



Università Politecnica delle Marche

Department of Information Engineering
Master Degree in Biomedical Engineering

Classification of hemiplegia by machine learning interpretation of EMG signal

Candidate :
Frasca Gennaro

Thesis advisor:
Di Nardo Francesco, PhD

Thesis supervisors:
Morbidoni Christian, PhD
Fioretti Sandro, PhD

Contents

Abstract	1
1 Introduction	2
2 Skeletal muscle	5
2.1 Skeletal muscle structure	5
2.2 Neuromuscular control	7
2.3 Excitation-contraction coupling	8
2.3.1 The action potential	8
2.3.2 The sliding filament theory	9
2.4 The origin of the EMG signal	10
3 Gait analysis	11
3.1 The Gait Cycle	11
3.2 Basographic sensors	14
3.3 EMG instrumentation	14
3.3.1 EMG signal	16
3.3.2 sEMG during gait analysis	17
4 Gait-based hemiplegia classification	19
4.1 Cerebral palsy	19
4.2 Hemiplegia and Winters classification	21
4.3 Cerebral palsy forms classification by machine learning: a state of the art	21

5	Neural Networks	25
5.1	Artificial neurons	25
5.2	Multi-layer Perceptron	27
5.3	Dataset partitioning	29
6	Classifier implementation	31
6.1	Dataset	31
6.1.1	Data acquisition	31
6.1.2	Data processing	33
6.1.3	Data windowing	34
6.1.4	K-fold validation	35
6.2	Neural network algorithm	36
6.3	Performance evaluation	40
7	Results	43
7.1	Multiclass classifier	43
7.2	Two-stage binary classifier	46
8	Discussion	55
	List of Figures	63
	List of Tables	65
	Bibliography	66

Abstract

Gait pattern recognition has a fundamental role in the process of diagnosis and clinical decision-making. Recent approaches have addressed this issue by applying machine learning techniques to achieve a rapid, reliable and objective interpretation of the data collected during gait analysis session. In the study have been proposed two machine learning approaches for surface electromyographic (sEMG)-based classification of the typical walking pattern of able-bodied children, Winters' type I and Winters' type II hemiplegic children. sEMG signals were taken from retrospective studies performed at Laboratory of Gait Analysis, Mocalieri (TO). The database consists of gait data related to 30 hemiplegic children (13 Winters' type I and 17 Winters' type II children) and 30 able-bodied children. The first approach proposed showed an average classification accuracy of 98.6% for learned subjects and 79.2% for unlearned ones in the discrimination of the three neuromuscular states considered. The second one showed an average classification accuracy of 93.8% and 88.9% in the discrimination between control and hemiplegic children in unlearned and learned conditions, respectively. In the recognition of the two forms of hemiplegia it reported an average classification accuracy of 94.4% for learned subjects and 75.0% for unlearned ones. Finally, for what concerns the discrimination between healthy, Winters type I and Winters type II subjects, it achieved and an average classification accuracy of 90.5% for learned subjects and 79.2% for unlearned ones. These promising performances highlight the reliability of machine learning predictions in the discrimination of different neuromuscular conditions. Moreover, these findings prove a correlation between the Winters' classification, based on the observation of kinematic data, and muscle recruitment during walking.

Chapter 1

Introduction

In recent years there has been an important increase in the application of machine learning methods in human movement biomechanics [1,2,3,29]. The significant rise of this field is addressed at the development and validation of automatic analysis systems of data derived from the gait analysis with the aim of identifying the presence of pathologies or discriminating different gait patterns. The exponential growth observed in this sector is associated with the spread and optimization of machine learning algorithms, the increase in data that wearable sensors can provide and the rapid improvement in computational power. Among the different pathologies whose main symptomatology is the presence of variations with respect to the standard gait pattern, particular interest has been given to hemiplegia [2].

Hemiplegia is a neurological disorder that is often detected in children with cerebral palsy. This pathology can cause impaired selective motor control, weakness and spasticity. While one side is affected by the disorder, the contralateral side appears to maintain complete functionality. The treatment of hemiplegia, like other forms of cerebral palsy, involves a multidisciplinary effort made by teams composed of different specialists to identify the clinical solutions appropriate for each patient. The therapeutic program is optimized by the correct recognition of the particular form of hemiplegia to which each patient belongs [4,5,6]. Therefore, a good classification system for the identification of groups that have comparable clinical presentations would have the

crucial effect of facilitating the diagnosis process, of defining in a univocal way the clinical picture of the subject, and of choosing the proper treatment protocol to ensure the best solutions for the child's development. Due to the variety of movement disorders associated with hemiplegia, more classification criteria have been proposed. A well-known classification was proposed by Winters and his colleagues [6] for the analysis of the most frequent walking patterns in children and young adult hemiplegic patients. This classification is based on the definition of four different classes, based on the progressive distal-proximal involvement of the paretic leg, defined through the observation of kinematic data in the sagittal plane. Winters' type I was defined by the presence of drop foot in swing phase, Winters' type II by the presence of equinism all through the gait cycle, associated with a possible knee hyperextension during stance, Winters' type III also presented a reduced knee flexion in swing phase and group IV, in addition, had a reduced hip displacement. In the literature, several studies have investigated the correlation between the Winters' classification and the data related to EMG signals acquired during a gait analysis session.

The purpose of the present work is the development of an automatic classification system for hemiplegia, according to Winters' classification, using machine learning interpretation of EMG signals during walking. In this way, this study aims to pursue two objectives: the validation of machine learning systems to automatically identify the typical gait pattern of hemiplegic subjects and the discrimination of two of the four forms of hemiplegia defined by Winters. Moreover, the proposed approach attempts to adapt a classification system defined on kinematic data to the muscular activity, in order to evaluate the correlation and the possibility of using, as index to discriminate healthy from pathological subjects, information obtained through the analysis of muscular recruitment. This information could support clinical decision, or otherwise could be used instead of kinematic data, if those are not available.

The manuscript is organized as follows. In the first chapters the generalities related to skeletal muscle and gait analysis are described. In the third chapter are reported the details of the Winters' classification and a review

of the principal works related to the classification of cerebral palsy through machine learning approaches. Chapter 5 and chapter 6 illustrate the usage of neural networks and the details of the proposed implementation. Finally the results are reported in the chapter 7, and discussed in the chapter 8 that concludes the document.

Chapter 2

Skeletal muscle

Muscles are made up of highly specialized cells for the conversion of chemical into mechanical energy. The three basic types of muscles are *cardiac muscle*, responsible for the mechanical activity of the heart, *smooth muscle*, under involuntary control, and *skeletal muscle*, responsible for posture, locomotion, reflexes and voluntary movement. Skeletal muscle is constituted by a set of parallel contractile units, the *muscle fibers*. Functionally, it is organized in *motor unit*, the smallest unit that can be activated by a volitional effort, in which all constituent muscle fibers are activated synchronously [7].

2.1 Skeletal muscle structure

As the name suggests, skeletal muscles usually act on the skeleton, to which they are attached by way of *tendons*, cords of elastic connective tissue that transmit force from muscle to the bone. The whole muscle is surrounded by connective tissue, that constitutes the *epimysium*, continuous with the tendons. Connective tissue outside the epimysium, *fascia*, surrounds and separates the different muscles. Moreover, connective tissue sheaths extend into the muscle, around groups of muscle fibers constituting numerous bundles called *fasciculus*. Finally, the single muscle fibers within the fascicle are encased in a thin sheath of connective tissue, called *endomysium* [7].

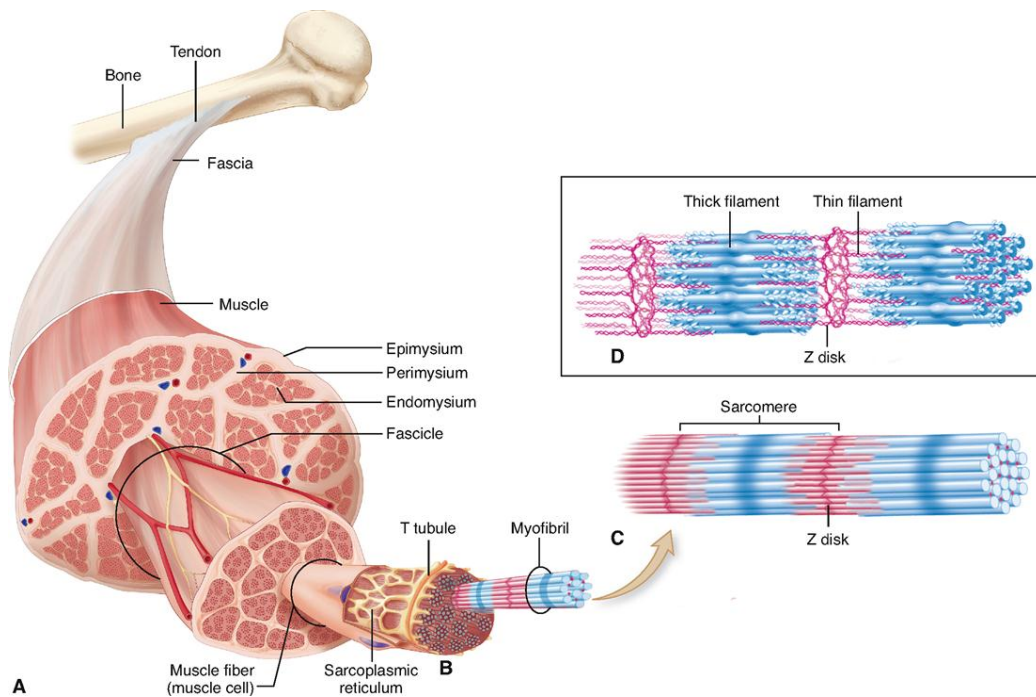


Figure 2.1: (A) Anatomy of the skeletal muscle (B) Muscle fiber (C) Myofibril (D) Sarcomere.

Each muscle fiber contains numerous *myofibrils*, made up of *sarcomeres*, which constitute the basic contractile unit in skeletal muscle. From a structural point of view, the term sarcomere designs the region comprises between two dark lines, called *Z lines*. This characteristic pattern of alternating dark and light region yield to skeletal muscle its characteristic striated appearance. The myofibrils' basic building elements are two contractile protein, *myosin* and *actin*, which constitute a thick and thin filament, respectively. The myofibrils are surrounded by an intracellular network, the *sarcoplasmic reticulum*, that regulates the intracellular Ca^{2+} concentration, fundamental for the contraction initiation and termination. This structure is closely associated with invaginations called *transverse tubule* (T tubule), that penetrate deep within the cell [7].

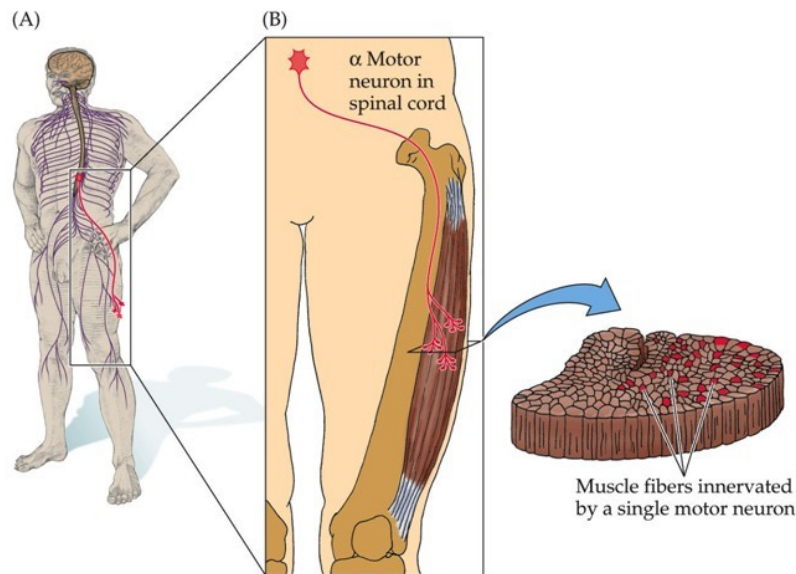


Figure 2.2: (A) α -motor neurons originate from the spinal cord (B) Each α -motor neurons innervate a group of muscle fibers, constituting a motor neuron [8].

2.2 Neuromuscular control

Skeletal muscle contraction is controlled by motor neurons, whose cell bodies are located in the ventral horn of the spinal cord. These neurons, also called α -motor neurons constitute the final pathway for transmitting information to the skeletal muscle and initiate its contraction. Each single motor neuron sends axon via the ventral root of the spinal cord and innervates a set of muscle fibers, constituting a motor unit. A single motor unit and its associated muscle fibers together constitute the smallest unit that can be activated to produce movement (Figure 2.2). The number of fiber innervated by a single motor neuron, and so the set of cells that are going to contract together when the motor neuron fires, can considerably change in different muscles. When the muscle fibers are stimulated through electrical signal delivered from the neural termination, all the constituent muscle fibers are activated synchronously and the whole motor unit is going to contract. The sequence of events that links the cell electrical stimulation action potential to the muscular contraction is referred to as *excitation – contraction coupling*. The connection between a motor neuron and a muscle cell is referred to as a

neuromuscular junction. Here, acetylcholine released from the motor neuron triggers a change in ion permeability in the innervated muscle fiber that initiates an action potential. The action potential then travels along the length of the muscle fiber and down the T tubules. There, the action potential propagation results in an increase in the intracellular Ca^{2+} concentration that initiates muscle contraction. [8]

2.3 Excitation-contraction coupling

2.3.1 The action potential

The excitable cells, the main constituent of nervous, muscular and glandular tissue, are able to generate and propagate an electrical stimulus as a result of electrochemical changes that occur at the cellular level when adequately stimulated. The action potential generation and conduction is related to the exchange of ionic species between intracellular and extracellular space that occurs through the cellular membrane. These two environments are characterized by a different ionic concentration to the extent that it exists a transmembrane electrical gradient. In the cell resting state, the electrical gradients are balanced with chemical gradients, and to keep this equilibrium, the cells are involved in different spontaneous mechanisms which contributes in the maintenance of a constant ionic concentration difference. Excitable cells are characterized, in addition to the state of rest, by a state of activity. The passage from the state of rest to that of excitement is determined by a transient electrical stimulus, which implies the opening and closing of voltage-gated ionic channels and, as a consequence, variation in the transmembrane potential. This implies the generation of an action potential, whose typical pattern, depicted in Figure 2.3, can be divided into three main phases: The action potential initiation (*phase 1*), is a rapid depolarization, the results of an electrical stimulus which, if higher than a threshold value, causes a sudden and dramatic increase in sodium permeability. The consequent Na^+ inflow implies a rapid depolarization during which membrane potential reaches its highest positive value. Subsequently, during the *phase 2* the reduction in

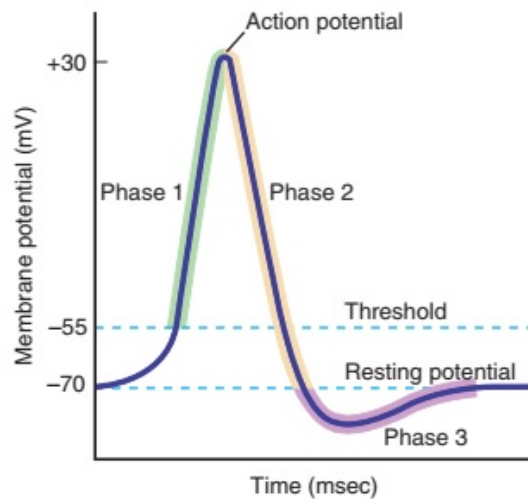


Figure 2.3: Action potential waveform.

sodium permeability, followed by an increase in K^+ permeability and then K^+ outflow, results in a reduction in the membrane potential, that comes back to negative values. In a transitory short period (*phase 3*), the trans-membrane potential reaches a more negative potential than at rest, before then restoring its resting state value.

2.3.2 The sliding filament theory

In the state of relaxation, in the muscle cell there is a very low cytosolic calcium concentration. When an action potential travels down the T tubules, it induces a conformational change in voltage-dependent receptors present over the T-tubule membrane, causing the calcium channels to open. Ca^{2+} then moves out of the sarcoplasmic reticulum and into the cytosol. The consequent rise in Ca^{2+} concentration initiates muscle contraction by promoting the myosin and actin interaction, resulting in the *cross – bridge* formation, a link between the myosin heads of the thick filament and the actin. After the cross-bridge formation, a ratchet action of the myosin head pulls the actin filament toward the centre of the sarcomere, causing sarcomere shortening and so muscle contraction. The sarcoplasmic reticulum and T tubule network arrangement permits nearly simultaneous delivery of calcium to all

sarcomeres of a muscle fiber, so the sarcomeres contract in unison, as does the entire muscle fiber. The contraction is prolonged until the muscle cell is stimulated. When it no longer receives input from its motor neuron, and action potentials no longer occur in the sarcolemma and so the muscle fibers are going to relax [7].

2.4 The origin of the EMG signal

Muscle contraction is the consequence of the propagation of electrical stimuli along the muscle fibers. The acquisition, through suitable electrodes, of those signals can provide useful information related to the muscular activity. This recording, called EMG signal, is the algebraic summation of the motor unit action potentials, the spatial and temporal summation of the individual muscle action potentials for all the fibers of a single motor unit, within the pick-up area of the electrode being used. EMG acquisition is performed using an instrument called electromyograph to produce a record called an electromyogram. An electromyograph detects the electric potential generated by muscle fibers when these fibers are electrically or neurologically activated, producing a contraction. EMG is a diagnostic tool used for the recording and evaluation of the functionality of muscles and nerves that controls them. EMG signal is commonly used in many fields such as in motor and postural control, biomechanics, kinesiology, rehabilitation and physical therapy, and it is used for the detection of medical abnormalities, activation level, or recruitment order, or to analyse the biomechanics of human movement. Its interpretation can assist doctors in the diagnosis of some neuromuscular disorders such as neuropathy and myopathy.

Chapter 3

Gait analysis

“Gait analysis is the systematic measurement, description and assessment of those quantities thought to characterize human locomotion”. [9] It is a well-established tool to analyse the human walking by determining some spatial and temporal parameters that characterize the human locomotion useful to make clinical evaluation. Gait analysis takes places in dedicated laboratory equipped with optoelectronic system, reflecting markers attached to patients in specific anatomical landmarks to acquire tridimensional motion of human segment, force plates, to measure the intensity and direction of the reaction force to body weight, electromyography, skin-electrode for the acquisition of the electrical signal generated by the muscle contraction and video system, made up of cameras to record the patient movement during a walking trial.

3.1 The Gait Cycle

Locomotion is the repetition of sequences of steps to allow body advancement in the desired direction. To analyse and quantify the human locomotion it is necessary to isolate the shortest and repeatable task during gait, the gait cycle, assumed by convention as the time between one foot strike to the subsequent foot strike of the same foot. In a first analysis, the entire gait cycle could be divided into two main periods, *stance phase* and *swing phase*. During stance phase, the foot is on the ground, whereas in swing phase that

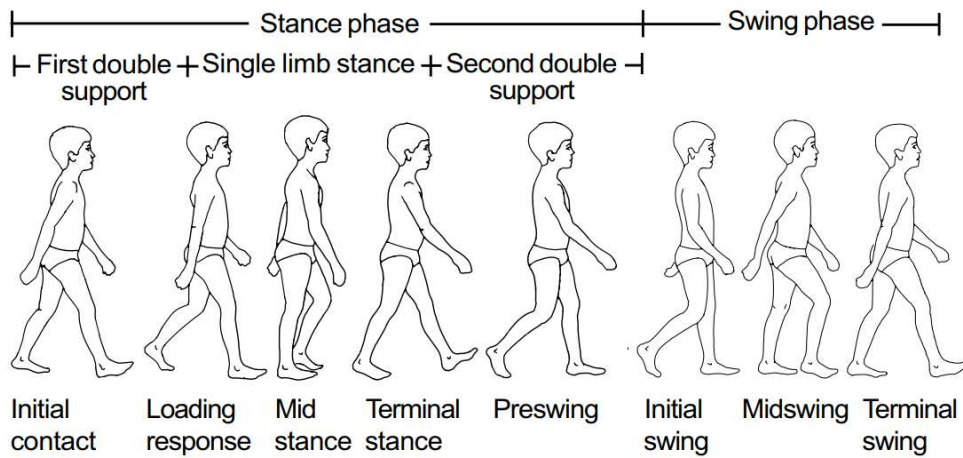


Figure 3.1: The normal gait cycle of an 8-yearold boy [3].

same foot is no longer in contact with the ground and the leg is swinging through in preparation for the next foot strike. In a more detailed analysis, for the overall duration of the gait cycle, different characteristic phases have been individualized. The traditional nomenclature that describes the different phases of the gait of a normal subject was developed by Perry and her associates at Rancho Los Amigos Hospital in California (Cochran, 1982). They subdivided the entire gait cycle in eight phases and developed a detailed self-descriptive nomenclature based on the functional interpretation of the movement of the foot. [9]

1. *Initial contact:* This phase, by convention the first event of the gait cycle, starts when the strike-foot touches the ground for the first time, impact that in not pathological condition occurs at heel level. After this initial impact, the foot totally comes into contact with the ground through ankle plantarflexion, which is accompanied by the ankle dor-siflexor contraction to break down the fall of the foot.
2. *Loading response (0-10%):* The loading response phase is the phase during which the weight is transferred from the training leg to the leading leg, and the lower limb muscles act to guarantee the limb stability, absorb the shock due to ground impact and preserve the body progression.

3. *Mid stance (10-30%)*: The mid-stance is the first sub-phase of the single support phase. It begins with the *opposite toe-off*, when the leading leg has completed the weight realize and leaves the ground. During this phase, the tibia rotates forward over the supporting foot to guarantee the forward progression of gait.
4. *Terminal stance (30-50%)*: The second stage of single support is the terminal stance. The event that initiates this phase is the heel rise and the contralateral foot swings in the direction of progression. As well as the heel loses the contact with the ground, there is a rapid plantarflexors' activity to allow the rise and the push off of the foot, so that the swing phase can start. During this phase it occurs the second double support phase with the *opposite initial contact*, the first impact of the contralateral foot with the ground, the instant that defines the swing phase begins and the contralateral swing phase end.
5. *Preswing (50-60%)*: Preswing is the transition phase between stance and swing phase, during which takes places the weight acceptance by the contralateral leg and the foot is pushed and lifted off the ground.
6. *Initial swing (60-75%)*: The initial swing phase starts with the *toe-off*., when the foot leaves the ground and the limb begin to move in the direction of progression. This instant defines the beginning of the right swing phase and left single support phase. During this first swing phase the hip is flexed, to guarantee the swing of the leg, while the knee flexes and the ankle is dorsiflexing to avoid that the swinging foot could touch the ground, reducing so the fall risk.
7. *Midswing phase (75-85%)*: During midswing, the thigh reaches its peak advancement: the foot passes directly beneath the body, coincidental with midstance for the other foot. This phase starts the when the two feet are adjacent and ends with the *tibia vertical*, the last event of the gait cycle that precedes the next initial contact.

8. *Terminal swing (85-100%)*: Terminal swing is the last gait cycle phase. During this phase, the thigh advancement is completed and the muscles slow the swing limb movement and stabilize the joints in preparation for the forthcoming next heel strike.

3.2 Basographic sensors

The basographic sensors consist of a rectangular membrane switch used to collect the temporal data relative to the foot-floor contact phase. To describe the support of the foot on the ground, three sensors are placed under the sole, in three independent zones, under the heel, the first and the fifth metatarsal head, in a way that it is possible to distinguish eight conditions of support. This sensor configuration allows to obtain the so-called eight-level baseline, too detailed to be used in the statistical analysis of the walking. The eight-level baseline can, therefore, be reduced to obtain the corresponding four-level baseline (foot contact, heel strike, push off and swing) or the two-level baseline (swing and stance phase). During the walk the basographical signals coming from the two feet are acquired continuously: the basographic cycle of the path is fundamental, as it constitutes the reference to evaluate the progress of all the other signals [10].

3.3 EMG instrumentation

The instrumentation needed for accurate recording of the EMG signal consists of recording electrodes, signal amplification, transmission, and a means for data display and storage. Myoelectric signals can be recorded either by surface electrodes, that has the benefit to be not invasive but provide low quality information on the muscle activation, and needle electrodes, employed for diagnosis of single muscle disorders and not recommended for routine gait analysis due their intensiveness. Surface electrodes are placed on the skin over the muscles of interest. In routine gait analysis are frequently used silver–silver chloride (Ag–AgCl) electrodes, that consist of both

a pickup and a reference electrode. These electrodes require the skin surface to be cleaned and dried to get a good signal and avoid artifacts, moreover a conductive paste gel is used to minimize interface impedance at the recording site. It has been recommended that the Ag–AgCl disks be separated by 1 cm to obtain improved signal quality from the targeted muscle. They should be placed between a motor point and the tendon insertion or between two motor points, and along the longitudinal midline of the muscle. Moreover, a reference electrode is placed as far away as possible and on electrically neutral tissue to provide a common reference to the differential input of the preamplifier in the electrode. The principal advantages of surface electrode is that they are not invasive, easy to apply, reusable, produce repeatable results, and can detect activities of muscle groups. However, their disadvantages are the inability to record activities from deep and specific muscles, being the signal that they provide affected by crosstalk from neighboring muscles. In situations where the monitoring of the activity of specific or deep muscles might be required, the use of needle electrodes results necessary.

Needle electrodes allows to measure myoelectrical activity directly from individual muscles. The electrode positioning is performed by inserting, directly within the muscle of interest, a small needle containing a pair of fine-wire electrodes, one active and one of reference. The needle is then pulled out, leaving the two fine wires positioned within the muscle. Confirmation of the correct placement of the electrodes is then determined by stimulating the muscle electrically through the fine-wire electrodes, through palpation and observation of the target muscle or tendon. The most common type of fine-wire electrode used in dynamic EMG is a nickel–chromium alloy wire (50 μm in diameter) with Teflon insulation. The greatest advantage of fine-wire EMG is muscle selectivity, and the possibility of recording electrical activity from specific muscles or those deeply located. However, despite the goodness of the signal that it provide, this method is invasive, requires skilled placement, and may require multiple insertions [11].

3.3.1 EMG signal

Both surface and fine-wire EMG signals have small amplitudes and do not allow direct interpretation without amplification and appropriate signal processing. A differential amplifier with high common mode rejection is used to eliminate electrical noise seen by both electrodes.

The myoelectric signals recorded by surface and wire electrodes have different spectral characteristics of known bandwidths. sEMG signals have a bandwidth of 10 Hz to 350 Hz, with a mean frequency of 50 Hz, while fine-wire EMG signals have bandwidth of 10 Hz to 1000 Hz, with a mean frequency of 350 Hz. The lower bandwidth observed in sEMG is the result of higher frequency signals attenuation as they travel through the tissues. The quality of the EMG signal is improved by appropriate signal processing, amplification and filtering rejected frequencies, in order to reduce some of the principal kinds of noises and artifacts that affects both sEMG and EMG signals such as

- *Line interference*, due to the power line, transmitted by electrical devices placed near the EMG data acquisition device.
- *Direct current offset*, caused by the difference in the impedance between the skin and the electrodes.
- *Motion artefacts*, related to two factors: to the interface between the detection surface of the electrode and the skin, and to the movement of the cable connecting the electrode to the amplifier. These artefacts has a low frequency content (10 to 20 Hz) and can be reduced by increasing the low frequency cut-off.
- *Electrocardiographic signal*, generated by the heart, that can be picked up with the EMG signal.
- *Crosstalk*, that mainly affects sEMG acquisitions, produced by EMG signals coming from other muscles than the ones being monitored.

3.3.2 sEMG during gait analysis

During the gait cycle, the largest movements occur in the lower limbs, whose activity acquisition, performed through EMG recording, plays an important role in evaluating gait in individuals with neuromuscular disorders. The current state of the art suggests that, during a session of gait analysis, a complete picture of the human walking pattern the activity of five superficial lower limbs might be monitored [5,12,13]. These are three muscles of the thigh, the rectus femoris and the vastus lateralis, located on the frontal part of the thigh and the hamstrings, located posteriorly, and two muscles of the shank, the tibialis anterior and the gastrocnemius, placed anteriorly and posteriorly, respectively (Figure 3.2).

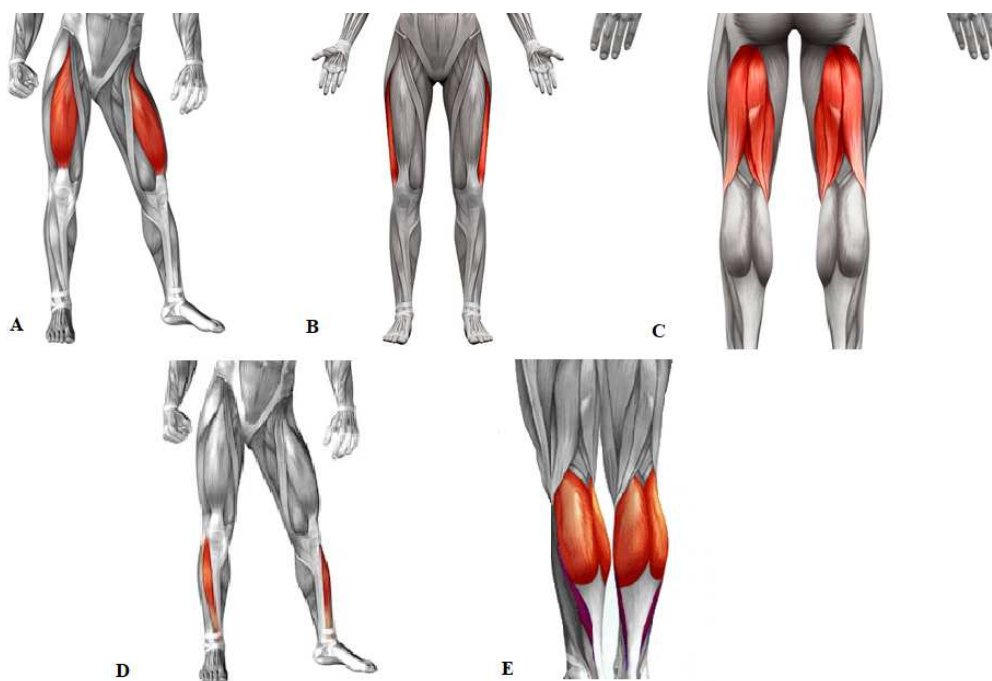


Figure 3.2: Lower limb muscles whose activation is acquired during sEMG session. (A) Rectus femoris (B) Vastus lateralis (C) Hamstrings (D) Tibialis anterior (E) Gastrocnemius.

The rectus femoris is situated in the middle of the anterior side of the thigh and the two insertion are located at the hip and at the base of the patella. This muscle is an hip flexor and a knee extensor. During the gait cycle it

is mainly active during stance phase, to allow the leg progression. Vastus lateralis is one of the three vasti (the other are the vastus medialis and vastus intermedius), powerful muscle that lies in the frontal side of the thigh that together with the rectus femoris constitute the quadriceps. Its insertion point lies between the femur and the patella; indeed, it acts only at the knee level and its activation is related to the knee extension. It has an important activation at the beginning of the gait cycle, indeed like the other muscle that act at knee level is recruited for the load absorption after the heel contact. Its activation, indeed, avoids an excessive knee flexion related and ensure the reduction of the load transmitted to the upper body segments when the foot impact the ground. Moreover, the quadriceps activation is necessary before the toe off, to avoid that the foot touches the ground during the stance phase. As for the hamstrings, it acts also at the end of the gait cycle, in order to prepare the leg for the next step. The hamstrings are a group of muscle made up of the bicep femoris, the semimembranosus and the semitendinosus, that lies in the posterior side of the thigh. These muscles act across the hip and the knee joint and are responsible for the hip extension and knee flexion. They are recruited, during the gait cycle, especially at its beginning and at its end: during the loading response, in order to stabilize the leg after ankle impact, and during terminal swing phase, in order to decelerate the swing of the limb and so to stabilize the leg for the forthcoming initial contact. The gastrocnemius muscle is a superficial muscle made up of two heads, the medial and the lateral one. It runs from its two heads just above the knee to the heel, indeed it is a two joint muscle: it is both a knee flexor and an ankle plantar flexor. It and the soleus are mainly recruited before the toe off, to generate power to rise the foot and for the push off phase. The tibialis anterior is situated on the lateral side of the tibia and act to dorsiflex and invert the foot. Ankle dorsiflexion many occurs after the initial contact, in order to break the tibia forward movement and the foot impact with the ground.

Chapter 4

Gait-based hemiplegia classification

The traditional gait assessment methods are based on clinical observation. However, it has the limit of being qualitative, whereby decisions were based on the subjective judgment of clinicians, and also on time-consuming processes. Recently, different machine learning approaches are used to automatically interpret kinetic and kinematic gait data and provide an objective decision. These gait classification tools would assist clinicians in the diagnosis of pathological gait patterns and in the validation of therapeutic treatment. In recent years, machine learning approaches have been increasingly used in gait disorders research. Cerebral palsy is one among the most commonly studied conditions in which the interest in developing gait pattern classification through machine learning interpretation of gait data is on the rise.

4.1 Cerebral palsy

Cerebral palsy is a group of permanent movement disorders characterized by paralysis resulting from abnormal development or damage to the brain either before birth or during the first years of life. Cerebral palsy does not necessarily include intellectual disability, indeed many children affected with cerebral palsy are mentally competent. The causes of cerebral palsy involve a

malfunctioning of the complex motor neuronal circuits. The cerebral damage causing cerebral palsy primarily affects the neurons and connections of the cerebral cortex, either of one cerebral hemisphere (contralateral to paralysis) or of both hemispheres, compromising the muscular function one side of both side of the body, respectively. There are not cures for cerebral palsy, and the clinical approach to this pathology includes physical therapy and medication to achieve muscle relaxation or functional surgery. The choice of the proper treatment depends on the identification of the cerebral palsy forms. One of the most striking features of cerebral palsy is the large variability of its clinical presentation. The symptoms vary among people, and the typical one are poor coordination, weakness and altered selective muscle control. The diversity of gait deviations observed in children with cerebral palsy has led to repeated efforts to develop gait classification systems to assist in diagnosis, clinical decision-making and communication. Gait classification is a system through which it is possible to allocate a gait pattern into groups that can be differentiated from one another based on a set of defined variables. The gait classification is achieved by the evaluation of gait indexes, assessment scores and scales, that provides an overall index to quantify deviations from normal gait. Gait classification may enable clinicians to differentiate gait patterns into clinically significant categories that assist with clinical decision-making. Moreover, the neuromuscular condition definition can also provide clinicians and researchers a common language that quickly refers to a clinical picture of an individual's gait impairments. There is a broad variety of movement disorders that belongs to the group of cerebral palsy, so multiple classification criteria have been proposed. According to the somatic location of prevailing neurologic symptoms, cerebral palsy can be distinguished as *tetraplegia*, in all four limbs are affected, *diplegia*, if the lower limbs are more compromised than the upper one, *hemiplegia*, if the muscle stiffness and paralysis affect only on one side of the body, or *monoplegia* if only a single limb is affected.

4.2 Hemiplegia and Winters classification

An acknowledged classification of the most frequent walking patterns in hemiplegic children and young adult was proposed by Winters et al. [6] who analysing the sagittal joint kinematics identified four different classes, with a progressive distal-proximal involvement of the paretic leg. Several studies tried, using such classification as a reference, to quantify the myoelectric activity of the lower limb muscles to identify the main distinctive gait pattern that distinguishes the hemiplegic children to able-bodied children and those present that characterize the different classes. The main characteristic of Winters type I subject is foot drop during swing phase resulting in a lack of first rocket at initial contact. This impairment is associated with the weakness or hypo-activation of the tibialis anterior during terminal swing and hyper-activation of ankle plantarflexors (gastrocnemius and soleus muscle) throughout swing. A more severe condition was detected in Winters' type II subjects: persistence of equinism all through gait cycle, associated with a possible knee hyperextension in stance phase. This is related to a curtailed activity of ankle dorsiflexors, too, and overactivity of the gastrocnemius and soleus muscle. Winters' type III also presented an excessive ankle plantar flexion over the entire gait cycle and reduced knee flexion in swing and Winters' type IV presents all the above-mentioned deviation, plus a reduced hip range of motion over the gait cycle [4,5,14,15,16].

4.3 Cerebral palsy forms classification by machine learning: a state of the art

The use of data science methods in human movement biomechanics studies that focus on neuromuscular and musculoskeletal pathologies has increased exponentially in recent years, as reported in the graph in Figure 4.1. From monitoring patients after stroke events through wearable devices, to predicting interventions outcomes cerebral palsy children, the range of applications in which artificial intelligence can improve rehabilitation research will con-

tinue to expand, particularly as wearable devices generate large amounts of data. The most common area of application of machine learning in movement analysis is the definition of automatic classification systems, able to automatically discriminate pathological kinematics from normal kinematics [1,2,3]. Automatic recognition and classification of the neuromuscular gait patterns could offer many potential benefits, especially for the diagnosis and treatment planning. Besides, to use the classifier output in phase of decision-making might be also useful to assess the goodness of therapy or rehabilitation intervention, indicating if the chosen treatment has improved the overall gait function. Studies propose as a future application of artificial intelligence learning in the field of biomechanics to employ the prediction of machine learning algorithm to drive robotic devices designed to assist the pathological subject in a program of training and rehabilitation. These systems are designed to acquire EMG or joint angular data, calculate the subject's intention and the gait deviation uses these data to drive a robotic gait expected to assist walking and improving walking functions in pathological subjects [17,18].

The cerebral palsy is one of the more largely neuromuscular dysfunctioning studied to achieve automatic gait pattern classification. Several researches have proposed different machine learning approaches for the classification

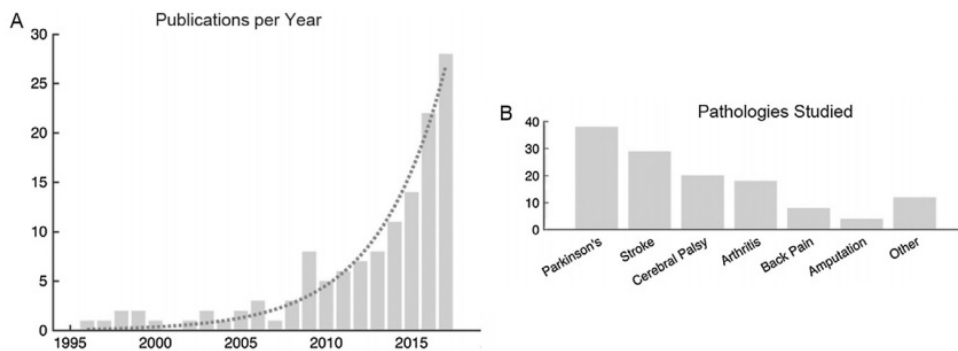


Figure 4.1: Summary of the main trends in machine learning in human walking. (A) Rise of the research in the usage of machine learning methods in movement biomechanics. (B) The most commonly studied neuromuscular conditions [1].

of cerebral palsy forms through the interpretation of electromyographic, kinetic and kinematic data. Among them, it is possible to identify studies that proposed supervised classification algorithms, with the neuromuscular state identified through specialists' interpretation of the children's gait pattern, and unsupervised approaches, that try to identify common gait pattern through cluster analysis. For what concerns the supervised approaches, studies proposed a classification of the different forms of cerebral palsy using as input features lower joint angles, derived via standard inverse kinematics using data collected by means of an optoelectronic system, such as the studies of Ferrari et al. [19], Salazar et al. [20] and Zhang et al. [21].

Ferrari et al. tried to solve a multi-class classification problem discriminating, through a set of angular parameters, four different forms of diplegia [19]. Similarly, the feature extraction methods using kinematics data were employed by Salazar et al. for the individualization of four different types of spastic hemiplegia [20]. The same approach of training a classifier fed with features extracted from angular time sequence was presented in the study of Zhang et al, who tried to identify four different gait patterns on a database of gait trials from cerebral palsy children with spastic diplegia [21]. Different kind of data was used in the classification systems proposed in the study of Hansi et al. [22], that implemented cerebral palsy forms identification using sEMG data. The machine learning system implemented in their study was a multi-class classifier, designed to discriminate three classes of neuromuscular states: healthy, neuropathic a myopathic condition. They set a 2-stage binary classifier approach, with input a set of time-domain features extracted from the sEMG signal. The first stage involved discrimination of healthy versus pathological subjects, while the second is intended to classify the data coming from the first stage into the myopathic and neuropathic classes. sEMG data was also used by Lahmiri et al. who in their study proposed a classification system to discriminate healthily versus neuropathic. [23] Moreover, a further classifier for the automatic detection of cerebral palsy children was proposed by Kamtuzzman et al., who design a binary classifier, intended to distinguish healthy versus pathological subject, that does not use kinematic or kinetic data as input features, as the above-mentioned studies, but stride

length and cadence, two temporal-spatial gait analysis derived parameters [24]. In addition to joint angle and sEMG data, different input features have been employed by Alaqtash et al. Their group of researches developed an automatic gait classification tool, to classify healthy, cerebral palsy and multiple sclerosis patients, using tridimensional ground reaction force data [25].

Chapter 5

Neural Networks

Around the middle of the last century a group of researchers in the field of computation developed a machine capable of simulating every aspect of learning and human intelligence. Since then the history of artificial intelligence had a rather sudden development. Machine learning, as the name suggest, is based on the idea that the machine might be able to elaborate a huge amount of data with the aim of define a set of rules to predict outcomes for unseen data. The process is performed by neural networks, composed of layers of neurons that act as on/off switches to create neural pathways similar to those built by the human brain. Deep learning is a machine learning tool designed to automatically extract features from raw data that is having a significant spread in recent times. Nowadays deep learning has applications in almost all areas of technology and information. Among them, medical diagnostics is a sector in which the interest in the development of neural networks is greater, given that in many cases algorithms are already used to support specialists in decision making.

5.1 Artificial neurons

The key elements of any neural network architecture are the artificial neurons. As the name suggest, the basic idea on which they are founded is the analogy with the activity of the biological neurons, the set of interconnected cells that

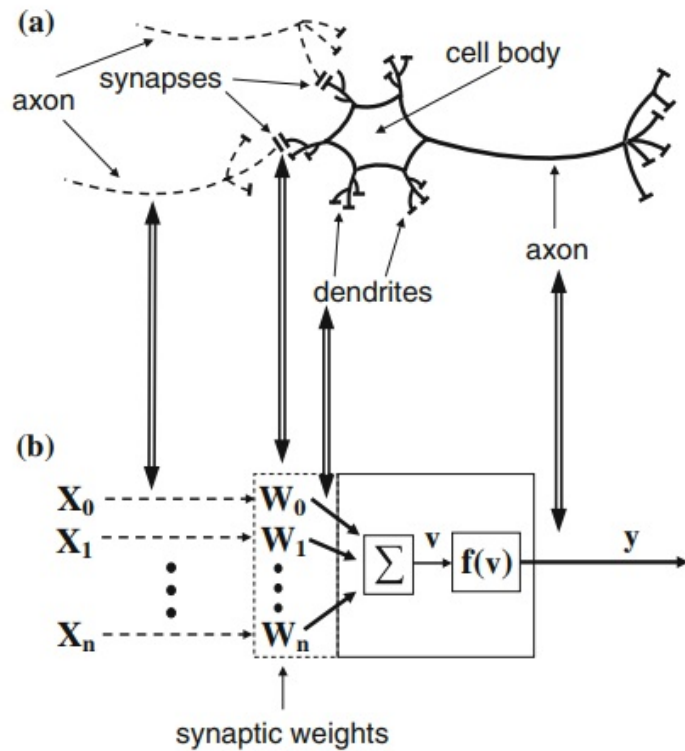


Figure 5.1: (a) Analogy between the biological neuron (b) and an artificial neuron scheme [26].

constitute the nervous system, involved in the transmission and elaboration of chemical and electrical signals. The first idea of the artificial neurons was reported by Warren McCullock and Walter Pitts in the paper "*A Logical Calculus of the Ideas Immanent in Nervous Activity*", [27], in which it is described the working principle of the first artificial neuron, implemented as a logical gate with a binary output, delivered to other neurons as the biological neurons send information through the axon, which value depends on the inputs value of the signal that are collected by neuron's body from a set of dendrites.

The analogy between artificial and biological neurons is presented in Figure 5.1. The impulses collected by the dendrites and delivered to the neuron cell body are modelled as a set of inputs X_0, X_1, \dots, X_n , and the action potential

delivered to the axon represents the output of the present model. This output is computed by multiplying the input signal by a set of weight W_0, W_1, \dots, W_n , that plays the role of the transmission properties of the synapses. The results of the overall summation of the multiplication of each input signals by each weight gives a signal v , that with the analogy of the biological neurons, corresponds to the excitatory postsynaptic potential (equation 5.1)

$$v = \sum_{i=0}^n W_i X_i \quad (5.1)$$

where X_i is the input matrix and W_i is the weight matrix.

The signal is transferred to the axon only if it reaches the proper threshold. This process is modelled with an activation function $f(\cdot)$, that generates, for each signal v , an output signal y , that is the final output of the considered model (equation 5.2).

$$y = f(v) \quad (5.2)$$

Several activation functions might be considered, linear and not linear, with different properties, and its choice is often guided by the outcome obtained using them [26].

5.2 Multi-layer Perceptron

During the years, several neural network architectures have been developed. One of the most common among these is the multi-layer perception (MLP). It is made up of a set of different neurons organized in multiple main layers, connected between them: one *input layer*, one or in more complex networks several *hidden layers*, where the data have been elaborated, and one *output layer*. The units present in the hidden layer are connected to the input layer and those present in the output layer are connected to those present in the hidden layer. Successive layers are connected by weights, and the MLP updates that weights iteratively to map the input vectors to a set of corresponding output vectors. The model output is computed throughout the *feedforward* process, which term refers to the fact that each layer serves as

an input for the next layer. Starting from the input layer, the input data are deeper and deeper propagated through the different hidden layer, in order to generate the model output. At the end of the feedforward process, the model error is computed as the difference between the real output and the model prediction, and the machine learning algorithm works to minimize this error in a second stage, the *backpropagation* process. During this process, it is computed the loss function, that estimates the difference between the model prediction and the true label and is the function that the algorithm has to minimize. The estimated error is then propagated back, and it is computed the loss function derivative with respect to each weight, the model is updated and a further output is generated, using a *gradient descending* technique. The goal of feedforward and backforward processes is the model parameters definition, the configuration of variables intrinsic to the model, which are estimated based on training data in order to minimize the error loss function. The model parameters updating is repeated several times, and each trial is called *epoch*. The amount that the weights are updated during training for each epoch is referred to as *learning rate*, one of the several configurable parameters used in training that has a small positive, often lower than 1.0. The execution of the gradient descent algorithm can be very expensive from the computational point of view in the presence of a very large dataset. In that situation, it is possible to apply a type of optimization called *mini-batch gradient descent*, which involves dividing the entire dataset into a small subset of training data called mini-batches. In batch gradient descent, the training occurs on the entire batch of training samples all at the same time, and the model parameters are updated averaging the gradients of all the training examples and then use that mean gradient. On the other side, in the mini-batch gradient descent the neural network processes at the same time a fixed number of training examples, less than the actual dataset, rather than working on the entire training set. The advantage over the descent of the batch gradient is the fact that convergence is achieved faster through mini-batches, due to the more frequent updating of model weights. Learning rate and batch size are only two examples of model hyperparameters, parameters that can be used to control the complexity and learning rate of a

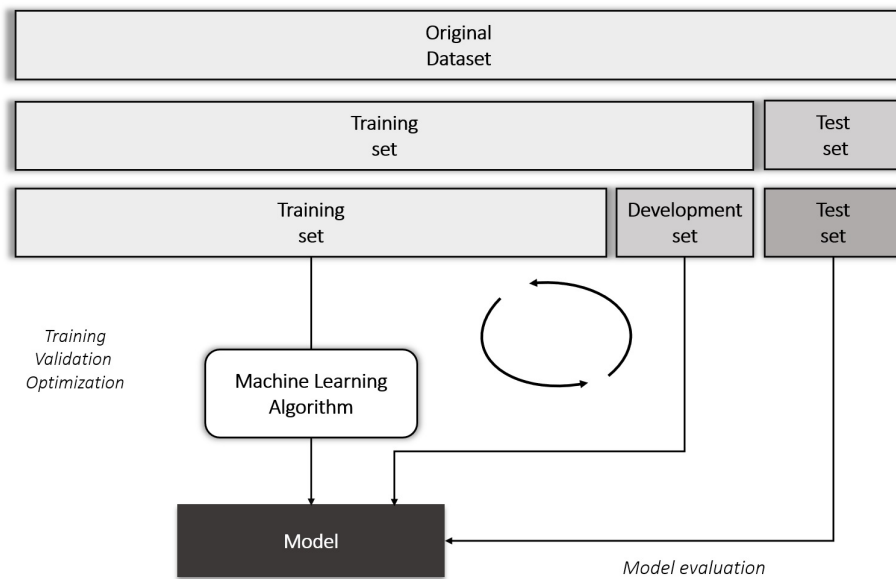


Figure 5.2: Partitioning of dataset into training and test set.

model. These parameters are not learned by the model but are set before the training phase or adjusted through a process called *hyper-parameter tuning*. This is an iterative process during which many different combinations of the different hyperparameters were tested, in order to find the set that gives better performances.

5.3 Dataset partitioning

The artificial neural network algorithms consist of a training and a test phase. The algorithm train is the process during which the network receive as input all the data present in a *training set*, and try, adjusting its internal parameter, to define a set of rules applicable for new instances. The model performances, then, have to be tested on a new set of data, different from the training set, the *testset*. For this reason it is necessary to run this separation. This workflow is depicted, in an exemplificative way in Figure 5.2. A common choice is to include the 90% of the total dataset into the *trainig set*, while the remaining 10% is used to evaluate the actual model performances.

Furthermore, the training test is split into *trainig set* and *development set*. The training set is used for the model train, while the development set is used during train for evaluate the model with different configuration, and is used for its validation. Indeed, it is called development set because it is used during the training and not to evaluate the final model performances, as the test set. When the model performances have been validated with the development set, the test set is used to give an unbiased estimation of the algorithm performances. One drawback of this implementation is that the model is sensible to the way on which the original dataset is split into training and test set. For this reason, to make the evaluation more reliable, the k-fold cross-validation might be considered. In this case, the above described workflow is repeated k times over k different sub-sets on training data. This process is based on the random subdivision of the whole dataset into k parts without reinsertion: k-1 parts have been used for the model train, while only one is employed for the testing. This procedure is repeated k times in order to get k different models and performances evaluations. The way on which the dataset is split is made to guarantee that wach data will be used for the training and for the testing at least once, to get an estimation of the model performances with lower variance (Figure 5.3).



Figure 5.3: A large variability of performances was expected, depending on the dataset partition into train set and test set. To make the model performances more reliable, a k-fold validation might be implemented.

Chapter 6

Classifier implementation

The aim of the present work is the implementation, on the Pytorch deep learning framework, of a machine learning algorithm, designed to handle a multi-class classification problem. The inputs were the sEMG data related to a group of healthy subjects, W1-hemiplegic subjects and W2-hemiplegic subjects, and the proposed algorithms were designed to provide information about the subjects' neuromuscular state through neural network interpretation of sEMG data. Two different solutions to handle this problem have been implemented and tested: the first one is a multiclass algorithm, while the second is a binary cascade approach.

6.1 Dataset

6.1.1 Data acquisition

Gait data from control and hemiplegic were taken from retrospective studies performed at Laboratory of Gait Analysis, Ospedale Santa Croce, Moncalieri (TO), Italy. The database included a total of sixty subjects, distributed into the three classes considered as follows: thirty able-bodied children (18 females and 12 males), of age ranging from 6 to 11 years, included in the database as control subjects, thirteen Winters' group I (W1, 6 females and 7 males), of age ranging from 5 to 13 years, and seventeen Winters' group II (W2, 7 females and 10 males), of age ranging from 4 to 10 years.

The different experiments have conducted using sEMG and basographic signals. The basographic signals have been obtained through three foot-switches (10 x 10 mm, thickness 0.5 mm, activation force 3 N) applied under the heel and on the first and the fifth metatarsal heads of each foot for measuring foot-floor contacts. The sEMG were recorded (sampling rate: 2 kHz; resolution: 12 bit) by the multichannel recording system Step32, positioning a single differential sEMG probes on tibialis anterior, gastrocnemius lateralis, vastus lateralis, rectus femoris and hamstring bilaterally following Winters' guidelines. Then, children were requested to walk barefoot back and forth over a 10-m straight walkway at their self-selected speed. From the database, have been selected only the sEMG and basographic signals related to one lower-limb: for the Winters' subjects it was taken the signals related to the pathological side, while for the healthy subjects it have

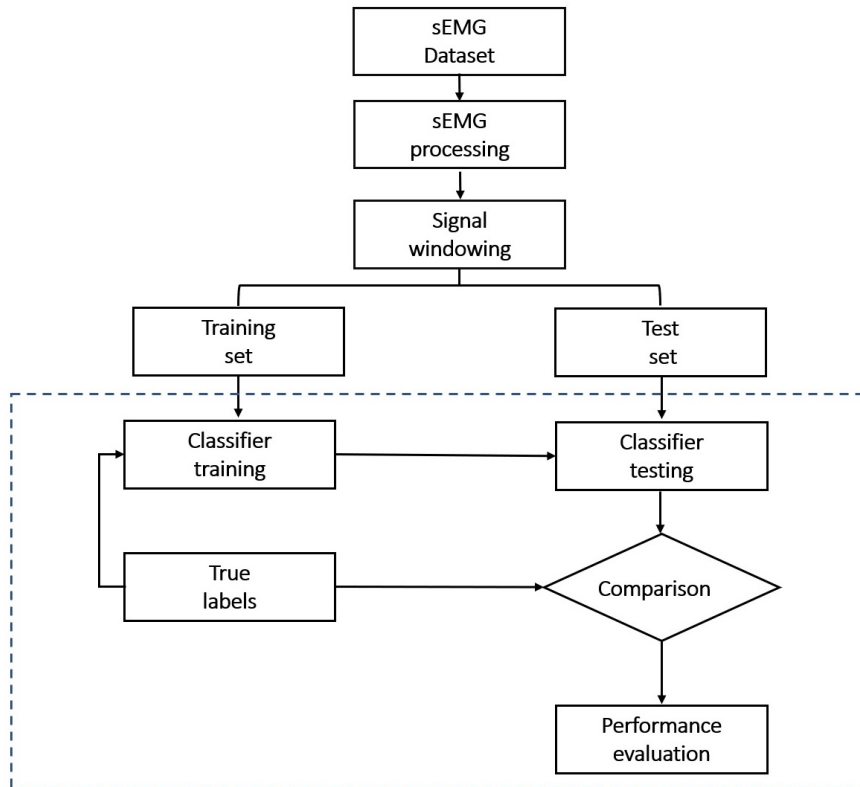


Figure 6.1: Proposed pipeline for the sEMG classifiers.

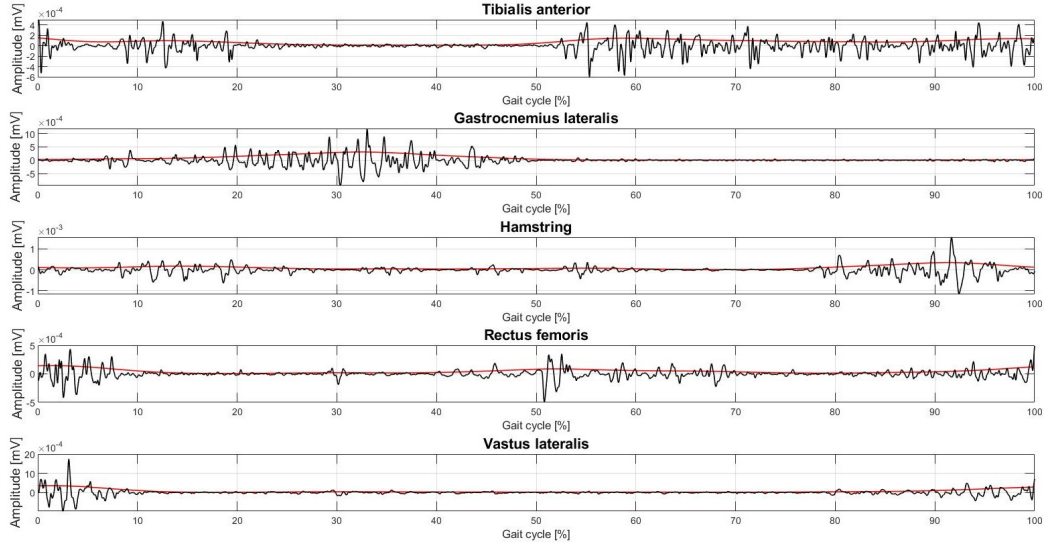


Figure 6.2: sEMG signal filtered (in black) and envelope (in red).

been chosen randomly the signal related to the left or the right side. This was done because not all the recordings related to the pathological subjects included both right and left side acquisitions, and the analysis cannot be performed on data with missing features.

6.1.2 Data processing

The initial phase of the proposed algorithm (Figure 6.1) is the data processing. The sEMG data processing was executed on the signals applying a band-pass 4th order Butterworth filter with cut-off frequencies of 20-450 Hz for the removal of external noises. The data were further processed to extract the signal linear envelope, needed to reduce the frequency content of the sEMG and to lower the sampling rates and memory storage; it also makes easier the interpretation and the detection of the onset of the muscular activity. The first step of the linear envelope was a the full-wave rectification computation, that consists in the derivation of the signal absolute value, the second step was a low-pass filtering stage performed through the application of a Butterworth filter of 1th order and cut-off frequency of 5 Hz. In Figure

6.2 are depicted the filtered signal and the signal envelope extracted for five muscles related to one single gait cycle. Finally, the last elaboration stage before feeding data into the classifier, was the *min-max normalization*, done to change the feature scale to the range $[0, 1]$, to guarantee that each feature belongs to the same range. The min-max normalization is applied to each sEMG sample $x^{(i)}$ as follows (6.1)

$$x_{norm}^{(i)} = \frac{x^{(i)} - x_{min}}{x_{max} - x_{min}} \quad (6.1)$$

where x_{max} and x_{min} are, respectively, the maximum and the minimum value of the sEMG signal associated to one muscle. The data normalization was done to reduce the variability that characterizes the signals related to the different muscles and, at the same time, preserve the information characteristic for each single muscular activation.

6.1.3 Data windowing

Following the approach proposed by Morbidoni et al. [29], the experiments were done with windows of equal length each constituted by an sEMG vector as input and a label that indicates if that window refers to a healthy, a W1 or a W2 subject.

The data preparation is done splitting the sEMG signal into consecutive segments each of which containing 2000 consecutive samples. A sliding approach was employed, to increase the number of data to feed into the classifier. This was done in a way that each window has an overlap of 200 samples with the following one. Each window is constituted of a sequence of five elements, related to the sEMG-signal value of each of the five muscles for a single instant of time. With this approach, each sEMG vector was composed of a 2000 sequence of five elements. After that the original sEMG signal has been subdivided into windows, these have been aggregated into a single vector of 10000 elements. The structure of the input so structured is illustrated in Figure 6.3. These sequences of sEMG samples are related only to the walking trial, because the network is intended to classify data related to a walking

event, and those muscle activation not correlated with the walking are considered as noise and then discarded. The reference for the identification of the first instant of activity was the foot-switch signals, that for this issue was converted in a two-level signal (stance and swing), to make possible the identification of the first swing phase of each acquisition.

6.1.4 K-fold validation

A large variability of performance was expected depending on the partition of training and test examples. For this reason, the model performance evaluation was improved using a k-fold cross-correlation criterion, considering different sets of train and test data. Six folds have been set, keeping the proportion between the different classes of subjects, chosen by randomly combining the data that constitute the present dataset. The dataset was split into training set (*Learned Dataset, LD*) and test set (*Unlearned Dataset, UD*). In the LD was included a number of subject equal to about 80% of the whole dataset, while the remaining 20% is used for the model testing. The composition of the six folds was done ensuring that each subject is included in the train set at least once. Additionally, the LD was further split into two subgroups: LD-train, used for the model training, that includes the 90% of

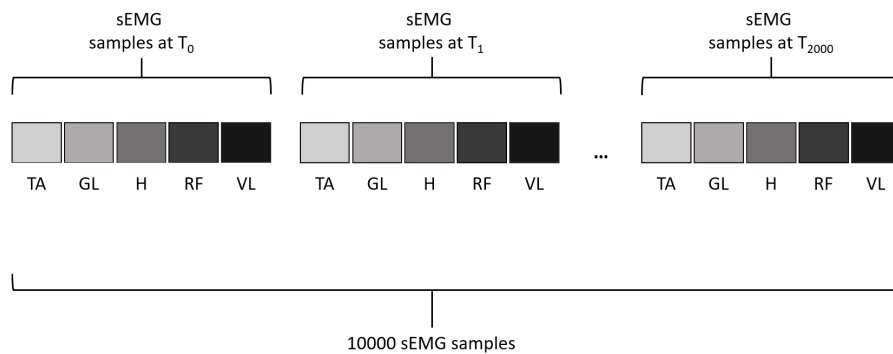


Figure 6.3: The structure of sEMG vector fed as input to the neural network. TA = Tibialis Anterior, GL = Gastrocnemius Lateralis, H = Hamstrings, RF = Rectus Femoris, VL = Vastus Lateralis.

Table 6.1: Overview of the MLP architecture.

Model Name	Model Structure	Dropout
MPL_1	1024,512,128	
MPL_2	128,64	
MPL_3	1024,521,128	$\rho = 0.5$

the sample for each subject, and LD-test, that constitutes the development set and includes the remaining 10% of LD samples.

The classifier has been trained only by the LD-train dataset, while its performances were tested both on LD-test and UD.

6.2 Neural network algorithm

In the study, different MLP have been trained and tested, some of them summarized in Table 6.1, to evaluate the effectiveness of different architecture in handling the considered classification problem. These models consist of one input layer, where the sEMG vectors and a label indicating the sEMG class are fed as input features and ground truth, a different number of hidden layers and neurons that constitute the single hidden layers and an output layer, that provides a prediction related to the neuromuscular condition identified. Dropout was tested, too. It is one of the most commonly used and the most powerful regularization techniques used in machine learning. Dropout is applied during the training time shutting down a certain percentage of the neurons present in the hidden layer, indicated by the dropout rate ρ , so that the model will not be dependent on a particular set of weights or patterns, thus overfitting. The architectures of the neural network for both multiclass than binary cascade classifier are the same, the difference is in the final output, that is a vector of dimension $y \in \mathbb{R}^3$ and $y \in \mathbb{R}^1$, respectively, and so the choice of the last activation function. The maximum possible epochs was set to 100. Furthermore, it was also chosen to impose an early-stop technique on the model updating rule, aiming to stop the model update if accuracy did not improve after 15 consecutive epochs. Once the training of the artifi-

cial neural network was completed, the best models obtained from each fold were saved. Due to the huge amount of data in training and test, these were grouped in batches of size equal to sixteen was considered. Different learning rate was tested, and the optimum value was chosen as 0.01.

The classification of the three neuromuscular conditions was handled following two different approaches. The first implementation was the design of a multiclass classifier, aimed at the identification of one of the three neuromuscular conditions considered on the basis of the machine learning interpretation of the sEMG data. Furthermore, an alternative to this first approach has been evaluated: in this case, it has been implemented a two-stage binary classifier. The first stage is designed to discriminate healthy versus hemiplegic subjects, while the second one was designed to discriminate the two forms of hemiplegia considered.

Multiclass classifier

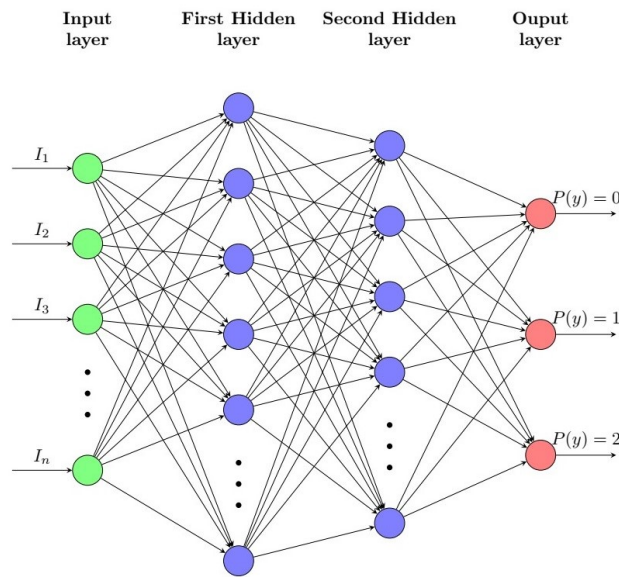


Figure 6.4: Architecture of the proposed multiclass MLP network. The input is constituted by the sEMG vectors while output computed through softmax, logarithmic softmax or tanh function, is a vector of dimension $y \in \mathbb{R}^3$.

The first neural network proposed is a multiclass classification system, designed for automatically distinguish healthy, W1 and W2 sEMG signals (Figure 6.4). This classifier has been trained and tested with all the data present in the considered dataset, that was split into LD, that includes 10 W1 subjects, 14 W2 subjects and 24 healthy subjects, and UD that includes 3 W1 subjects, 3 W2 subjects, and 6 healthy subjects. For each input data, in the output layer three values have been obtained. Each of these values is the probability that the considered input refers to each of the three possible output classes considered. Higher is that value, higher is the probability that the example belongs to a certain class. The prediction is given taken the maximum of all these three probabilities. The activation function tested to get the model prediction are the softmax, the logarithmic softmax and the hyperbolic tangent, which differs is the shape, the slope and the range in which the outputs are mapped. The one that gives better results was the Logarithmic Softmax.

Binary cascade classifier

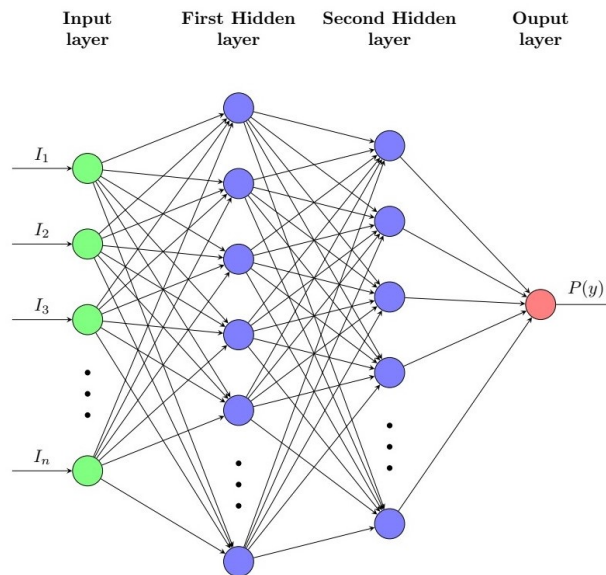


Figure 6.5: Architecture of one stage of the binary classifier MLP network. The input is constituted by the sEMG vectors while output computed through sigmoid function, is a vector of dimension $y \in \mathbb{R}^1$.

The second approach consists of two stages of binary classification: the first stage was designed in order to classify only healthy and hemiplegic pattern, while the second one is designed to discriminate the two forms of hemiplegia for those subjects identified pathological in the first stage (Figure 6.5). The activation function used to get the class prediction of either stage is the sigmoid function, which output is a scalar value comprises in the range $[0, 1]$. If the sigmoid function result is lower than a certain threshold, set to 0.5, the output is equal to zero, otherwise, it is equal to one. In this way it was possible to associate to a given sEMG input vector the respective label referred to the healthy (0) or pathological (1) condition, or W1 (0) or W2 (1) condition, respectively in the first and second stage of the binary classifier. The two models are therefore cascaded: each sEMG vector enters the first neural network, and in the case it is recognized as healthy, that input window is labeled as healthy. Otherwise, the data is fed as input to the second stage, and at this point this second neural network assigns a binary value to the given input, depending on whether the window considered is predicted as W1 or W2.

The dataset used to train and test the first stage of the binary classifier comprises all the data, split into training set and test set by keeping the proportion among the considered classes: the LD is composed by 24 pathological subjects (10 W1 and 14 W2) and 24 healthy subjects, and the LD includes the remaining 6 pathological subjects (3 W1 and 3 W2 subjects) and 6 healthy subjects.

The second stage was trained exclusively with hemiplegic subjects. The dataset preparation was done the subdivision into train and test set keeping the same proportion between the two classes as above: thirteen subjects, 13 W1 and 17 W2 have been considered, the LD is composed by 10 W1 subjects and 14 W2 subjects and the remaining 6 subjects constitute the LD. Particular attention was paid in the preparation of the folds to perform those experiment, making sure to keep the same division of the subject into LD and UD both in the first and second stages: the hemiplegic subjects that belong to a determined fold was included in the same fold to set the second binary classifier.

6.3 Performance evaluation

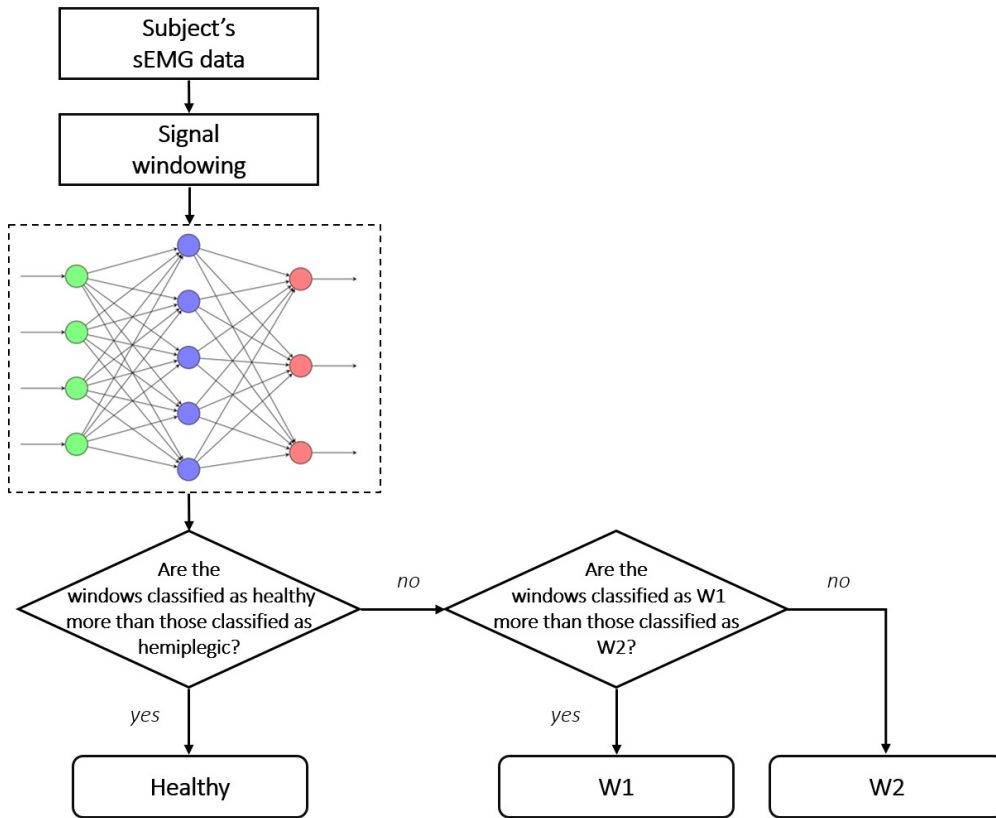


Figure 6.6: Schematization of multiclass classification process as two-stage cascade binary classifier.

The model's input is an array of windows related to 1 s of acquisition, and neural network training is done to automatically determine if that signal belongs to a healthy subject, W1 or W2 subject. The classifier has been trained only by the LD-train dataset, while its performances were tested both on LD-test and UD, on those data that it has not seen in the phase of training. The classifier performances reported in the following sections have been evaluated both on the basis of its ability to correctly identify the single windows, and further, more important from the clinical point of view, the neuromuscular condition of each subject. This evaluation was done giving as input to the neural network all the windows that belong to any individual

subject, and it works by assign to each of those windows a label. After that, it is counted the number of windows labeled as class 0, W1 or W2. At this point, following a binary classification approach, the subject is stated healthy if the overall number of windows labeled as healthy is higher than those labeled as pathological, as depicted in the pipeline reported in Figure 6.6. If the subject is assumed to be pathologic, the discrimination between the two forms of hemiplegia is done counting if the number of windows identified as W1 is higher than those predicted as W2. The correctness of the classifier was then evaluated by counting the number of subjects correctly identified (true positive TP), the number of correctly recognized subjects that does not belong to the predicted class (false positive FP), and subjects that either are incorrectly assigned to the class (false positive FP) or that were not recognized as class subjects (false negative FN). These for counts are used for the computation of a set of variables that makes possible to assess the goodness of the model performances. Moreover, these variables are collected in the confusion matrix, a squared matrix that gives an immediate understanding of the classification ability in the correct class identification. This is a $k \times k$ matrix, where k is the number of total classes, and the rows and columns represent predicted versus true classes. The elements in the diagonal are the TP, while the other ones are the subjects incorrectly classified. Ideally diagonal, the number of elements out of diagonal, that indicate the subject not correctly identified, is useful in the evaluation of the distribution of the errors. The results presented in the following section the result of the average over the six fold. The assessment of the model performances is defined, for any classes, by computing a series of measures that qualify the goodness of a multiclass classifier. The measures computed, for any classes, are precision, recall, and f1-score [28]. The precision define the classifier’s ability to accurately recognize the correct class. It is the proportion of identified positives that are indeed positive, is computed, as reported in 6.2 as the number of correctly classified positive examples by the number of subjects labeled by the system as positive

$$precision = \frac{TP}{TP + FP} \quad (6.2)$$

The recall (6.3) would indicate the ability of the classifier of not to generate false detection. It is the proportion of TP that are correctly identified as such and is computed as the ratio between the number of correctly classified positive examples and the number of positive examples.

$$recall = \frac{TP}{TP + FN} \quad (6.3)$$

The f1-score, computed as above (6.4) is the harmonic mean of precision and recall.

$$f1 - score = 2 \frac{(precision \cdot recall)}{recall + precision} \quad (6.4)$$

Chapter 7

Results

The present chapter is divided into two main sections: in the first sections are presented the performances achieved by the multiclass classifier, while in the second sections are shown the results of the binary cascade classifier, reporting the details related to the first and the second stage of binary classification.

7.1 Multiclass classifier

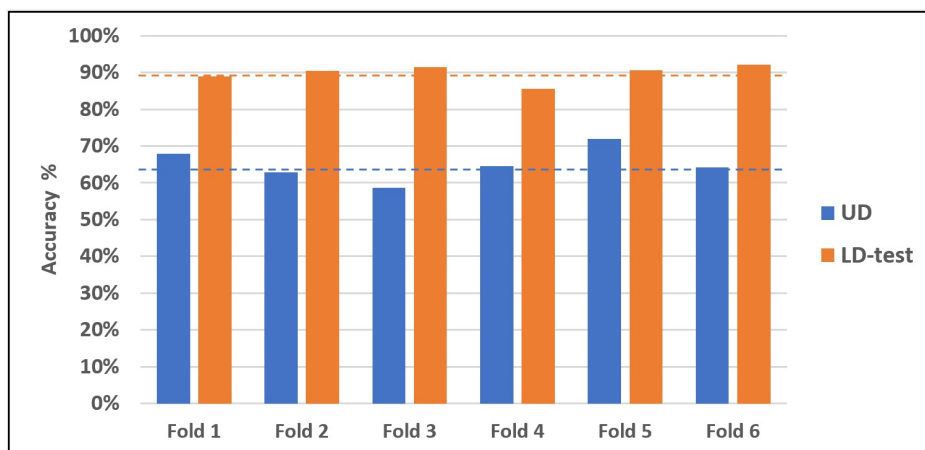


Figure 7.1: Accuracy in the windows classification achieved by the multiclass classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.

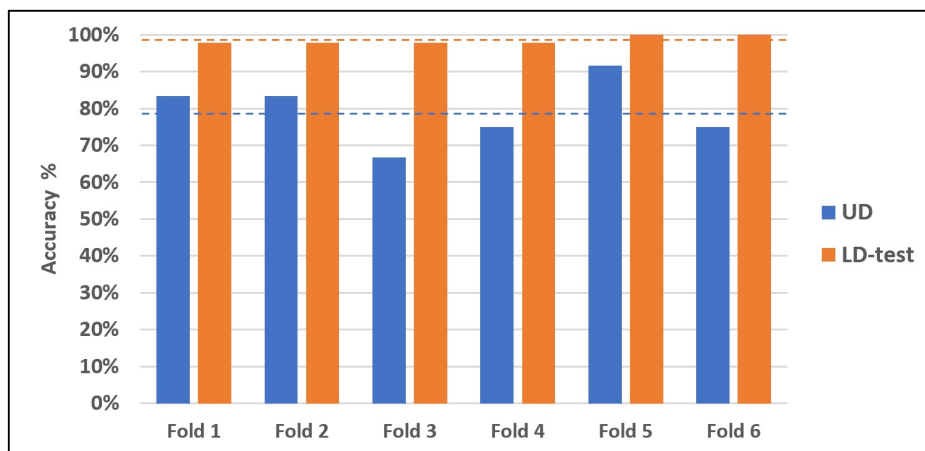


Figure 7.2: Accuracy in the subjects classification achieved by the multiclass classifier on the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.

Table 7.1: Multiclass system performances in the windows classification averaged on the six folds. The three rows are related to the scores on windows related to healthy (first row), W1 (second row) and W2 (third row) subjects.

(a) Score on UD

Precision	Recall	F1-score
0.72	0.83	0.77
0.59	0.48	0.52
0.56	0.50	0.53

(b) Score on LD-test

Precision	Recall	F1-score
0.93	0.90	0.91
0.86	0.88	0.87
0.88	0.92	0.90

Table 7.2: Multiclass system performances in the subjects classification averaged on the six folds. The three rows are related to the scores on healthy (first row), W1 (second row) and W2 (third row) subjects.

(a) Score on UD

Precision	Recall	F1-score
0.83	1.00	0.90
0.81	0.50	0.59
0.67	0.67	0.65

(b) Score on LD-test

Precision	Recall	F1-score
1.00	0.97	0.99
0.95	1.00	0.98
0.99	1.00	0.99

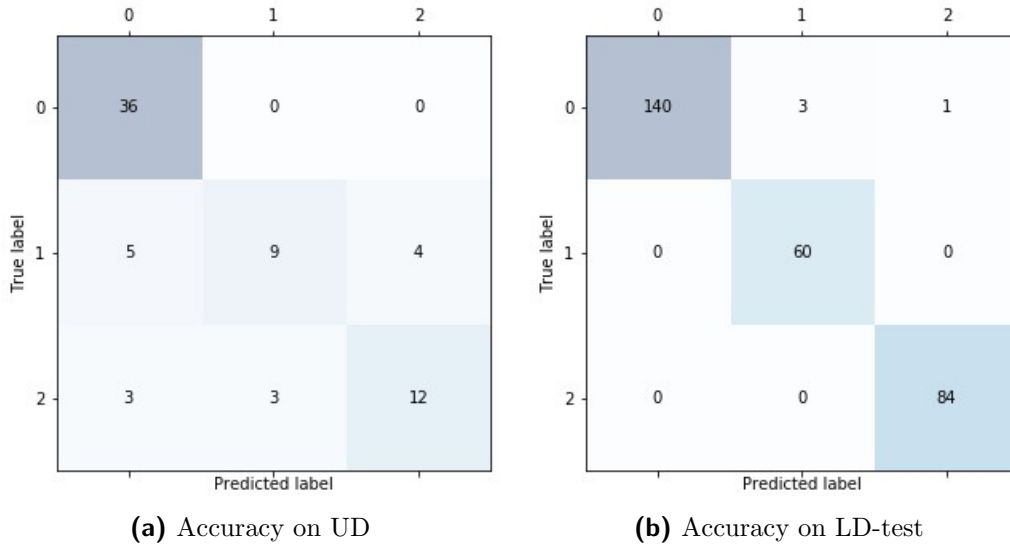


Figure 7.3: Confusion matrix related to the discrimination between healthy subjects (0), W1 subjects (1) and W2 subjects (2) performed on either UL set (left) and LD-test (right).

Among the different MPL architecture trained and tested, the MLP_1 is the one that gave the best performances as multiclass classifier. The classification accuracy achieved by this classifier for the task of subjects classification was higher with respect to the accuracy of the windows classification on every fold. Moreover, in both cases, as reported in Figure 7.1 and in Figure 7.2, the system achieved higher accuracy for those data belonging to LD-train with respect to those belonging to the UD. Indeed in learned conditions, the mean accuracy across the six folds is equal to 89.9% and 98.6% in the analysis carried on the identification of windows and subjects, respectively, while in unlearned conditions those averages values are equal to 65.1% and 79.2%. In unlearned condition precision, recall and f1-score associated with the correct identification of healthy subjects and healthy windows were higher than those associated with the correct identification of the two forms of hemiplegia, as indicated in Table 7.1(a) and Table 7.2(a). In fact, for those subjects belonging to the UD, the classifier was able to correctly identify all the healthy subjects, while there was a higher number of hemiplegic subjects that the system is not able to identify: out of a total of 18 W1 and 18 W2

subjects present in the UD are correctly recognized as such only 9 and 12, respectively (Figure 7.3). This lower ability in the identification of the different hemiplegic forms was determined by the presence of a total of 8 hemiplegic subjects (5 W1 and 3 W2) that the system had labeled as healthy, and in addition to this, the presence of some mistakes in the discrimination between the two forms of hemiplegia. Differently in learned conditions, as reported in Table 7.1(b) and Table 7.2(b), classification score related to the correct identification of the three neuromuscular states was similar for either control subjects, W1 and W2 subjects. Indeed, for what concerns the LD-test subjects, the system classified in the correct manner almost all the control subjects, with a total of 284 out 288 correct predictions.

7.2 Two-stage binary classifier

First stage of binary classification

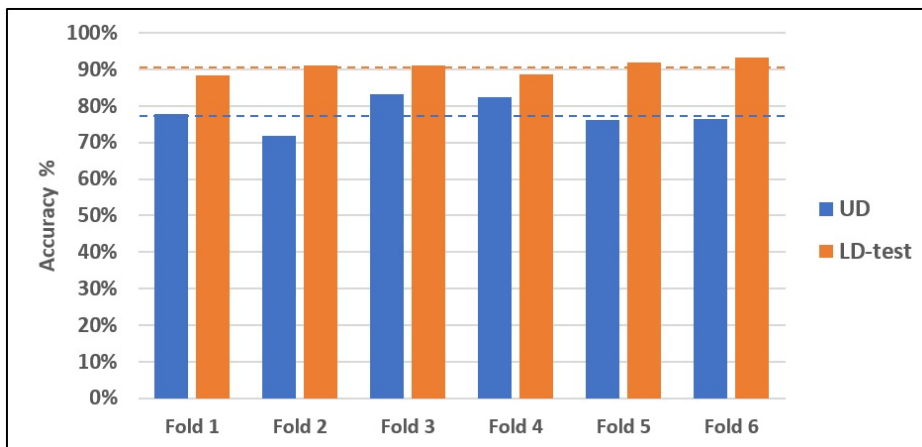


Figure 7.4: Accuracy in the windows classification achieved by the first stage of binary classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.

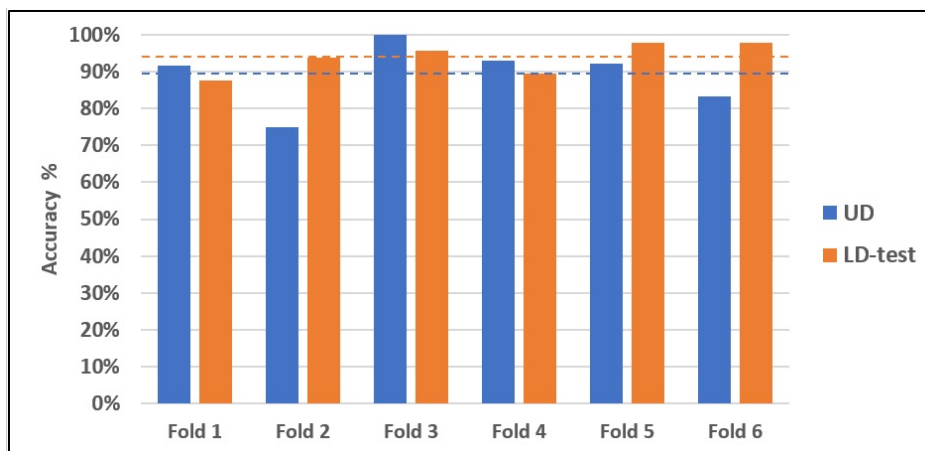


Figure 7.5: Accuracy in the subjects classification achieved by the first stage of binary classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.

Table 7.3: First stage of binary classification performances in the windows classification averaged on the six folds. The two rows are related to the scores on windows related to healthy (first row) and hemiplegic subjects (second row).

(a) Score on UD			(b) Score on LD-test		
Precision	Recall	F1-score	Precision	Recall	F1-score
0.75	0.80	0.77	0.94	0.86	0.90
0.82	0.76	0.79	0.88	0.95	0.91

Table 7.4: First stage of binary classification performances in the subjects classification averaged on the six folds. The two rows are related to the scores to healthy (first row) and hemiplegic subjects (second row).

(a) Score on UD			(b) Score on LD-test		
Precision	Recall	F1-score	Precision	Recall	F1-score
0.84	0.97	0.90	1.00	0.87	0.93
0.97	0.81	0.88	0.89	1.00	0.94

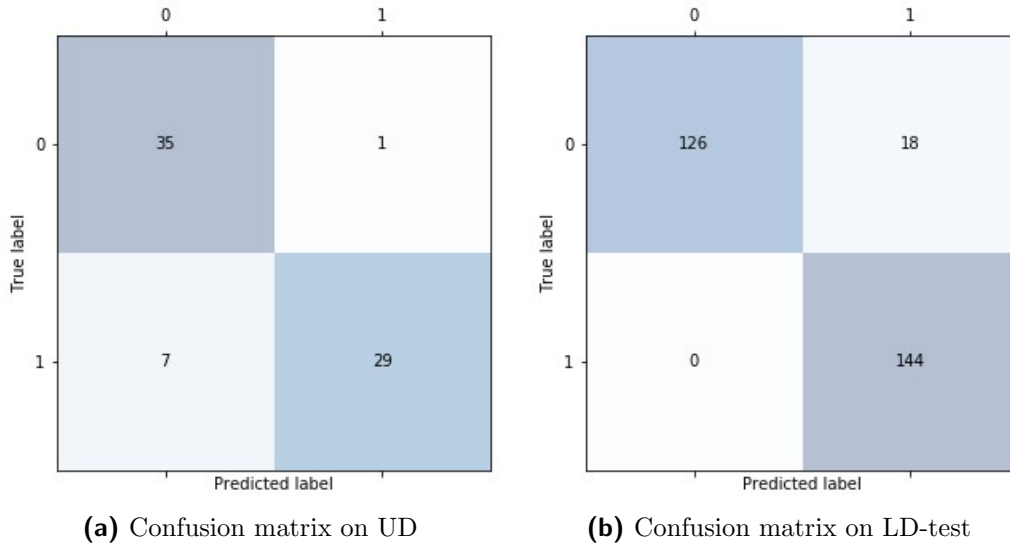


Figure 7.6: Confusion matrix related to the discrimination between healthy subjects (0) and pathological subjects (1) performed on either UD set (left) and LD-test (right).

The binary classifier that achieved better discrimination of healthy and pathological subjects is the MPL_2 . This MLP architecture in the first stage of binary classification reported higher accuracy for the task of subject classification (mean accuracy = 88.9% and 93.8%, in learned and unlearned conditions, respectively) - Figure 7.5, with respect to the accuracy for the task of windows classification (mean accuracy = 78.0% and 90.7%, in learned and unlearned conditions, respectively) - Figure 7.4. In both cases, the accuracies were higher for learned data than for unlearned data. The MLP_2 binary classifier reached optimal values for the discrimination of the windows related to control and hemiplegic subjects, as reported in Table 7.4. Moreover, this system was able to distinguish healthy from hemiplegic subjects in either learned than unlearned conditions, as reported in Table 7.5. The high rate of correct predictions is reported in the confusion matrix in Figure 7.6, which illustrates that the number of total correct predictions is equal to 64 out 72 subjects and 268 out 284 subjects, respectively in learned and unlearned conditions.

Second stage of binary classification

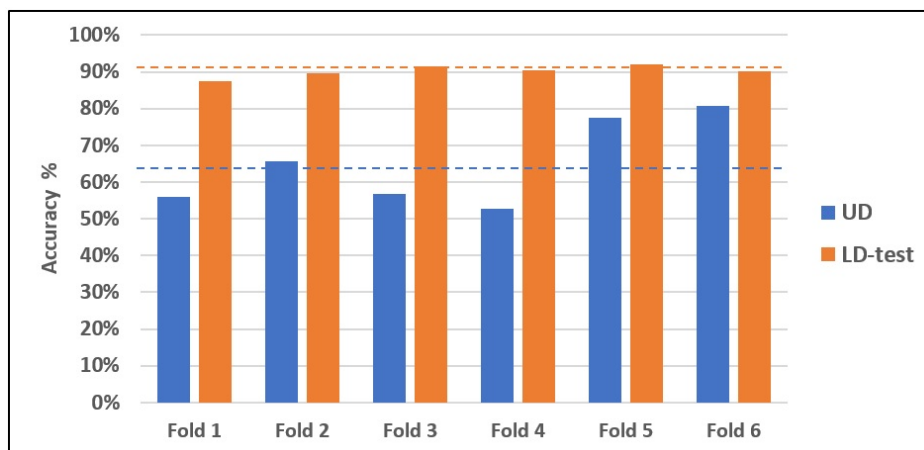


Figure 7.7: Accuracy in the windows classification achieved by the second stage of binary classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.

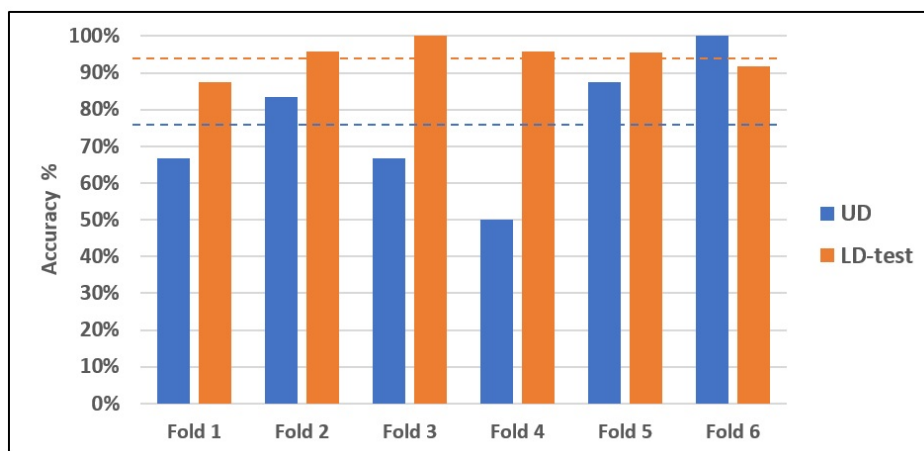


Figure 7.8: Accuracy in the subjects classification achieved by the second stage of binary classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.

Table 7.5: Second stage of binary classification performances in the windows classification averaged on the six folds. The two rows are related to the scores on windows related to W1 (first row) and W2 subjects (second row).

(a) Score on UD			(B) Score on LD		
Precision	Recall	F1-score	Precision	Recall	F1-score
0.55	0.50	0.52	0.92	0.88	0.90
0.58	0.47	0.52	0.81	0.89	0.85

Table 7.6: Second stage of binary classification performances in the windows classification averaged on the six folds. The two rows are related to the scores related to W1 (first row) and W2 subjects (second row).

(a) Score on UD			(b) Score on LD-test		
Precision	Recall	F1-score	Precision	Recall	F1-score
0.70	0.62	0.65	0.87	0.92	0.89
0.62	0.68	0.64	0.93	0.89	0.91

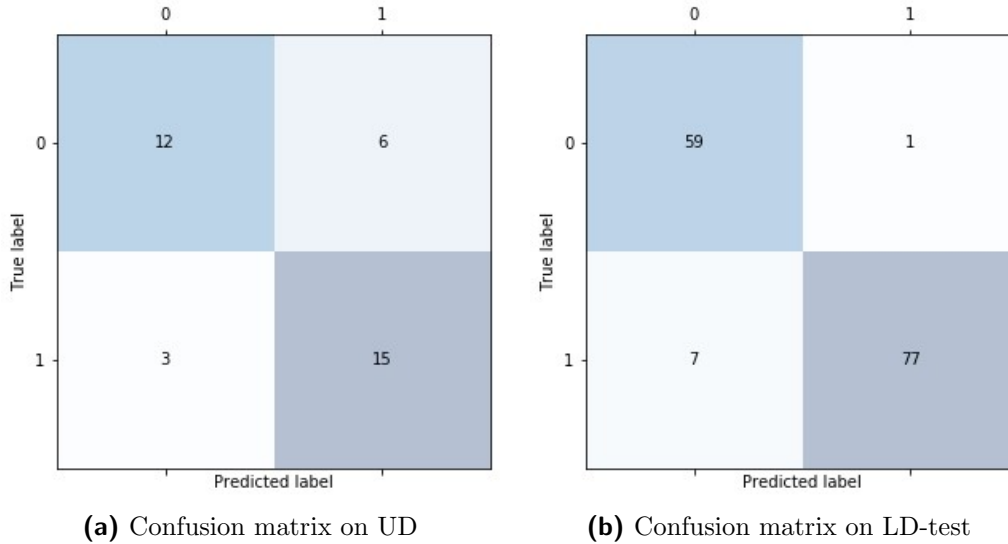


Figure 7.9: Confusion matrix related to the discrimination between W1 subjects (0) and W2 subjects (1) performed on either UD set (left) and LD-test (right).

For the implementation of the second stage of the binary cascade approach, the MLP architecture that reported the best results is the MLP_3 . This binary classifier, trained for the discrimination between the two forms of hemiplegia, in LD-test, as reported in the Figure 7.7 and Figure 7.8 achieved across the six folds an average accuracy of 90.2% and 94.4%, in the analysis conducted per window and subjects, respectively. Regarding the performance under unlearned conditions, the same Figures reported an average accuracy of 64.9% and 75% for the analysis by subject and by window. The analysis shows that higher scores were achieved in the subjects classification than in the windows classification task. Furthermore, better performances were reported in the conditions learned. In both learned that unlearned data a certain rate of error is reported in the discernment between the two forms of hemiplegia, which turned out to be more evident in unlearned conditions. The discrimination between W1 subjects and W2 subjects, as reported in Table 7.5 reached, approximately, the same results. The rate of correct and incorrect prediction for the unlearned subjects is reported in the confusion matrix in Figure 7.9(a). It indicates that the system correctly identified as W1 subjects 12 out of 18 subjects, reporting a number of 12 TP and 6 FP, while it was more robust in identifying the form of hemiplegia W2, being able to correctly identify 15 of 18 subjects W2. Moreover, in the same Figure it is reported the errors distribution in the discrimination of hemiplegic subjects performed in learned conditions, in which were correctly identified 59 out of 60 Winters' type I subjects and 77 out of 84 Winters' type II subjects.

Binary cascade classification

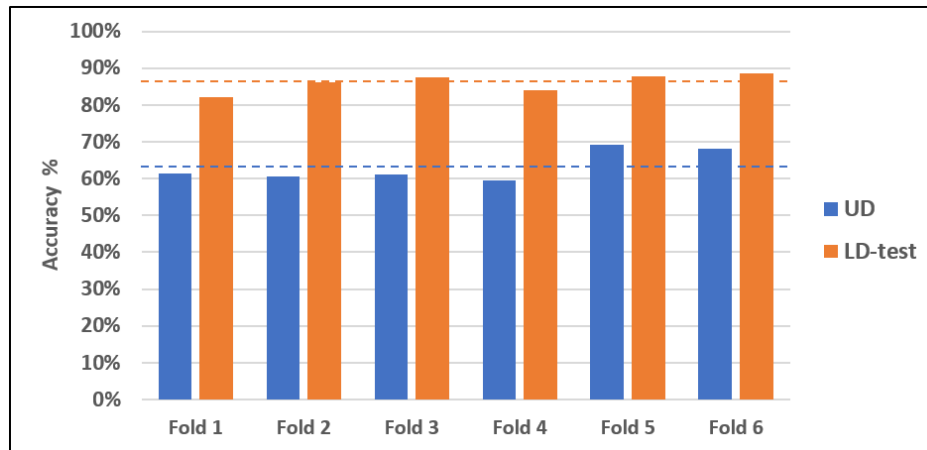


Figure 7.10: Accuracy in the windows classification achieved by the binary cascade classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.

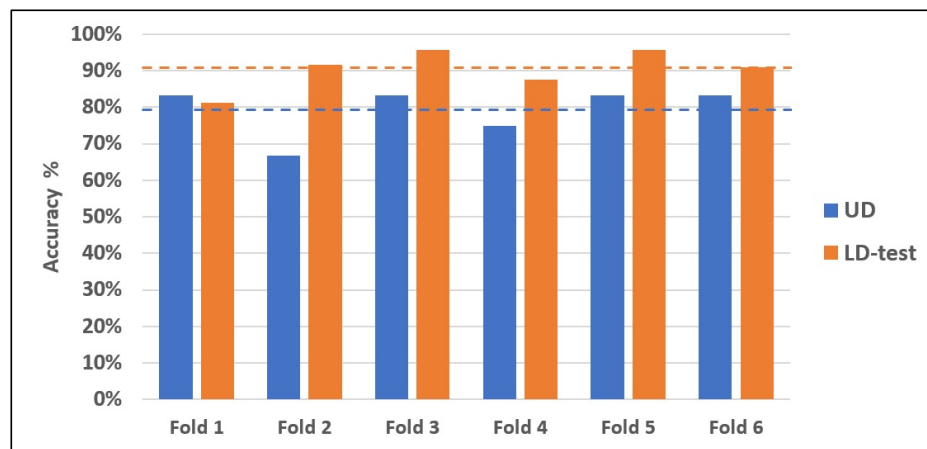


Figure 7.11: Accuracy in the subjects classification achieved by the binary cascade classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.

Table 7.7: Binary cascade classification performances in the windows classification averaged on the six folds. The three rows are related to the scores on windows related to healthy (first row), W1 (second row) and W2 (third row) subjects.

(a) Score on UD			(b) Score on LD-test		
Precision	Recall	F1-score	Precision	Recall	F1-score
0.75	0.80	0.77	0.94	0.86	0.90
0.57	0.46	0.50	0.76	0.86	0.80
0.52	0.56	0.52	0.84	0.85	0.85

Table 7.8: Binary cascade classification performances in the windows classification averaged on the six folds. The three rows are related to healthy (first row), W1 (second row) and W2 (third row) subjects.

(a) Score on UD			(b) Score on LD-test		
Precision	Recall	F1-score	Precision	Recall	F1-score
0.84	0.97	0.90	1.00	0.87	0.93
0.81	0.50	0.59	0.77	0.98	0.86
0.75	0.72	0.73	0.94	0.93	0.93

The cascade of the first and second binary classifiers performed better in the subjects than in the windows discrimination. This happens, in either case, as reported in Table 7.10 and Table 7.11 with higher accuracy on the learned test set (reporting mean accuracy across the six folds equal to 86.1% and 90.5% in the analysis taken on windows and on subjects, respectively) than in the unlearned test set (reporting mean accuracy equal to 63.4% and 79.2% in the analysis taken on the windows and on the subjects, respectively). The analysis on the learned set, which details are reported in Table 7.7(b) and in Figure 7.12, indicates that in these conditions the classification scores were similar for the three classes. On the other hand for what concern the analysis carried out on the unlearned data, the Table 7.11 and Figure 7.12(a) assess that the classifier achieved good performances in the identification of the healthy subjects (f1-score = 0.90) - Table 7.8(a) being able to correctly identify almost all the healthy subjects, reporting only 1 FP.

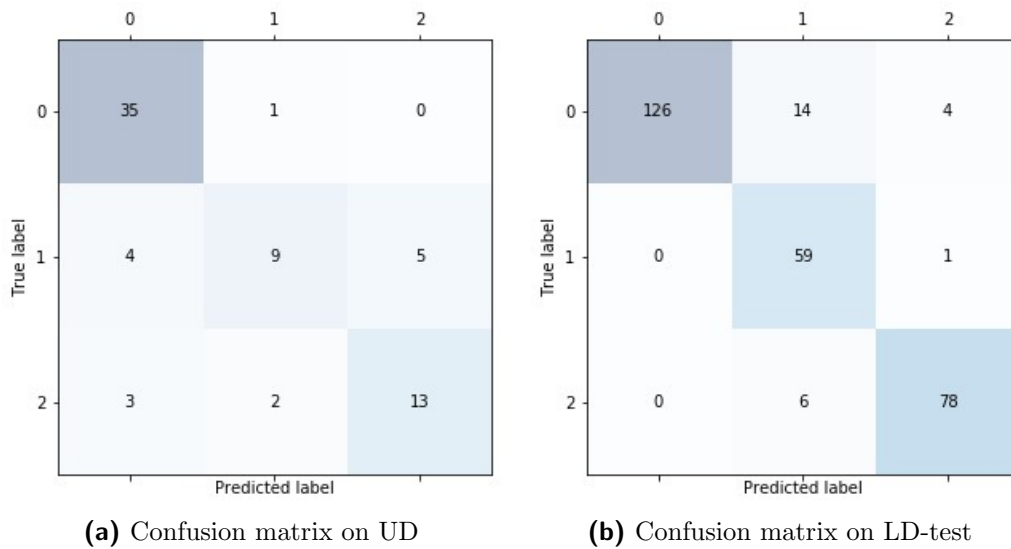


Figure 7.12: Confusion matrix related to the discrimination between healthy subjects (0), W1 subjects (1) and W2 subjects (2) performed on either UD set (left) and LD-test (right).

The identification of form I and form II of hemiplegia, on the other side, were associated to a lower number of correct predictions. Indeed, the confusion matrix in Figure 7.12(a) illustrates that the number of correctly identified W1 subjects was equal to 9 out of 18, and the number of correctly W2 subjects was equal to 13 out of 18.

Chapter 8

Discussion

The purpose of this work was the implementation of a machine learning algorithm to address a multiclass classification problem through the automatic interpretation of EMG signals. The goal, in fact, was the development of a system able to automatically discern, through the analysis of the EMG data, if a subject could be considered healthy or hemiplegic. In addition, it has been proposed to optimize the prediction on the pathological subjects by distinguishing two forms of hemiplegia on the bases of the Winters' classification [6]. To our knowledge, in the current state of the art no attempts were made for the automatic classification of Winters' type I and Winters' type II subjects through machine learning interpretation of EMG signal. The aim of the study was pursued through the implementation of different MLP architectures. The EMG signals fed as input to the neural networks were subjected to a pre-processing stage, at the end of which the EMG signal envelope has been computed. The training of the neural network was based on a supervised learning approach, in which the inputs were a 1-second-long EMG vector and the ground truth, from which the artificial neural network learned, had different values, each one associated with the three different neuromuscular states considered. Two different implementations have been proposed, working as follows. The first approach was designed to automatically identify the subjects belonging to the three neuromuscular conditions proceeding through a multiclass approach, while the second method is based

on the combination of two binary classifiers: the first classifier trained to discern healthy subjects from hemiplegic ones, and the second to distinguish the form I and form II of hemiplegia. As reported in Figure 7.1 and 7.10, the results that both classifiers achieved for LD-test data in assign the correct label to each window are comparable, reporting the first approach an average accuracy across the six folds equal to 89.9% and the second approach an accuracy equal to 86.1%. Also for what concern the classifier ability in the identification of the neuromuscular state of the different subjects both systems show optimal results, with accuracy equal 98.6% and 90.5% (Figure 7.2 and 7.11). Either the first designed classifier and the second one were able to generalize the predictions on data unseen in phase of training, reporting in those condition for the windows discrimination an overall accuracy equal to 98.6% and 90.5% (Figure 7.1 and 7.10) for the first and second classifier, respectively, while for the recognition of the subject both systems achieved an accuracy equal to 79.2% (Figure 7.2 and 7.11). This demonstrates the greater ability of both classifiers in the correct assignment of labels for the data associated with subjects on which the model has been trained. However promising performances were also reported in the analysis in unlearned conditions, demonstrating the classifiers' ability to generalize even on individuals never seen before. Furthermore both classifiers reported better performances on the analysis by subject compared to the analysis by windows, which highlights the actual difficulty of correctly labeling the single window passed in input to the neural network. This improvement, evident both in the learned and unlearned conditions, is obtained through a system of results optimization consisting in the usage of a mechanism with binary decision thresholds, through which it was possible to increase the rate of correct predictions. This mechanism works by counting the percentage of the entire gait pattern associable with that of a healthy, W1 and W2 subjects, and assigning the class healthy if the percentage of the entire signal classified as healthy is higher than half of the total myoelectric recording. Otherwise, if the greatest percentage of the walk is recognized as belonging to a W1 subject, the classifier would assign this label to the subject, and vice versa. This choice is explained with the greater difficulty of identifying patterns

associated with hemiplegic subjects with respect to patterns associated with control subjects. In fact, as reported in the literature, the typical pattern of hemiplegic subjects can alternate, with different percentages, to the healthy walking pattern, making the effective identification and discrimination from healthy subjects more challenging.

The neural architecture that made possible to get optimal performances for the multiclass classifier consisted of three hidden layers, consisting of 1024, 512 and 128 neurons, respectively, and each fold trained with a test set consisting, in mean, of 57.870 labeled data. In learned conditions, it reported an optimal correct prediction rate, being able to assign the correct label to a total of 284 subjects out of 288, as reported in Figure 7.3(b). For what concerns the classification of unseen data, it achieved excellent performances in the correct identification of healthy subjects, reporting precision equal to 1.00 - Table 7.2(b). Lower accuracies were reported in the task of assigning the correct labels to the hemiplegic subjects. Indeed, if this task was done almost correctly in learned conditions, the system is less accurate in generalizing data related to the pathological subjects that the network has never seen. This issue is evident in the distinction of hemiplegia with respect to control subjects, and in addition to the presence of pathological subjects not correctly identified as such, there is a difficulty in the discernment between the two categories of hemiplegia. This issue could be explained by recalling that it has not yet demonstrated the presence of a single muscular recruitment characteristic for each form of hemiplegia. Moreover, the clinical presentations and the symptoms of the two forms of hemiplegia, for some traits, tend to overlap [15,16]. This problem is more evident in the correct identification of W1 subjects, a task that reports a high rate of incorrect predictions, with a recall equal to 0.56 - Table 7.2(a), indicating the effective difficulty in correctly identifying the mildest form of hemiplegia.

The second system of classification solved, even if in part, the problem of discriminating form I and form II of hemiplegia, providing a more efficient tool for the distinction between the class W1 and W2. This solution is offered by the use of two levels of binary classification.

The first binary classifier, consisting of two hidden layers of 128 and 64 neural

units respectively, has been trained with a total of 57.869 labeled EMG vectors per fold. It, designed for the discernment between healthy and pathological subjects, reported optimal accuracy in learned and unlearned conditions, as reported in Figure 7.5. Indeed, even if the proportion of correctly labeled subjects is higher in learned conditions, the classifier is able to generalize on data unseen in the phase of train. On this dataset, the better results were achieved in the recognition of healthy subjects, task that reported a recall value equal to 0.97 - Table 7.4(a). On the other side, the recognition of the hemiplegic subjects is more challenging, with a lower corrective prediction rate (recall = 0.81) - Table 7.4(a). In fact, the number of pathological subjects correctly identified is equal to 29 out of 36, with 7 hemiplegic subjects incorrectly labeled as healthy, as reported in Figure 7.6(a), in a similar way to what happens for the multiclass classifier.

The second binary classifier was trained with a mean value of 29.958 labeled windows per fold and achieved better performances using the MPL_3 network (3 hidden layers, composed of 1024, 512, 128, and drop out rate $\rho = 0.5$). It, aimed at distinguishing the two forms of hemiplegia, enhanced the identification of subjects W1 and W2. This occurred with optimal results in unlearned conditions, reporting a mean accuracy equal to 94.4%. In unlearned conditions, the system reported a mean accuracy equal to 75.7% (Figure 7.8). In this case, there is a greater error rate in the discrimination between the two forms of hemiplegia, justified taking into account that the two forms of hemiplegia could share some clinical presentation. Nevertheless, the system reached a promising value of correct prediction rate, being the number of correct predictions equal to 27 out of 36 subjects, as reported in the confusion matrix in Figure 7.9(a).

Between the first and second stage of binary classification, the ones that achieves better performances is the first one, with an accuracy equal to 88.9%, a support of the actual higher difficulty in discerning the form I and form II of hemiplegia than discriminate a pathological form an healthy subjects though the analysis of myoelectric signal. The combination of these two binary classifiers achieved, in learned conditions, a high accuracy, being able to assign the correct label to 263 out of 288 subjects. However, the combination

of the two systems, although it found optimal values in the identification of healthy subjects, it suffered from the bad recognition of pathological subjects which the first stage of the binary classification considered healthy, and for this reason, they do not pass into the second classification stage, resulting in a propagation of the error and in a reduction of the overall performances of the classifier. A deeper understanding of the error distributions is provided by the confusion matrix reported in Figure 7.12. This matrix shows that the form I of hemiplegia (class 1 in the Figure) was the most difficult to identify, being in part confused with the healthy condition (class 0 in Figure), in part confused with the form II of hemiplegia (class 2 in figure), in accordance with the clinical observation that the W1 subjects share a certain percentage of their gait pattern with both able-bodied and W2 children.

Finally, both classifiers reported comparable performances in the identification of the three muscular states on unseen data: in both cases the optimal performances were reported for the identification of the healthy subjects, while the main trouble was the predictions of Winters' type I subjects. To this concern, the second approach offers an additional tool for discernment. In fact, splitting the task of classification into two different stages gave the opportunity to reach, with the second stage of binary classification, better results for the identification of the W1 subjects. Indeed this stage eliminates the possibility of confusing a hemiplegic subject with a healthy one, that is the crucial issue in both approach proposed.

The promising results achieved by both classifiers suggest that artificial neural networks can be successfully used to automatically classify neuromuscular states from EMG signal, without the need for human intervention in selecting relevant signal features. While in clinics hemiplegia diagnosis will probably remain a human-based task performed by specialized physicians, machine learning approaches demonstrate a promising level of accuracy and would help the clinician in prognosis, providing a powerful tool to have a non-objective, rapid and automatic gait data interpretation.

The second fundamental point which these results demonstrate is the possibility to adapt a clinical classification based on the sagittal joint angles, to an analysis based on the evaluation of muscle recruitment. Therefore, despite

the popularity of Winters' classification, the present state of the art does not recognize any predominant muscle activation patterns, although this information would have crucial clinical implications in patient management. The present study, is supporting similar results present in the literature [5,13,14,15,16], assessing a certain correlation between Winters' classification and the muscular activation pattern during walking. These findings define the possibility of using EMG data for the classification of hemiplegia in support or substitution of kinematic data, reducing the need for employing additional sensors, associated with costs and clinical encumbrance.

Future studies could be aimed at testing and evaluating the possibility of having a different application in clinics. Indeed, the algorithm is designed for diagnostic purposes and is based on the percentage of walking patterns recognized as the walking pattern typical of standard subjects, W1 or W2 subjects, respectively. In clinics this index might be used for the analysis of clinical program, through the evaluation of variation of this index before and after specific therapeutic treatment. The variation of this index before and after a certain therapeutic treatment could be useful to quantify the goodness of a certain therapeutic program. Another possible future prospective would be aimed at improving the proposed classifiers through the use of a broader EMG dataset, with the aim of obtaining a better generalization of the hemiplegic pattern, evaluating the possibility of implementing a classifier that considers also the Winters' class III and IV.

List of Figures

2.1	(A) Anatomy of the skeletal muscle (B) Muscle fiber (C) Myofibril (D) Sarcomere.	6
2.2	(A) α -motor neurons originate from the spinal cord (B) Each α -motor neurons innervate a group of muscle fibers, constituting a motor neuron [8].	7
2.3	Action potential waveform.	9
3.1	The normal gait cycle of an 8-yearold boy [3].	12
3.2	Lower limb muscles whose activation is acquired during sEMG session. (A) Rectus femoris (B) Vastus lateralis (C) Hamstrings (D) Tibialis anterior (E) Gastrocnemius.	17
4.1	Summary of the main trends in machine learning in human walking. (A) Rise of the research in the usage of machine learning methods in movement biomechanics. (B) The most commonly studied neuromuscular conditions [1].	22
5.1	(a)Analogy between the biological neuron (b) and an artificial neuron scheme [26].	26
5.2	Partitioning of dataset into training and test set.	29
5.3	A large variability of performances was expected, depending on the dataset partition into train set and test set. To make the model performances more reliable, a k-fold validation might be implemented.	30
6.1	Proposed pipeline for the sEMG classifiers.	32

6.2	sEMG signal filtered (in black) and envelope (in red).	33
6.3	The structure of sEMG vector fed as input to the neural network. TA = Tibialis Anterior, GL = Gastrocnemius Lateralis, H = Hamstrings, RF= Rectus Femoris, VL = Vastus Lateralis.	35
6.4	Architecture of the proposed multiclass MLP network. The input is constituted by the sEMG vectors while output computed through softmax, logarithmic softmax or tanh function, is a vector of dimension $y \in \mathbb{R}^3$	37
6.5	Architecture of one stage of the binary classifier MLP network. The input is constituted by the sEMG vectors while output computed through sigmoid function, is a vector of dimension $y \in \mathbb{R}^1$	38
6.6	Schematization of multiclass classification process as two-stage cascade binary classifier.	40
7.1	Accuracy in the windows classification achieved by the multiclass classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.	43
7.2	Accuracy in the subjects classification achieved by the multiclass classifier on the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.	44
7.3	Confusion matrix related to the discrimination between healthy subjects (0), W1 subjects (1) and W2 subjects (2) performed on either UL set (left) and LD-test (right).	45
7.4	Accuracy in the windows classification achieved by the first stage of binary classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.	46

7.5	Accuracy in the subjects classification achieved by the first stage of binary classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.	47
7.6	Confusion matrix related to the discrimination between healthy subjects (0) and pathological subjects (1) performed on either UD set (left) and LD-test (right).	48
7.7	Accuracy in the windows classification achieved by the second stage of binary classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.	49
7.8	Accuracy in the subjects classification achieved by the second stage of binary classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.	49
7.9	Confusion matrix related to the discrimination between W1 subjects (0) and W2 subjects (1) performed on either UD set (left) and LD-test (right).	50
7.10	Accuracy in the windows classification achieved by the binary cascade classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.	52
7.11	Accuracy in the subjects classification achieved by the binary cascade classifier over the six folds, for either UD (blue) and LD-test (orange). Mean value across six folds is represented by the two horizontal dashed lines.	52
7.12	Confusion matrix related to the discrimination between healthy subjects (0), W1 subjects (1) and W2 subjects (2) performed on either UD set (left) and LD-test (right).	54

List of Tables

6.1	Overview of the MLP architecture.	36
7.1	Multiclass system performances in the windows classification averaged on the six folds. The three rows are related to the scores on windows related to healthy (first row), W1 (second row) and W2 (third row) subjects.	44
7.2	Multiclass system performances in the subjects classification averaged on the six folds. The three rows are related to the scores on healthy (first row), W1 (second row) and W2 (third row) subjects.	44
7.3	First stage of binary classification performances in the windows classification averaged on the six folds. The two rows are related to the scores on windows related to healthy (first row) and hemiplegic subjects (second row).	47
7.4	First stage of binary classification performances in the subjects classification averaged on the six folds. The two rows are related to the scores to healthy (first row) and hemiplegic subjects (second row).	47
7.5	Second stage of binary classification performances in the windows classification averaged on the six folds. The two rows are related to the scores on windows related to W1 (first row) and W2 subjects (second row).	50

7.6	Second stage of binary classification performances in the windows classification averaged on the six folds. The two rows are related to the scores related to W1 (first row) and W2 subjects (second row).	50
7.7	Binary cascade classification performances in the windows classification averaged on the six folds. The three rows are related to the scores on windows related to healthy (first row), W1 (second row) and W2 (third row) subjects.	53
7.8	Binary cascade classification performances in the windows classification averaged on the six folds. The three rows are related to healthy (first row), W1 (second row) and W2 (third row) subjects.	53

Bibliography

- [1] E HALILAJ, A RAJAGOPAL, M FITERAU, JL HICKS, TJ HASTIE and SL DELP. *Machine learning in human movement biomechanics: best practices, common pitfalls, and new opportunities*. Journal of biomechanics 81 (2018): 1-11.
- [2] E PAPAGEORGIOU, A NIEUWENHUYS, I VANDEKERCKHOVE, AV CAMPENHOUT, E ORTIBUS and K DESLOOVERE. *Systematic review on gait classifications in children with cerebral palsy: An update*. Gait & Posture 69 (2019): 209-223.
- [3] J FIGUEIREDO, CP SANTOS and JC MORENO. *Automatic recognition of gait patterns in human motor disorders using machine learning: A review*. Medical Engineering and Physics 53 (2018): 1–12.
- [4] S ARMAND, G DECOULON and A BONNEFOY-MAZURE. *Gait analysis in children with cerebral palsy*. EFORT open reviews 1.12 (2016): 448-460.
- [5] F DI NARDO, A STRAZZA, A MENGARELLI, S CALDARELLI, A TIGRINI, F VERDINI, A NASCIMBENI, A VALENTINA, M KNAFLITZ and S FIORETTI. *EMG-Based Characterization of Walking Asymmetry in Children with Mild Hemiplegic Cerebral Palsy*. Biosensors 9.3 (2019): 82.
- [6] TF WINTERS, JR GAGE and R HICKS. *Gait patterns in spastic hemiplegia in children and young adults*. J Bone Joint Surg Am 69.3 (1987): 437-441.
- [7] MN LEVY and BM KOEPPEN and BA STANTON. *Berne & Levy Principles of Physiology*. Elsevier Health Sciences, 2005.

- [8] DE PURVES, GJ AUGUSTINE, D FITZPATRICK, WC HALL, AS LAMANTIA, JO MCNAMARA, LE WHITE. *Neuroscience*. Sinauer associates, 2008.
- [9] CL VAUGHAN, BL DAVIS and JC O'CONNOR *Dynamics of human gait*. Human Kinetics, 1992.
- [10] A CAPPELLO, A CAPPOZZO and PE DI PRAMPERO. *Bioingegneria della postura e del movimento*. Patron Editore, 2003.
- [11] J G WEBSTER and J W CLARK. *Medical instrumentation: application and design*. John Wiley & Sons, 1995.
- [12] F DI NARDO, A MENGARELLI, A STRAZZA, V AGOSTINI, M KNAFLITZ, L BURATTINI and S FIORETTI. *A new parameter for quantifying the variability of surface electromyographic signals during gait: The occurrence frequency*. Journal of Electromyography and Kinesiology 36 (2017): 25-33.
- [13] F DI NARDO, V AGOSTINI, A STRAZZA, A NASCIMBERI, M KNAFLITZ, L BURATTINI and S FIORETTI. *Gait asymmetry in Winters' group I hemiplegic children*. EMBEC & NBC (2017): 976-979.
- [14] F DI NARDO, A MENGARELLI, M MALAVOLTA, A STRAZZA, V AGOSTINI, A NASCIMBERI, M KNAFLITZ, L BURATTINI and S FIORETTI. *Ankle muscle co-contractions in Winters I hemiplegic children during gait*. Gait & Posture, 57 (2017): 4-5.
- [15] V AGOSTINI A NASCIMBENI, A GAFFURI and M KNAFLITZ. *Multiple gait patterns within the same Winters class in children with hemiplegic cerebral palsy*. Clinical Biomechanics 30.9 (2015): 908-914.
- [16] V AGOSTINI, M KNAFLITZ, A NASCIMBERI and A GAFFURI. *Gait measurements in hemiplegic children: an automatic analysis of foot-floor contact sequences and electromyographic patterns*. 2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA) (2014): 1-4.

- [17] F ZHANG, ZG HOU, L CHENG, W WANG, Y CHEN, J HU, L PENG and H WANG. *iLeg—A Lower Limb Rehabilitation Robot: A Proof of Concept*. IEEE Transactions on Human-Machine Systems 46.5 (2016): 761-768.
- [18] P WEI, R XIE, R TANG, C LI, J KIM and M WU. *sEMG Based Gait Phase Recognition for Children with Spastic Cerebral Palsy*. Annals of biomedical engineering 47.1 (2019): 223-230.
- [19] A FERRARI, L BERGAMINI, S CALDERARA, N BIOCCHI, G VITETTA, C BORGHI, R NEVIANI, and A. FERRARI. *Gait-Based Diplegia Classification Using LSMT Networks*. Journal of healthcare engineering 2019 (2019).
- [20] AJ SALAZAR, OC DE CASTRO and RJ BRAVO. *Novel approach for spastic hemiplegia classification through the use of the support vector machines*. The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol. 1. IEEE, 2004.
- [21] Y ZHANG and Y MA. *Application of supervised machine learning algorithms in the classification of sagittal gait patterns of cerebral palsy children with spastic diplegia*. Computers in biology and medicine 106 (2019): 33-39.
- [22] H HASNI, N YAHYA, VS ASIRVADAM and MA JATOI. *Analysis of Electromyogram (EMG) for Detection of Neuromuscular Disorders*. 2018 International Conference on Intelligent and Advanced System (ICIAS). IEEE, 2018.
- [23] S LAHMIRI and M BOUKADOUM. *An Accurate Automatic System for Distinguishing Neuropathy and Healthy Electromyography Signals*. 2017 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 2017.
- [24] J KAMRUZZAMAN and R K BEGG. *Support Vector Machine and Other Patter Recognition Approaches to the Diagnosis of Cerebral Palsy Gait*. IEEE Transactions on Biomedical Engineering 53.12 (2006): 2479-2490.

- [25] M ALAQTASH, T SARKODIE-GYAN, H YU, O FUENTES, R BROWER and A ABDELGAWADI. *Automatic Classification of Pathological Gait Patterns using Ground Reaction Forces and Machine Learning Algorithms*. 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2011.
- [26] M FLASINKI. *Introduction to Artificial Intelligence*. Springer, 2008.
- [27] WS McCULLOCH and W PITTS. *A logical calculus of the ideas immanent in nervous activity*. The bulletin of mathematical biophysics 5.4 (1943): 115-133.
- [28] M SOKOLOVA and G LAPALME. *A systematic analysis of performance measures for classification tasks*. Information processing & management 45.4 (2009): 427-437.
- [29] C MORBIDONI, A CUCCHIARELLI, S FIORETTI and F DI NARDO. *A Deep Learning Approach to EMG-Based Classification of Gait Phases during Level Ground Walkings*. Electronics 8.8 (2019): 894.