SAINI, Vipin; JAIN, Y; DHIMAN, Surender; TECH, M; SCHOLAR, 2011. A review of gait cycle and its parameters. *IJCEM Int J Comput Eng Manag*. Vol. 13.

KRUSKALL, J.; LIBERMAN, Mark, 2022. The symmetric time warping algorithm: From continuous to discrete.

LAFUENTE, R; BELDA, JM; SÁNCHEZ-LACUESTA, J; SOLER, C; PRAT, J, 1998. Design and test of neural networks and statistical classifiers in computer-aided movement analysis: a case study on gait analysis. *Clinical Biomechanics*. Vol. 13, no. 3, pp. 216–229. issn 0268-0033. Available from doi: https://doi.org/10.1016/S0268-0033(97)00082-X.

LIANG, Shengyun; NING, Yunkun; LI, Huiqi; WANG, Lei; MEI, Zhanyong; MA, Yingnan; ZHAO, Guoru, 2015. Feature Selection and Predictors of Falls with Foot Force Sensors Using KNN-Based Algorithms. *Sensors (Basel)*. Vol. 15, no. 11, pp. 29393–29407.

MARSLAND, S., 2014. *Machine Learning: An Algorithmic Perspective*. 2st. Chapman and Hall/CRC. isbn 1107123526. Available from doi: https://doi.org/10.1201/b17476.

MARTÍN ABADI; ASHISH AGARWAL; PAUL BARHAM; EUGENE BREVDO; ZHIFENG CHEN; CRAIG CITRO; GREG S. CORRADO; ANDY DAVIS; JEFFREY DEAN; MATTHIEU DEVIN; SANJAY GHEMAWAT; IAN GOODFELLOW; ANDREW HARP; GEOFFREY IRVING; MICHAEL ISARD; JIA, Yangqing; RAFAL JOZEFOWICZ; LUKASZ KAISER; MANJUNATH KUDLUR; JOSH LEVENBERG; DANDELION MANÉ; RAJAT MONGA; SHERRY MOORE; DEREK MURRAY; CHRIS OLAH; MIKE SCHUSTER; JONATHON SHLENS; BENOIT STEINER; ILYA SUTSKEVER; KUNAL TALWAR; PAUL TUCKER; VINCENT VANHOUCKE; VIJAY VASUDEVAN; FERNANDA VIÉGAS; ORIOL VINYALS; PETE WARDEN; MARTIN WATTENBERG; MARTIN WICKE; YUAN YU; XIAOQIANG ZHENG, 2015. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Available also from: https://www.tensorflow.org/. Software available from tensorflow.org.

MATLAB, 2021. *9.11.0.1873467 (R2021b)*. Natick, Massachusetts: The MathWorks Inc.

MATSUSHITA, Yume; TRAN, Dinh Tuan; YAMAZOE, Hirotake; LEE, Joo-Ho, 2021a. Recent use of deep learning techniques in clinical applications based on gait: a survey. *Journal*

of Computational Design and Engineering. Vol. 8, no. 6, pp. 1499–1532. issn 2288-5048. Available from doi: 10.1093/jcde/qwab054.

MATSUSHITA, Yume; TRAN, Dinh Tuan; YAMAZOE, Hirotake; LEE, Joo-Ho, 2021b. Recent use of deep learning techniques in clinical applications based on gait: a survey. *Journal of Computational Design and Engineering*. Vol. 8, no. 6, pp. 1499–1532. issn 2288-5048. Available from doi: 10.1093/jcde/qwab054.

MCKINNEY, Wes, 2010. Data Structures for Statistical Computing in Python. In: WALT, Stéfan van der; MILLMAN, Jarrod (eds.). *Proceedings of the 9th Python in Science Conference*, pp. 56–61. Available from doi: 10.25080/Majora-92bf1922-00a.

MCMULKIN, Mark L.; MACWILLIAMS, Bruce A., 2015. Application of the Gillette Gait Index, Gait Deviation Index and Gait Profile Score to multiple clinical pediatric populations. *Gait & Posture*. Vol. 41, no. 2, pp. 608–612. issn 0966-6362. Available from doi: https://doi.org/10.1016/j.gaitpost.2015.01.005.

MEZGHANI, Neila; BOIVIN, Karine; TURCOT, Katia; HAGEMEISTER, N; AISSAOUI, Rachid; GUISE, Jacques de, 2007. Asymptomatic and knee osteoarthritis automatic gait pattern analysis using a wavelet representation of kinetic data and the nearest neighbor classifier, pp. 6–8.

NORDIN, NH; MUTHALIF, Asan; RAZALI, M., 2018. Control of transtibial prosthetic limb with magnetorheological fluid damper by using a fuzzy PID controller. *Journal of Low Frequency Noise, Vibration and Active Control*. Vol. 37, p. 146134841876617. Available from doi: 10.1177/1461348418766171.

NUTTALL, Frank Q, 2015. Body Mass Index: Obesity, BMI, and Health: A Critical Review. *Nutr Today*. Vol. 50, no. 3, pp. 117–128.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E., 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. Vol. 12, pp. 2825–2830.

SADEGHI, Heydar; ALLARD, Paul; PRINCE, François; LABELLE, Hubert, 2000. Symmetry and limb dominance in able-bodied gait: a review. *Gait & Posture*. Vol. 12, no. 1, pp. 34–45. issn 0966-6362. Available from doi: https://doi.org/10.1016/S0966-6362(00)00070-9.

SANCHEZ-FERNANDEZ, M.; DE-PRADO-CUMPLIDO, M.; ARENAS-GARCIA, J.; PEREZ-CRUZ, F., 2004. SVM multiregression for nonlinear channel estimation in multiple-input multiple-output systems. *IEEE Transactions on Signal Processing*. Vol. 52, no. 8, pp. 2298–2307. Available from doi: 10.1109/TSP.2004.831028.

SCHUTTE, L.M.; NARAYANAN, U.; STOUT, J.L.; SELBER, P.; GAGE, J.R.; SCHWARTZ, M.H., 2000. An index for quantifying deviations from normal gait. *Gait & Posture*. Vol. 11, no. 1, pp. 25–31. issn 0966-6362. Available from doi: https://doi.org/10.1016/S0966-6362(99)00047-8.

SCHWARTZ, Michael H.; ROZUMALSKI, Adam, 2008. The gait deviation index: A new comprehensive index of gait pathology. *Gait & Posture*. Vol. 28, no. 3, pp. 351–357. issn 0966-6362. Available from doi: https://doi.org/10.1016/j.gaitpost.2008.05.001.

SOBRAL, H.; FERREIRA, J. P.; VIEIRA, A.; PAULO COIMBRA, A.; CRISÓSTOMO, M.; LEMOS, P.; PINHEIRO, J.; LIU, Tao, 2018. Two New Indices to Assess Gait Disturbances Applied to Anterior Cruciate Ligament Reconstructed Knees. In: *2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, pp. 701–706. Available from doi: 10.1109/CYBER.2018.8688159.

STANSFIELD, B.W.; HILLMAN, S.J.; HAZLEWOOD, M.E.; ROBB, J.E., 2006. Regression analysis of gait parameters with speed in normal children walking at self-selected speeds. *Gait & Posture*. Vol. 23, no. 3, pp. 288–294. issn 0966-6362. Available from doi: https://doi.org/10.1016/j.gaitpost.2005.03.005.

STERGIOU, Nicholas; GIAKAS, Giannis; BYRNE, Jennifer E.; POMEROY, Valerie, 2002. Frequency domain characteristics of ground reaction forces during walking of young and elderly females. *Clinical Biomechanics*. Vol. 17, no. 8, pp. 615–617. issn 0268-0033. Available from doi: https://doi.org/10.1016/S0268-0033(02)00072-4.

TEAM, The pandas development, 2020. *pandas-dev/pandas: Pandas*. Zenodo. Latest. Available from doi: 10.5281/zenodo.3509134.

VAN ROSSUM, Guido; DRAKE, Fred L., 2009. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace. isbn 1441412697.

VIEIRA, Alexandra; RIBEIRO, Bernardete; SOBRAL, Heloísa; COIMBRA, A. Paulo; CRISÓSTOMO, Manuel; FERREIRA, João P.; LIU, Tao, 2018. Dynamic Human Gait VGRF Reference Profile Generation via Extreme Learning Machine. In: *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7. Available from doi: 10.1109/IJCNN.2018.8489062.

WASKOM, Michael L., 2021. seaborn: statistical data visualization. *Journal of Open Source Software*. Vol. 6, no. 60, p. 3021. Available from doi: 10.21105/joss.03021.

WATT, Jeremy; BORHANI, Reza; KATSAGGELOS, Aggelos K., 2016. *Machine Learning Refined: Foundations, Algorithms, and Applications*. 1st. USA: Cambridge University Press. isbn 1107123526.

WU, Jianhua; BEERSE, Matthew; AJISAFE, Toyin, 2014. Frequency domain analysis of ground reaction force in preadolescents with and without Down syndrome. *Research in Developmental Disabilities*. Vol. 35, no. 6, pp. 1244–1251. issn 0891-4222. Available from doi: https://doi.org/10.1016/j.ridd.2014.03.019.

ZHANG, Kai, 2022. *Code for the Extreme learning machine and adaptive sparse representation algorithm (EA-SRC)*. MATLAB Central File Exchange. Available also from: https://www.mathworks.com/matlabcentral/fileexchange/58005-code-for-the-extreme-learning-machine-and-adaptive-sparse-representation-algorithm-ea-src.