POLITECNICO DI TORINO

Corso di Laurea in Ingegneria Biomedica

Tesi di Laurea Magistrale

A study on sleep parameters in Parkinson's Disease patients

Analysis on EEG and inertial signals in patients and controls

Relatore
Prof.ssa Gabriella Olmo

Correlatore:
Ing. Irene Rechichi

Laureandi
Camilla Cervetto
Sara Mangione

Marzo 2020

Contents

List of Figures

List of Tables

Abstract

Parkinson's Disease (PD) is a neurodegenerative disorder that affects a considerable number of people all over the world. Symptoms are the results of the dopaminergic neurons death in the substantia nigra pars compacta and the dopamine lack leads to the classical parkinsonian movements such as tremor, irregular gait, paralysis and a low muscular strength.
Actually the disease is associated also to other symptoms that are not specifically involved in the motor system and can precede the movements disorder. Non-motor features are: olfactory dysfunction, fatigue, pain, psychiatric and sleep disorders; they have an high impact on the quality of life.
This study is focused on the analysis of the electroencephalogram signals (EEG) and the inertial signals, both of them collected during the night. Regarding EEG, three algorithms are implemented: one for distinguishing the REM phases from the others and two for distinguishing all the sleep stages. Features extraction and machine learning techniques are employed, reaching an accuracy of 98.7%, regarding the REM-NREM differentiation, and an accuracy of 93.1% for all the stages recognition. Wavelet transform is used in the last algorithm for all stages identification; an accuracy of 93.5% is obtained.
The second part is centred on the comparison between PD patient's and healthy subjects' inertial signals. The effect of PD on sleep is undeniable: the patient performs less and slower movements with respect to controls. For evaluating these differences an activity index is calculated; moreover other parameters are extracted: the number of turning in bed, the angular velocity and the turning duration.
Since the EEG data come from an existing on-line database, the future implementation of this work is to associate EEG and inertial data from the same patient, in order to make an evaluation about his quantity of motion in relation to the disease status.
Finally, two user interfaces are proposed for assuring an easier employment of this work.

Chapter 1

General Introduction

1.1 Introduction

This Master thesis is focused on the study of sleep parameters in healthy subjects and in a patient affected by Parkinson's Disease.
Parkison's Disease (PD) is one of the most spread diseases all over the world: it affects 1% of the global population over 60 years old, not taking into account the juvenile cases [1]. PD is often recognizable by motor impairments such as tremor, but there are also non-motor issues that are relevant as well. Among non-motor symptoms there are sleep-related disorders that have an high impact on the patients' quality of life. Moreover these symptoms appear in the early stages of the disease and an immediate recognition could be crucial for treating and slowing down the course of the illness.
For this reason, the research has centred on sleep parameters in order to give our contribute to achieve a better understanding of the PD dynamics.
Sleeping is fundamental in everybody's life; therefore, the approach begins with a study of brainwaves during sleep. In particular, a method for distinguishing the sleep stages through just one electroencephalogram (EEG) channel has been implemented. In addition to the EEG, data were collected via an inertial sensor that recorded body accelerations and angular velocities. This device allowed to make a comparison between a healthy subject's sleep and a patient's sleep; it was studied how much a disease as PD actually affects a peaceful moment like sleep.
The choice of considering just one channel derives from the prospective of a future applications in which both the electroencephalogram and inertial signals will be collected from patients. A single channel is more comfortable to keep all night long.

Finally, it is hoped that the integration between signals will monitor patients effectively, in order to give an idea about the quality of sleep in relation to their disorders.

The thesis is organized in four chapters: the first chapter provides a general view on sleep and PD through a physiology excursus. The state of art section includes an explanation about the monitoring sleep current procedures but also some new techniques.
In the second chapter details about materials and a focus on how the proposed algorithms work are introduced.
The results appear in the third chapter: each algorithm outcomes are presented.
At the end, future developments and conclusions, in which we review the entire work, are introduced. The acronyms list and the bibliography can be found at the end of entire work.

1.2 Sleep physiology

Sleep is an almost unconscious state which is the complementary of the awake state; it takes up about one third of a person's life. During sleep, sensory activity is inhibited, muscular tone is reduced and the interaction with the external environment is nearly suppressed.
The main regulators of sleep are the brainstem and the hypothalamus: the first one maintains wakefulness through the reticular activating system and the sensory pathway activity [3], while hypothalamus preoptic nucleus are supposed to initiate sleep through noradrenergic, cholinergic and serotonergic systems. Other areas give a contribute to sleep modulation, for instance the posterior hypotalamus with the histaminergic system and the basal forebrain with the cholinergic system.
When active, all of these systems promote wakefulness while their inhibition promotes sleep. [5]

1.2.1 Sleep stages

The wake-sleep cycle is regulated by circadian rhythm and homeostatic processes; in specific, the circadian rhythm determines the temporal distribution of these two states while the homeostatic process determines the necessity and

the proportion of sleep in relation to the previous wake period. In general, a circadian cycle is an internal body clock that marks several mechanisms, such as hormones production or body temperature and also wake-sleep sequence by referring to light-dark cycle.
During sleep there is a succession of non-REM (non-rapid eye movement) and REM (rapid eye movement) phases that results to be extremely different. Non-REM sleep is characterized by a decrease in body temperature and in heart rate; moreover the brain is consuming less energy than during the wake state.
REM sleep is a limited portion of the total sleep time: in this phase, there is the dream generation associated with rapid brain waves, ocular movement and a complete loss of muscular tone.
It is possible to divide sleep in stages that are deeply related to the brain waves within it as can be seen from the figure 1.1.

- *Wake*: when the subject is relaxed with closed eyes *alpha rhythm* is present. The brainwaves are "sinusoids" with a bandwidth between 8 and 12 Hertz. In these there are both slow and rapid eye movement and the muscular tone is still present.

- *Stage 1*: normally it lasts a few minutes; the alpha rhythm decreases and brainwaves in $3-7$ Hz bandwidth (*theta rhythm*) slowly take place. Eye movements are still present but now are slow and circular. As regard the muscle tone, the activity is present but in a lower level than during the wake stage.

- *Stage 2*: brainwaves activity is mainly represented by the theta rhythm. During this stage, K-complex waveforms and sleep spindles are recognizable. A K-complex is a high voltage waveform (more than 100 μV) that can be seen in the electroencephalogram (EEG) and it seems to perform two important aims: the suppression of cortical excitation as a result of stimuli that are not a danger signal and the promotion of sleep-based memory consolidation.
  Sleep spindles are bursts of activity which last around one seconds with high frequencies, that are useful to inhibit unnecessary information elaboration.

- *Stage 3*: during this phase a *delta rhythm* is reached; the frequencies are included in the $0.5-4$ Hz band and the brainwaves amplitude is over 75 μV.

Muscular tone is slightly decreased and there are almost no eye movements. K-complex are sometimes distinguishable and sleep spindles can occur.

- *Stage 4*: the subject is experiencing a deep sleep with brainwaves mainly in the delta rhythm.
  Eye movement are absent and the muscle tonic activation is very low. The metabolic brain activity is minimized.

- *REM stage*: this stage is characterized by low voltage desynchronized brainwaves and mixed frequencies; moreover this phase is combined with an increase in heart rate, blood pressure and breathing rate. Due to this similarities to the wake stage, REM sleep is also named the "paradoxical sleep".
  In this stage the body loses muscles tone completely and this event is called "REM atonia". The beginning of the REM stage can be recognized from PGO waves (ponto-geniculo-occipital waves) since it is demonstrated they precede other signals of REM sleep by 30-90 seconds [2]. PGO waves are related to the act of dreaming; the signal to noise ratio is greater in pre-REM epochs with respect to the REM ones, and it seems to be the reason of the vivid imagery of that period.
  Moreover these waves probably have a role in the structural growth and brain development. [2]
  In addition, it has been suggested that the brain uses the REM sleep to wake itself up after experiencing a sufficient time in NREM sleep. In fact, as the night progresses, the REM periods appear more frequently; the awakening occurs usually after a REM period.

Generally NREM sleep is believed to be the period during which the brain recovers itself to the prior wakefulness state. There is a correlation between the amount of delta waves and the period of wakefulness that has preceded it. Some studies suggested that the NREM waves are not only related to the duration of wakefulness but also to the quality; when the subject experiences stress there is an increase in the need of sleep [6].

1.2.2 Function of sleep

There are different theories regarding the function of sleep. One of the most common theories concerns the energy conservation: during sleep the biological functions are reduced so that energy can be conserved and used in the

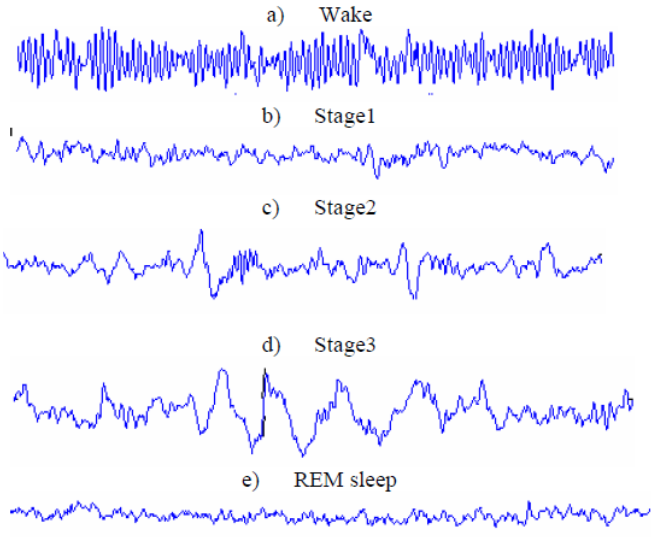


Figure 1.1: EEG sleep stages

awake period. Actually the amount of gathered energy is negligible and it suggests that this is not the primary function of sleep. Moreover the presence of REM sleep, in which the metabolic rate increases is in contradiction to this theory.
Another theory is about the energy allocation based on the evolutionary principle by which organisms have evolved to allocate energy in basic functions such as growth, maintenance and reproduction in order to maximize the reproductive output [8]. According to this, there are three allocation strategies: sleep-wake cycle, torpor and continuous wakefulness. During the sleep state, energy is reallocated in order to be used in the biological process during the next period of wake.
It is possible to see sleep also as a process to stimulate the cellular repair caused by the metabolic processes. In fact, it is proven that a consistent number of genes changes their expression during sleep and some sleep-associated proteins are involved in intracellular transport and endo/exocytosis [29]. This hypothesis can be extended to all organisms and tissues and it seems to be helpful to fix cellular changes that occur during the wake stage.
Moreover some evidences suggest a decisive role of sleep in the learning and memory process; sleep could be related to brain plasticity and it is helpful to improve abilities in learning and remembering. Both REM and NREM seem to be dedicated to different types of memories [10]. In fact, memory activated in performing a task is improved and settled in REM sleep; the one

which involves facts and ideas is improved during NREM sleep [11].
In conclusion, the theories briefly explained above, suggest the absolute necessity of sleep: sleep duration is related to many issues as strokes [12], cardiovascular diseases [13], obesity [14] and carcinogenesis [15].
Therefore it is necessary to study and diagnose possible sleep diseases through specific medical tests and procedures.
It is proved that sleeping a proper amounts of hours and eventually improving sleep quality is crucial for preventing diseases and generally for health encouragement in the population.

1.3 State of art

As explained previously, in order to monitor the entire night and the stages succession, a few parameters should be considered together for assigning the proper sleep stage. For assessing that, a specific examination is performed: the *polysomnography* (PSG). The PSG is a multiparametric exam in which patients are continuously monitored overnight, physiological data are collected and studied for making a diagnosis.
PSG collects the electroencephalogram (EEG) signal, in order to follow the brainwaves evolution, electrooculogram (EOG) signal, for understanding eyes movements, electromyogram (EMG) signal, for measuring the skeleton muscle activity overnight and electrocardiogram (ECG), for recording heart rate.
After the discovery of the sleep disorder *sleep apnea*, in which patients stop breathing for a while, breathing functions, respiratory airflow, and respiratory effort indicators were added to PSG, plus a peripheral pulsi oximetry. This exam is executed from technicians or specialized personnel.
Information as the sleep onset latency (SOL)(which is the time duration to accomplish the transition from full wakefulness to sleep), the number of awakenings during night, total sleep time (TST) and many other features are directly extrapolated from the PSG data and they are determinant for many diagnostics. Other information such as movements during night or respiration, that are not directly related to sleep, are often used too, according to the patients' need.
Polysomnography is mainly used to find out or verify sleep disorders, as instance: narcolepsy, periodic limb movement disorder (PLMD), REM behavior disorder (RBD), parasomnias, and sleep apnea.

The PSG exam requires at least twelve channels with wires directly attached to the patients; a minimum of three EEG channels are requested, then one or two channels for recording airflow, chin muscle tone, one or more for measuring leg movements, two EOG channels, one or two ECG electrodes and one for the oxygen saturation. Moreover, there are two channels applied on the belts that typically measure, through piezoelectric sensors, chest and upper abdominal movements. All of these channels are wired from the patients to a central box that is connected to a computer system which records and stores data and finally shows them on a display. Some laboratories support also a video camera thus it is possible to observe the patients during the examination.

1.3.1 EEG

More in specific, as regards the EEG recording, normally six exploring electrodes are placed in the frontal, central and occipital portions of the scalp and two references electrodes, placed according to the international 10-20 system (Figure 2.2); these channels record the brain activity (which is the electrical field produced by cortical pyramidal neurons) that will be scored into the sleep stages, as explained above. Usually a conductive gel is interposed between the electrode and the scalp. The EEG signal is then divided in 30 second epochs and each epoch is scored following specific scoring rules. The main scoring methods are the *Rechtschaffen & Kales* (R&K) and the guidelines from the American Academy of Sleep Medicine (AASM).
R&K is the first widely used sleep scoring manual and for approximately forty years it was the only accepted rule. It divides sleep into seven discrete stages: wake, stage 1 (S1), stage 2 (S2), stage 3 (S3), stage 4 (S4), stage REM and movement time [16]. The R&K scoring was then criticized because of its subjectivity that consequently can results in different interpretations. In addition to this problem, this scale was developed for young healthy adults and its application to patients or elderly people could be uncertain. The American Academy of Sleep Medicine suggested a new guideline by modifying terminology, re-coding methods and scoring. Regarding the terminology changes, sleep is divided in four main stages: N1, N2, N3 and REM (R). N3 stage includes S3 and S4 from the R&K and represents the slow wave sleep and the movement time stage is abolished.
In summary, the new AASM guideline simplifies the R&K technique in the scoring system as well as the whole PSG setting.

1.3.2 EOG

The electrooculogram consists of two electrodes placed one 1 *cm* above the outer canthus of the right eye (ROC) and the other one is placed 1 *cm* below the outer canthus of the left eye (LOC). In fact during the eye movement there is a change in electrical potential: the cornea, positively charged, moves toward one electrode while the retina, negatively charged, takes away, causing a electropotential difference. When the eye is still, the change in position is zero and there is no output signal. During the first sleep stages there is a slow and circular eye movement that leads to long waves in the output signals while in REM stage the signal is characterized by bursts that identify this stage [17]. Blinking produces a rapid vertical movement.
In Figure 1.2 there is the representation of the electrodes positioning.

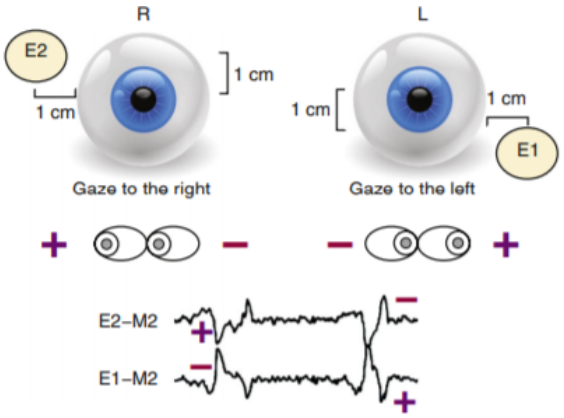


Figure 1.2: E1 is the LOC previously explained, E2 is the ROC, M2 is the right mastoid electrode location [17]

1.3.3 EMG

EMG is recorded in order to measure the muscle tension and to assess the sleep stages, considering that during sleep, movement is reduced. Especially during REM sleep there is no movement at all, if the subject is healthy, otherwise it can be symptom of disease.
Two electrodes are normally placed on the chin, one above and the other below the jawline; other two electrodes are placed on the anterior tibialis in order to detect leg movements.

Sleep leads to relaxation, therefore it is expected the subject to have a muscle tone that decreases during the night. The EMG channel measures the total atonia during REM stage.

1.3.4 ECG

The electrocardiogram is normally made of ten electrodes but during PSG only two or three electrodes are used. The aim is to monitor the electrical signal produced by the heart when it pulses: three main waves are necessary to be distinguished, the P wave, the QRS complex and the T wave.
The study of the heart behaviour during sleep is necessary to assess the sleep stage in conjunction with the other signals and to evaluate heart pathologies.

1.3.5 Respiratory airflow and blood oxygenation

The respiratory flow, nasal and oral, is measured with pressure transducers and thermocouples. Breathing rate is another characteristic useful to confirm the sleep stage but the main application is in supervising and revealing sleep apnea episodes. In addition, belts can be employed for measuring breathing rate by exploiting the chest expansion.
Pulsioximetry detects blood oxygenation and it is directly related to respiration issues. The pulse oximeter fits on an earlobe or over the fingertip.

1.3.6 Procedure

For executing the PSG the patient should go to the hospital or to a specific clinic where a technician apply the electrodes and wires. The personnel in charge will monitor the patient overnight through monitors that show the recorded signals.
Some important feature are then extracted after the scoring procedure:

- *Sleep stages*: as explained above, the sleep stages are fundamental to understand the phases the patient has gone through. The diagram that represents the stages is called *hypnogram*.
- *Sleep efficiency*: it is calculated by dividing the number of minutes of sleep by the number of minutes in bed.

- *Sleep onset latency:* it is the onset of sleep from time the lights were turned off.
- *Arousals:* the arousals are unexpected change in the EEG. This episodes could be caused by different factors, for instance, leg movements, external noises or breathing abnormalities.
- *Breathing and cardiac abnormalities*
- *Leg movements*

1.3.7 Other techniques

Despite its undisputed utility in the sleep diseases field, the PSG technique has some drawbacks, such as its high cost, its short recording time (one or two nights, while daytime recording is complicated) and the amount of cables. For all these reasons alternative techniques to PSG are sought. In particular, non-traditional modalities for instance actigraphy, audio, video and temperature recording are becoming increasingly important: the goal is to achieve a comparable results to PSG without its complications.

- Audio: Audio recording is a useful method for monitoring sleep. It is not invasive because the microphone does not touch the patient: this allows not to disturb the patient's natural sleep. It is also not expensive. Audio recordings are used to identify snoring, normal breathing or obstructive events.
  The study of sleep by audio is carried out using speech analysis techniques: in order to generate the elements of the speech, the vocal tract changes form; in particular the created sounds are shaped in the frequency domain by the frequency response of the vocal tract itself [19].
  This kind of change occurs less noticeably in snoring; therefore a frequency analysis, useful to identify sleep disorders, such as the Obstructive Sleep Apnea (OSA), can be carried out.
- Video: by recording the patients it is possible to monitor a lot of parameters without disturbing the subject. The video recording is often used in combination to PSG for understanding patients disorders. In fact it is possible to observe patients' position, respiratory and body movements; moreover the cost associated to it is not excessive, even if a specialist has to watch the video recordings.
  Limb movements and apneas event can be confirmed with this technique.

One drawback is that often the patient is covered by bed sheet; therefore most of the body is obscured and the light level is normally low. In order to solve this problem, infrared lights or infra-red-sensitive cameras can be employed [19].

- Actigraphy: also known as actimetry or accelerometry, it is a non-invasive method exploiting wearable sensors. Actigraphy is easy-to-use and has been validated for estimating parameters of night-time sleep and circadian rythms.
  The actigraph allows to measure accelerations and subjects' movements. Moreover, it allows to determine if it is a phase of waking or sleep. In the simplest view, sleep is assigned when the sensor does not detect any movement.
  The use of actigraphy is preferred to PSG in case of long-term sleep/wake monitoring or in some specific situations, such as with children when EEG waves are not fully formed yet. However, accelerometry is less specific (less than 50%) than PSG in identifying the waking state.
  The actigraph has multiple fields of application in addition to the determination of sleep-wake phases and sleep stages; examples can be found in the study of physical activity, in the sport context or in rehabilitation. Moreover, it is often used in Parkinson's disease studies, especially for detecting episodes of akinesias and hypokinesias.

- Temperature: body temperature changes during the course of the night, therefore it can be a good marker to be monitored. In fact, temperature is regulated by circadian systems as well as sleep. The core body temperature (CBT), the distal skin temperature (DIST) and the proximal skin temperature (PROX) are often used as indicators. Together with the termo-regulatory model and with the melatonin hormon influence, sleep can be monitored.
  CBT trend decreases during night and increases during arousals; this value is correlated to PROX and it is anti-phase with DIST. Surely the sensor placement has a role in the obtained values.

During PSG, the environment in which the patient has to sleep in, is often felt as uncomfortable and unfamiliar, leading to problems in recording reliable data during the test. This is known as the *first night effect* (FNE) [18]. As a results the sleep onset latency is longer and the sleep efficiency is lower. On the other hand, there is the possibility that patients could sleep better than

usual, resulting in the reverse first night effect (RFNE). A solution can be found in the "serial night PSG".

1.4 Parkinson's disease

Parkinson's disease (PD) is a neurodegenerative disorder that is rather common in our society nowadays. The first description dates back to two centuries ago when James Parkinson published an essay where he reported cases of this disease, describing patients with a continuous tremor, irregular gait, paralysis and a low muscular strength; the complete picture of this disorder is still evolving today.
Symptoms are the results of the death of dopaminergic neurons in the substantia nigra pars compacta (SNpc) (Figure 1.3); in fact, the dopamine lack leads to the classical parkinsonian movements [20]. Actually the disease is associated to other symptoms that are not specifically involved in the motor system and can precede the movements disorder by more than a decade.

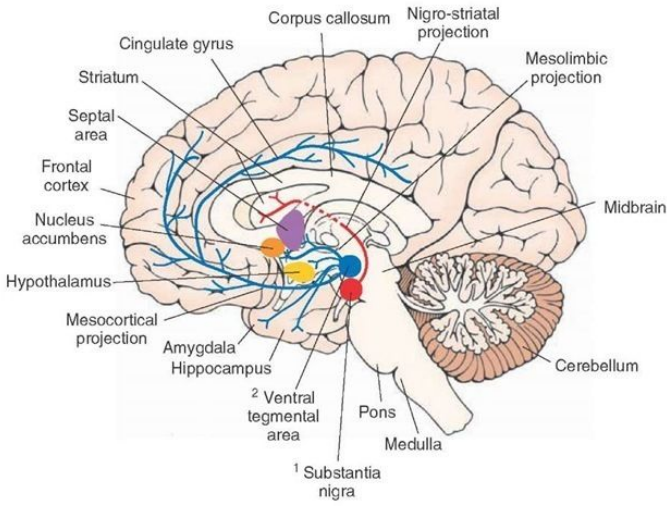


Figure 1.3: Brain anatomy

The classical motor dysfunctions, as described by Parkinson, are mainly: bradykinesia, muscle rigidity, rest tremor, and gait and postural impairment [20]. It is important to highlight that not all the patients show the same characteristics; therefore, two subclasses of the disease has been described with empirical observations: tremor-dominant Parkinson's disease and non

tremor-dominant Parkinson's disease. In the first category, subjects show mainly tremor without other symptoms, with a slower progression and a lesser disability [22], while in the second category akinetic-rigid syndrome and postural instability gait disorder are included [20].

Non-motor characteristics are: olfactory dysfunction, psychiatric symptoms, fatigue, pain and sleep disorders; they have an high impact on the quality of life [23]. The pathogenic pathway that leads to PD is thought to be in progress during the pre-motor phase involving the nervous system and the SNpc; this period can be used for adjusting the therapy in order to get the progression of the disease slower. In fact, over the course of time there is a worsening of the motor skills that can be managed with long term therapies that have drawbacks too. Patients in the late-stage become medications-resistant and motor and non-motor features are noticeable, for instance freezing of gait (FOG), speech dysfunctions and dyskinesia. In Figure 1.4 there is a scheme of typical symptoms evolution.

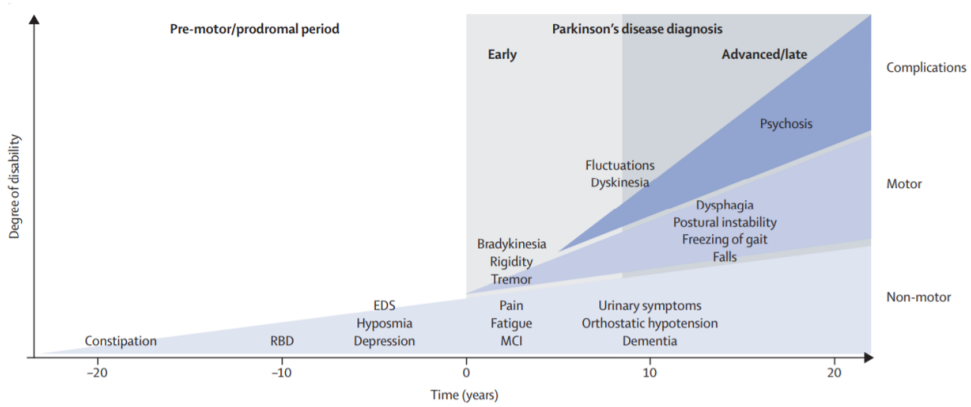


Figure 1.4: Symptoms of Parkinson's disease over years [20]

As anticipated, PD main features are due to the dopaminergic neurons loss in the SNcp, specifically in the ventrolateral tier, which includes neurons that project to the dorsal putamen of the striatum [20], but also in the nucleus basalis of Meynert, amygdala, dorsal motor nucleus of the vagus and in the pedunculopontine nucleus [24]. The microscopical alterations can be observed in the melanin pigmentation in the substantia nigra and in the other regions.

Another distinctive characteristic of PD is the Lewy pathology, which provides for a anomalous aggregation of a protein, the $a-$synuclein. In fact, it has been discovered a mutation in the $a-$synuclein gene, SNCA, that causes a mendelian form of this illness [25]. This protein, in a misfolded state, is insoluble and it accumulates in neurons provoking inclusions called "Lewy bodies".
Lewy pathology is suggested to progress always in the same manner; Braak proposed a PD classification based on the pathological picture: Lewis bodies appear first in the peripheral nervous system and in the olfactory system, then, with the progress of the disease, the aggregations spread in the central nervous system [26]. In addition, another feature is the neuroinflammation, which is mediated by resident astrocytes and microglia designated to the clearance of the extracellular detritus [27]. When the microglia is activated, there is a trophic factors release, but also dangerous reactive nitrogen and oxygen species and cytokines are produced. This process is still under study in order to understand the real effect on neurons [28].

1.4.1 Risks factors

During the last decade researches has been done on the genetics of Parkinson's disease. The first gene related to the illness is the SNCA, that codifies $a-$synuclein: mutation to SNCA lead to autosomal dominant parkinsonism. Other genes involved are parkina (PRKN), dardarin (LRRK2), PTEN-induced kinase 1 (PINK1), DJ-1 e ATP13A2 [29].
Several risks factors has been proposed through the years:

- Gender: there is a ratio male-to-female of about 3:2 [30].
- Ethnicity: In the USA the most affected are the Hispanic people, followed by non-Hispanic Whites, Asians and Blacks [31].
- Heredity: Having many relatives with Parkinson's disease increases the chances of a genetic disease. Generally, juvenile (atypical) cases have a genetic origin.
- Age: the incidence of PD is higher as the age increases [32].
- Environmental exposure: it has been demonstrated that insecticides and herbicides have a role in the PD risk [33].

Although there are also factors related to a decreased risk, as instance:

- Tobacco smoking [34]
- Coffee drinking [34]
- Calcium channel blocker use [34]

Therefore there is an interaction between genes and environmental factors that should be taken into account as shown in Figure 1.5.

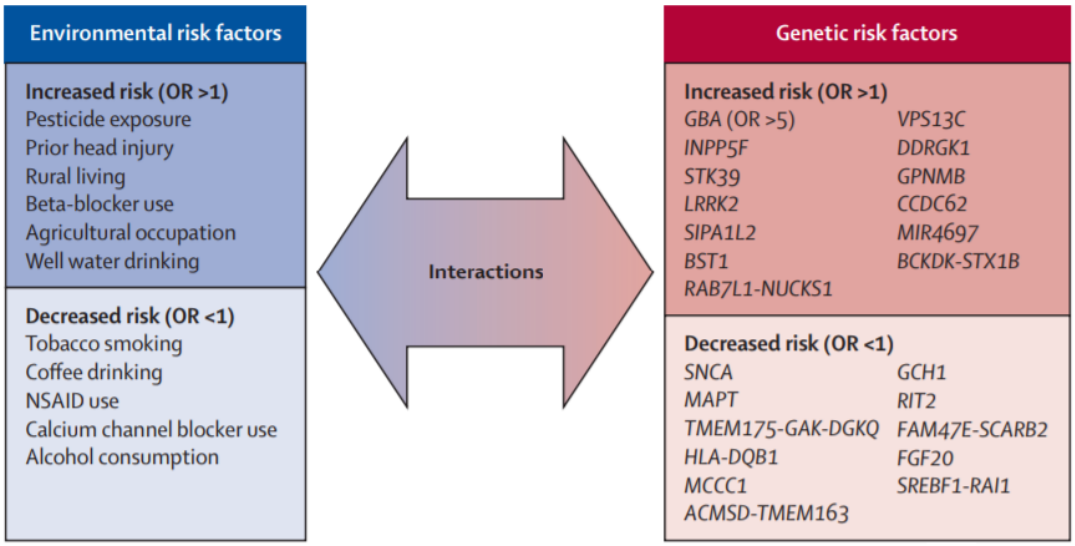


Figure 1.5: Interaction between risks factors, where OR is the odd ratio [20]

1.4.2 Diagnosis

Regarding the Parkinson's disease diagnosis, the presence of the main feature of the illness, such as bradykinesia, rigidity and tremor, are clinically evaluated; in any case the gold standard is the neuropathological assessment. The UK Parkinson's Disease Society Brain Bank Clinical Diagnostic Criteria are the most broadly used, but recently a new set of diagnostic criteria was proposed by the International Parkinson and Movement Disorder Society (MDS) in order to improve the diagnostic accuracy.
Tomographic cerebral scansion can be used to discriminate and possibly exclude pathologies with similar symptoms. Positron emission tomography (PET) or single photon emission computed tomography (SPECT) are used

to estimate the reduction of the dopaminergic terminals; one of the advantages of this methods is the non-invasiveness, on the other hand they are not specific in distinguishing from other disorders associated with SNpc neurodegeneration.
Different biomarkers has been identified for investigation and they can be divided in clinical, pathological, imaging, genetic and biochemical. In Figure 1.6 some of them are depicted; the combination is likely to be useful for a precise diagnosis. The Lewis body observation during autopsy is considered the final proof that the patient suffered from Parkinson's Disease.

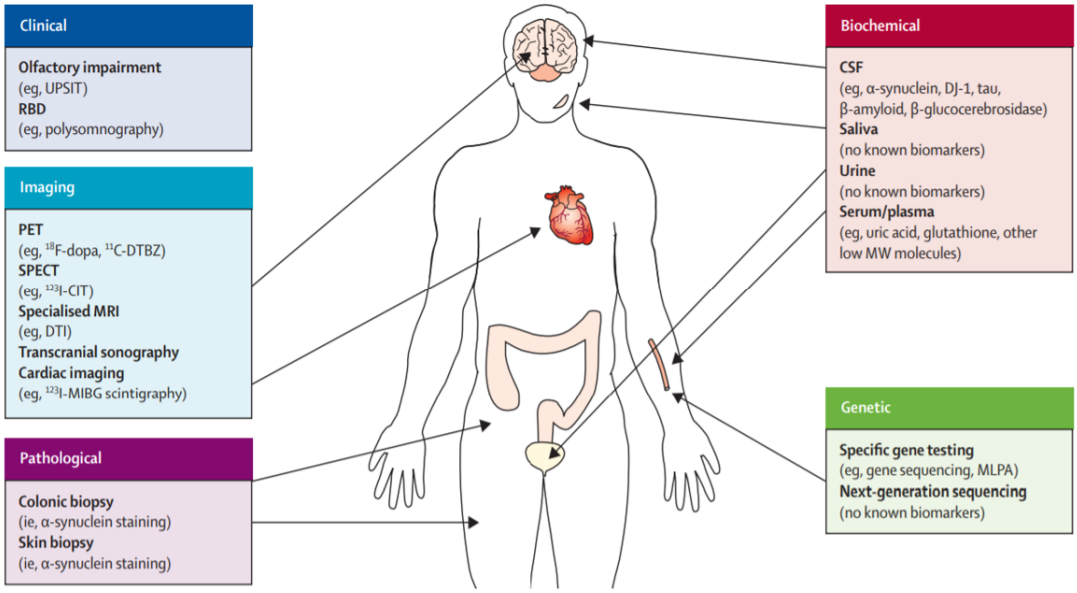


Figure 1.6: Biomarkers useful for parkinson's disease diagnosis [20]

1.4.3 Medications

At the moment, a definitive cure does not exist, pharmacological treatments and surgery can only alleviate symptoms.
Since the dopamine loss causes motor symptoms, the neurotransmitter should be re-integrated. Dopamine does not cross the blood-brain barrier, but a precursor of dopamine, *levodopa*, can. Levodopa, after going through the barrier, is converted in dopamine and it reduces for a short period motor symptoms. Not all the levodopa crosses the barrier, most of it is metabolized elsewhere and it causes some side effects like orthostatic hypotension and nausea.

Other medications can be used in combination to levodopa: *carbidopa* and *benserazide*; they are dopa-decarboxylase inhibitors that reduce the conversion of levodopa to dopamine outside the brain.
In the long term, the use of these medications can cause dyskinesias and fluctuations in the efficacy of treatments; over the course of time the body get used to this medications, consequently making them less effective; this issue is called "wearing off".
Alternatives to levodopa are some dopamine agonists which bind to the dopamine receptors [35]; Initially these medications were used as a therapy for patients that showed levodopa complications, but right now are more used as therapy before starting levodopa in order to delay the use of this latter and therefore its side-effects. Some of the dopamine agonists are *pergolide*, *ropinirole*, *bromocriptine*, *cabergoline* and *apomorphine*.
As regards surgery, thanks to the great improvements in this field, PD patients whom medications are no more sufficient (or they experience mainly dyskinesia and motor fluctuations) undergo to an operation. There are two: deep brain stimulation (DBS) and lesional surgery. DBS is used in patients that suffer from the disease in a moderate to a severe state. The targeted area for both the surgical techniques is the subthalamic nucleus or the globus pallidus [20]. In the deep brain stimulation a neurostimulator is implanted: it is a medical device that sends electrical impulses to the zone in which is placed. It can significantly improve quality of life in Parkinson's patients by blocking the abnormal nerve signals that cause the main symptoms such as tremors, rigidity, bradykinesia, and postural instability. Especially resting tremor is suppressed within seconds, while relief from bradykinesia takes minutes, improvements in gait and posture can require continuous stimulation from hours to days before achieving a maximum effect.
The DBS implantation occurs on average 13 years after the beginning of the disease. It would be optimal to implant the device at an earlier disease-stage to prevent psychosocial decline and to maintain quality of life for a longer period of time[21]. In Figure 1.7 there is an illustration of DBS positioning.

1.4.4 Disease evaluation

It is fundamental to evaluate the disease stage of progress for every single patient. For this reason a specific scale has been introduced: the *unified Parkinson's disease rating scale* (UPDRS).
It has fours parts as reported in the table 1.1.

A revised version of UPDRS named MDS-UPDRS was then introduced:

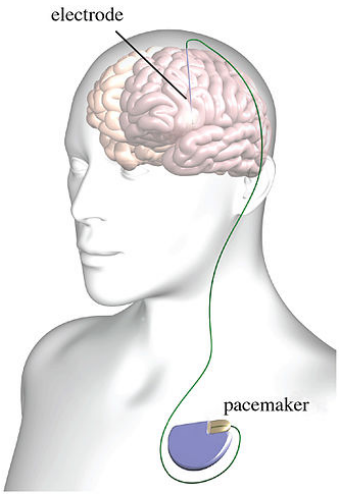


Figure 1.7: Electrode and device used for deep brain stimulation [37]

UPDRS	
UPDRS-I	Regards the nonmotor aspects of daily life
UPDRS-II	Regards the motor aspects of daily life
UPDRS-III	Motor subscale: concretely assigns specific parkinsonian marks, such as tremor, rigidity, bradykinesia and gait impairments.
UPDRS-IV	Regards motor complications.

Table 1.1: UPDRS stages

it still maintains the same 4 parts structure of UPDRS, but it solves some problems identified in the original version such as some PD-related issues that were poorly expressed [38].
The Hoen and Yahr scale is another rating scale that has five stages related to which part of the body the disease is mostly affecting, at which level and if it affects the balance too.
Moreover, Schwab and England scale is procedure for assessing abilities of impaired patients; it exploits percentages in order to evaluate the complications patients experience for completing daily life activity and if they need support for carrying out them.

All these criteria refer to the main PD manifestations as a result of nigrostriatal damage, such as bradykinesia, tremor and rigidity. PD diagnosis is generally made when these symptoms begin to appear, during the so called

motor stage.
Bradykinesia occurs when one body part is affected by a progressive decrease in speed and amplitude of movements. This is the primary consequence related to dopamine decrease and what specialists seek to quantify in order to diagnose parkinsonism and, as a consequence, optimize the therapy.
Typical parkinsonian tremor is an intermittent oscillatory movement which is mainly present on the limbs. Generally, the rest tremor (frequency 4-6 Hz) is inhibited by movement and may reappear a few seconds after a position change. On the contrary to what is commonly thought, not all PD patients exhibit tremor since the disease manifests itself more in other ways (bradykinesia or gait problems).
Rigidity is the resistance of any body part to the passive mobilization because of an increased tone. Rigidity can affect head, trunk and lower or upper limbs.

1.4.5 Parkinson's disease sleep disorders

As already introduced, PD patients suffer from non-motor disorders that have a high impact on everyday life and reduce life quality.
These disorders include for instance *insomnia*, *restless leg syndrome* (RLS), *excessive daytime sleepiness* (EDS), *rapid eye movement sleep behaviour* (RBD) and *sleep fragmentation*. The majority of sleep issues are experienced in the late stages of PD but RDB and EDS can be observed also in the early stage, and even in the premotor phase [39]. There are three main categories that explain the abnormal behaviour during sleep that can happen during day and night, as shown in Table 1.2

Category	**Sleep issue**
Parasomnias	REM parasomnias (as instance RBD) NREM parasomnias (as instance sleepwalking)
Sleep difficulties	Insomnia (initial, maintenance and terminal insomnia)
Sleepiness	EDS

Table 1.2: Main sleep disorders in Parkinson's disease

Parasomnias are undesired behaviours that happen not only the first

phases of sleep, but also during the awakening (REM and NREM).

Rapid eye movement sleep disorder

This disorder is very common in PD patients, affecting a percentage from 20 to 72% [40]. Normally during REM sleep muscles tone is almost suppressed. People who suffer of RBD have dream-related vocalizations and are able to move; they shout, talk, scream, have complex motor movements like kicking or punching. Surely this is a problem for themselves, because they can be walking into injuries and for the people who sleep with the patients. RBD often appears years before the parkinsonian motor symptoms. It is related to some feature like gender, cognition, age, disease severity, disease duration and medications.
The disorder diagnosis is performed with a clinical manifestations or questionnaire: however, a history of vocalizations and motor behaviours in REM sleep is relevant for diagnosing. The PSG can measure an abnormal chin muscle tone in the EMG during the REM phase but it is not mandatory for the clinical diagnosis.
Regarding the pathophysiology, RBD is related to the structures that provide the atonia by inhibiting the motor neurons of the spinal cord, the pedunculopontine nucleus (PPN) and the medullary magnocellular reticular formation [40]; the degeneration of the neural structures makes difficult to control the tone.
Usually two kinds of medication are used in this cases: melatonin and clonazepam. The first one is commonly used in patients that suffer from mental impairment or sleep apnea. The mechanism by which melatonin works is still not clear; it has some side effects though, like headache or daytime excessive sleepiness. On the other hand clonazepam can worsen sleep apnea symptoms and mental disorders.

Insomnia

Insomnia can be seen as a difficulty in beginning sleep, in maintaining it or in early awakenings. Insomnia can occur alone or in association to mental or systemic illnesses and it is related to female gender and the PD duration. Sleep fragmentation happens when the patient cannot maintain the sleep integrity and he wakes up many times during the night.
The diagnosis of this disorder is derived from the patients' clinical history and it is necessary for making a medication plan. The most used drugs are

hypnotics or sedating anti-depressant, used for maintaining sleep longer [40]. One of the causes of insomnia is the restless leg syndrome: in fact the patients, perceiving an unconformable feeling at legs, need to move them for making it goes away. This continuous movement does not let the patients get to sleep.

Sleep-Related Movement Disorders - RLS and PLMD

Restless legs syndrome (RLS) and periodic limb movement disorders (PLMD) are two sleep-related movement disorders. The first refers to unpleasant or uncomfortable feelings in legs and an urge to move them; while PLMD is characterized by periodic limb movements during sleep. Both disorders are due to dopamine dysfunction and respond well to dopamine medication.

Excessive daytime sleepiness

It is a chronic sleepiness that occurs during the day, and can be the results of different causes such as depression, anxiety, changes in sleeping habits and in the circadian rhythms. Sleepiness is related to the deterioration of the normal night sleep but also to the degeneration of the sleep-wake control centre and to the medications side effects.
A specific scale is used for evaluating the level of EDS: the *Epworth Sleepiness Scale* (ESS). It has statements that can be scored to a maximum of 3 points; the higher is the score, the higher is the sleepiness level [40].

Nocturnal hypokinesia

It is a very common condition in which patients are impaired or totally enabled to move during the night, to turn over in bed or to get out of bed. This could be cause of immobilization that can lead to additional problems, for instance predisposition to aspiration pneumonia, pressure ulcers or asphyxia [50]. In literature this problem has been often pointed out and it is characterized by few episodes of turning in bed that results to be very slow. Sometimes, this disorder can be associated to the wearing off effect but also to an ageing process.

1.5 Purpose of this thesis

As suggested in the previous section, PD is a disabling disease not only involving the motor system. In fact, sleeping is highly affected by disorders, derived from PD, that make it difficult to pursue.
This disorders are often underestimated but they have an important role in the everyday life. The diagnose comes from the PSG exam which is expensive and not always guarantees the desired results. Cables and electrodes do not make the patient's sleep comfortable and relaxed, moreover the unfamiliar environment makes sleep even more complicated.
In PD patient, hypokinetic events are frequent and it is not easy to recognize them from the PSG; patients are not able to move during the night resulting in long periods of immobilization.
This thesis is focused on two main fronts, a study on the EEG signals and the inertial signals processing.
The aim is in finding a simple method for revealing sleeping issues with a compact and easy-to-wear system. With the EEG it is possible to distinguish and characterize every sleep stage, while with inertial signals an amount of movement quantification can be done. Therefore, an idea of the patients' behaviour during the night is obtainable from the combination of these two signals.
In particular, hypokinesia can be easily pointed out because accelerometer and gyroscope data return all the changes in position. Another disorder that the union of the two signals identifies, is RBD; in fact, during the REM stages, muscles are supposed to be atonic but if movements are detected then a disorder could be revealed.
The solution proposed in this study is a simple way for monitoring the patient's sleep without the complications given by a full PSG exam. This additional information extracted by signals can be helpful for doctors to have a complete medical picture. In this thesis we could not join the two signals for testing our idea on patients but we hope in future implementation.

Chapter 2

Materials and methods

2.1 Materials

In this work, both EEG and inertial data has been processed. In particular, the EEG data belonged to an existing dataset, the PhysioNet database, while the inertial data have been collected from a patient and from healthy subjects, through the SensorTile.box.

2.1.1 PhysioNet database

In order to understand how the biological signals behave during the night, PSGs from an existing database have been used. Since most of the papers in literature uses the *Sleep-EDF Database Expanded* database uploaded on the PhysioNet website, it has been employed here too.
This database contains a total of 197 recorded polysomnographies, including EEG signals (from Fpz-Cz and Pz-Oz electrode positions), EOG signal, chin EMG and event markers. Only some of the records have respiration signals and body temperature.
The epochs have been scored by technicians following the Rechtschaffen and Kales rules. In fact, the stages present in the annotation columns are:

- $W \rightarrow Wake$
- $R \rightarrow REMSleep$
- $1 \rightarrow Stage1$
- $2 \rightarrow Stage2$

- $3 \rightarrow Stage3$
- $4 \rightarrow Stage4$
- $M \rightarrow Movement\ time$
- $? \rightarrow Not\ scored$

Signals are in .EDF format while hypnograms appear in the .EDF+ format; they have an header that defines the patients in a anonymous way; in particular only gender and age are codified.
The 197 recordings are named as SCssNEO-PSG.edf, in order: Sleep Cassette (SC), ss which is the patient number and N which is the number of the night.
All files belong to a study dated between 1987-1991 of healthy Caucasians, from 25 to 101 years old and not assuming any sleep-related medication.
The EEG and EOG signals are sampled at 100 Hz; the EMG signal is electronically filtered with an high pass filter, rectified and filtered again with a low-pass filter therefore the envelope is in µV root mean square and sampled at 1 Hz.

2.1.2 SensorTile.box

SensorTile.box is an evaluation system assembled by STMicroelectronics which contains ST MEMS devices, in specific a six axis inertial measurement unit (IMU), a temperature sensor, a magnetometer, a pressure sensor, a microphone and a humidity sensor. It is closed in a rigid plastic casing.
As regards the energy source, this box contains a Li-Ion rechargeable battery: it provides 3.7 V and 500 mAh with a dimension of 35.5x25.5mm. All the sensors are managed thanks to a low power microcontroller STM32L4R9. It exploits a Bluetooth connection for communicating with its app on the smartphone but all the data are stored in an 8 Gb SD card. The app allows the user to create a personalized application by selecting the sensors to be used, the frequencies and full scale range; then it is easily uploaded on the device.
In this work, the used application consists on a desired output with date, time, accelerations, gyroscopes and magnetic field sensor.
The device is equipped with three leds on the front side, a blue one, that start flashing when a Bluetooth connection is available, a red led that flashes when connected to an energy source and a green led that turns out by when

a new firmware is uploaded.
The main structure of this device is in Figure 2.1.

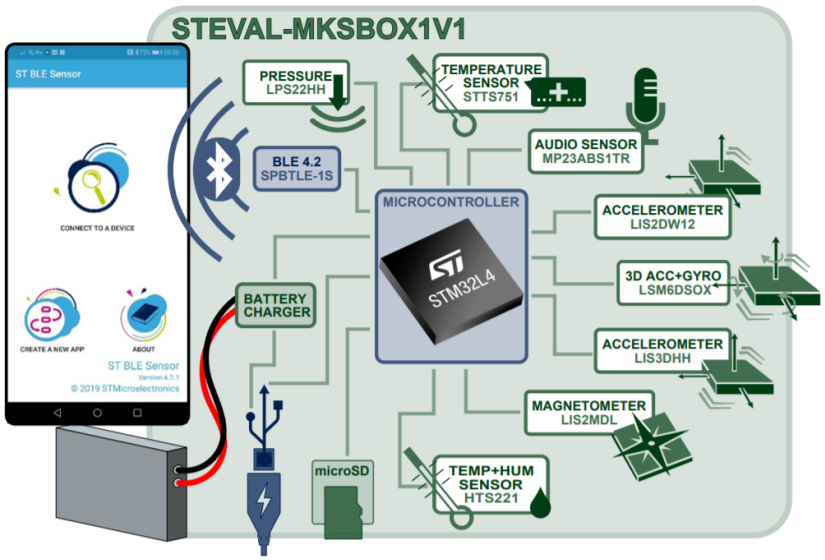


Figure 2.1: Sensortile.box by STMicroelectronics

2.2 Methods

In this section the implemented algorithms will be explained; all the codes has been written in MATLAB R2019b.

2.2.1 Exploring EEG signal

Three main algorithms regarding the EEG signal have been implemented: the first one is a recognition algorithm that can distinguish the REM stages from the NREM ones. It uses specific features extracted from the signal and it reaches promising results. Since the EEG contains a lot of information, an all-stages recognition has been attempted. For exploring this path, two methods have been used: a first one in which specific features have been extracted, and a second one in which the wavelet transform has been implemented. Both of them exploit machine learning techniques.
The Physionet database provides two EEG channels, but for this work only

one channel is used: Fpz-Cz. This channel position (Figure 2.2) is chosen because it is easier to reach and more comfortable for the patient when data are going to be collected.

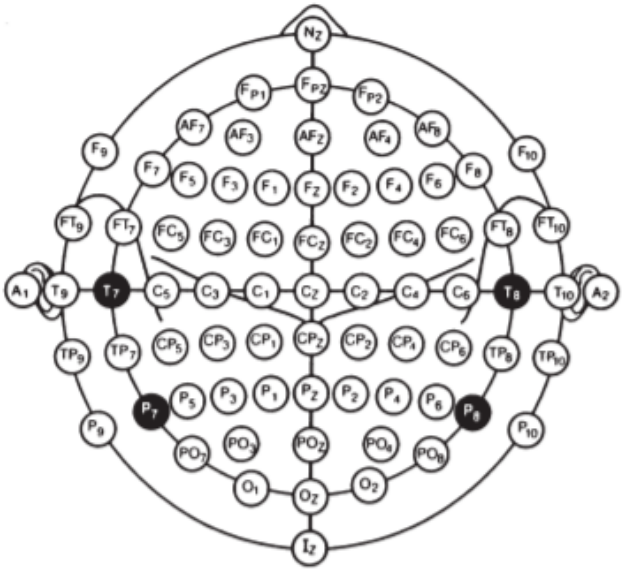


Figure 2.2: Ten-twenty system for EEG electrodes positioning [36]

REM-NREM epochs recognition

In this first phase an algorithm for recognizing REM epochs has been implemented.
Since the data have different extensions, it is necessary to convert the hypnogram files into a compatible extension. In fact, the .EDF+ format is not directly importable, therefore these files are transferred through a free software called *Polyman* and converted into .csv files which are easier to process.
A selection of the total number of patients is considered: 45 out of 197 recordings are taken into account, because not all the data presented suitable characteristics and correct annotations; moreover, since the computational cost is already quite high, by increasing the number of patients it would be unbearable for the supplied personal computers.
The annotation vector has been done following the *R&K* rules, but, in order

to decrease the number of stages to classify, it is changed into the more objective AASM classification. Moreover, since there were not many not-scored epochs, they were not taken into account.
The aim of this algorithm is to extract features that can easily identify the REM phase.
A summary of the method is depicted in Figure 2.3, a flow chart is proposed; then each part is described in detail below.

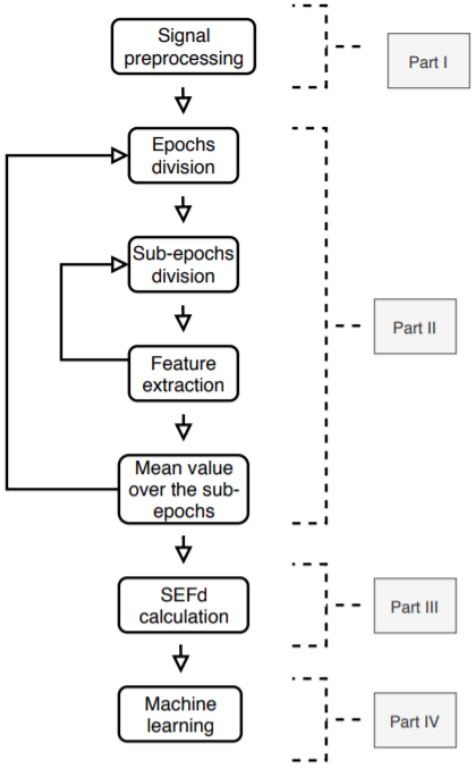
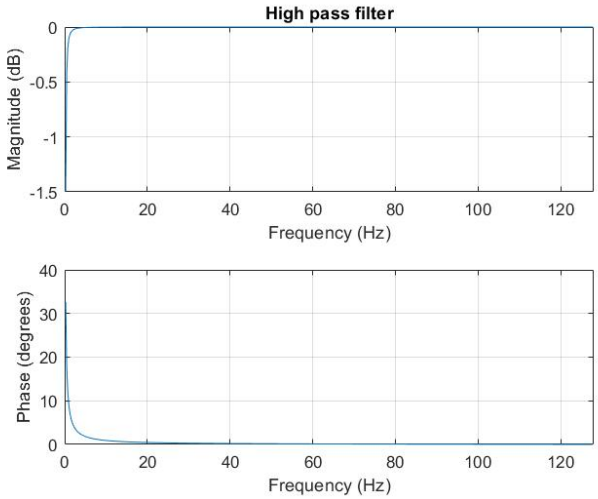


Figure 2.3: General flow chart of the implemented algorithm

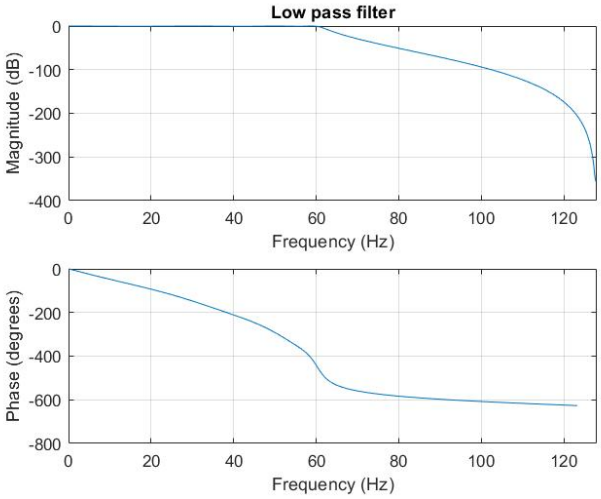
Part I:

After uploading the data in the proper way, a preliminary processing of the signal is performed. The EEG signal is filtered with a pass band filter; the cut-off frequencies are 0.16Hz and 60Hz. The pass band filter has been realized with a combination of a high pass filter (Figure 2.4(a)) and a low

pass filter (Figure 2.4(b)). With these filters the continuous frequency is removed and then the band is limited up to 60Hz, because the EEG content does not extent beyond.
Originally the sampling frequency was 100Hz but the signal has been resampled at 256Hz to improve the accuracy.



(a) High-pass filter



(b) Low pass filter

Figure 2.4: Filters composition for the EEG preprocessing

Most of the recordings in the dataset begins during the afternoon but,

since only the EEG during sleep is necessary, the signal is cut; all wake epochs at the beginning of the signal are removed while the wake epochs at the end are maintained. This step is important also because it helps in balancing the dataset used in the second part of machine learning.

Part II:

The signal is divided into 30 seconds epochs, as the state of art dictates. The epochs are not overlapped.
The difficulty in distinguishing NREM and REM stages is mainly due to the similarities in frequency of N1 and REM, specifically in the bandwidth 13-17Hz [41]. The bandwidth 10-13Hz seems to be helpful for dividing REM and N1 while the bandwidth 12-16Hz is helpful for discriminating REM from other stages. Following the work [41], the range chosen for understanding better the different stages is 8-16Hz, that includes the two bands expressed before.

Every 30 seconds epoch is divided again into 2 seconds epochs, non-overlapped. In order to move from the time domain to the frequency domain, the fast Fourier transform (FFT) is applied:

$$X(f) = \int_{-\infty}^{+\infty} x(t) e^{(-j2\pi t)} dt \tag{2.1}$$

Theoretically the Fourier transform is expressed from $-\infty$ to $+\infty$, but in practical signals are multiplied by a rectangular window that limits in time the signal.
The resolution of this transform is given by:

$$df = \frac{1}{T} \tag{2.2}$$

Where T is the length in time of the segment that is 2s, consequently there is a 0.5Hz resolution.

From this result two main feature are extracted: the spectral edge at 50% (from now on called SEF50), and the spectral edge at 95% (SEF95) in the interval of $8 - 16$Hz.
The SEF is the specific frequency below which a certain amount of signal power is retained; in SEF50 what is looked for is the frequency below which

there is half of the spectral power. This calculation is the same as calculating the median frequency, from the FFT coefficient the frequency is extracted using this method [42]:

$$\sum_{1=1}^{x} |mag_i|^2 = 0.50 \times \sum_{i=1}^{n} |mag_i|^2 \tag{2.3}$$

$$SEF50 = freq(x) \tag{2.4}$$

N is the number of the FFT coefficients while x is the index that it is looked for. The x-th frequency from the FFT vector is required.
Behind this equation there is the assumption that the signal in 2 two seconds window is wide sense stationary (WSS). By doing the quadratic modulus of the Fourier transform the direct method of the spectral power calculation is applied. If the process is WSS and ergodic, which means the process will not change its statistical features in time and these can be deduced from a single, long enough sample, then the following equality holds:

$$P(f) = \lim_{M\to\infty} E\left\{ \left[\frac{1}{(2M+1)T} \right] \left| T \sum_{-M}^{M} x[n] e^{-j2\pi f n T} \right|^2 \right\} \tag{2.5}$$

This means that the spectral power is equal to the statistical average of the quadratic modulus of the discrete time Fourier transform (DTFT) of the windowed signal and divided for the window length when the latter tends to infinity.

SEF95 is defined as:

$$\sum_{1=1}^{x} |mag_i|^2 = 0.95 \times \sum_{i=1}^{n} |mag_i|^2 \tag{2.6}$$

The output is the frequency below which there is the 95% of the spectral power.

Since every 30s epoch is divided into 2s epoch, at the end of every loop 15 measures of SEF50 and SEF95 are obtained; therefore, these values are averaged and then, for every epoch, there will be just one output.
The last extracted feature is the difference between SEF95 and SEF50, called SEFd. For every epoch e, SEFd is calculated as follows:

$$SEFd = \frac{1}{15} \times \sum_{i=1}^{15} (SEF95[se_n] - SEF50[se_n]) \tag{2.7}$$

Where i is the subepoch while n is its index. The signal is then smoothed for flattening its variability.
In Figure 2.5 is shown the SEFd and the correspondent hypnogram. It can be noticed that SEFd values are higher in the REM phases and that is why it could be a good discriminant feature from other phases. Also during some awake stages there is an high SEFd, but still lower than REM in most of the cases. This pattern appears because of the high values of SEF95 and low values of SEF50 in REM stages.
The numeric codes assigned at every epoch stage in this work are:

- N1 → 1
- N2 → 2
- N3 → 3
- REM → 5
- WAKE → 6

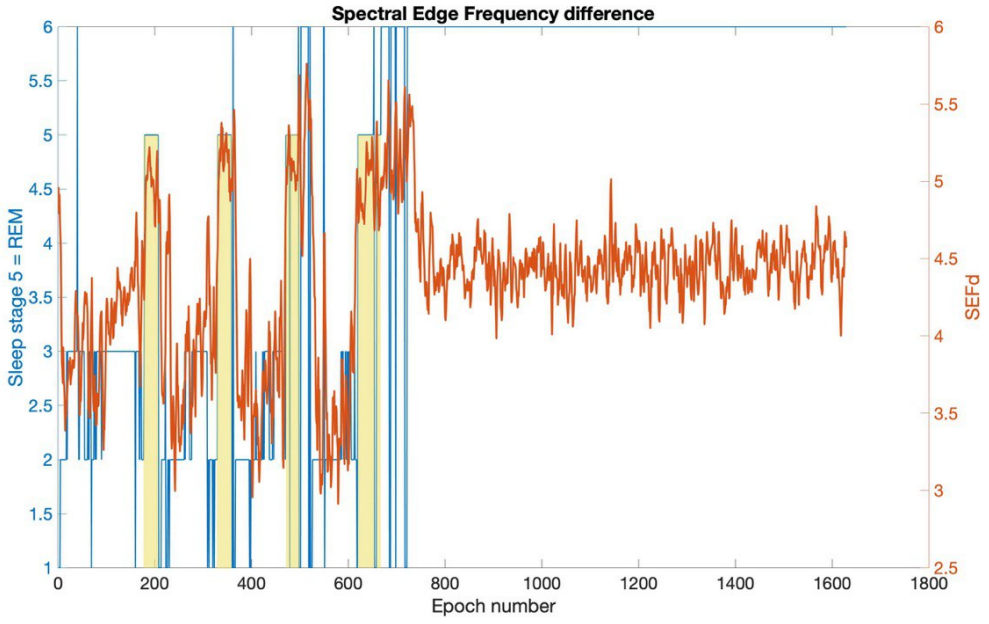


Figure 2.5: SEFd during the epochs of the first patients; REM stage is identified with the value 5, wake is 6

The explanation underneath Figure 2.5 is given by the fact that power changes between REM and NREM stages; in particular, powers are similar at around 8Hz then the NREM power increases: the SEF50 of the REM stages is expected to be lower than NREM. After 15Hz the trend changes and the power of the REM stage starts increasing. Therefore SEF95 will be higher in the REM stages. The Figure 2.6 is taken from the signal of the first patient and it shows briefly the explained concept.

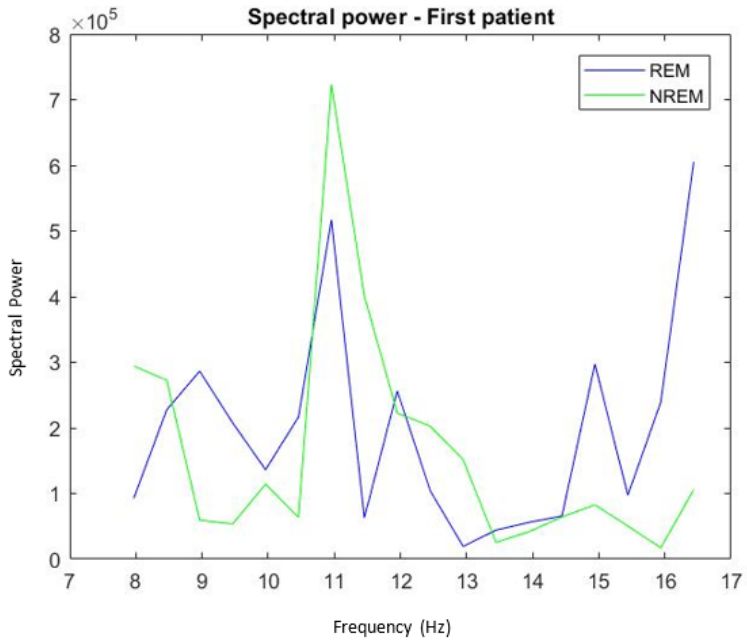


Figure 2.6: Difference between the spectral power of a REM epoch and a NREM epoch

Due to the fact that some epochs are not well distinguished with only the SEFd, especially the wake ones, other characteristics have been extracted for improving the differentiation.
The absolute power (AP) is calculated with the Fourier coefficients of the signal between two selected frequency, in our case 8 and 16 Hz.

$$AP = 20 \times \log\left(\sum_{i=n(f1)}^{n(f2)} (|mag_i|)\right) \tag{2.8}$$

AP has been calculated for every sub-epoch and then averaged for every epoch. During REM stages the AP exhibits the lowest values as shown in Figure 2.7.

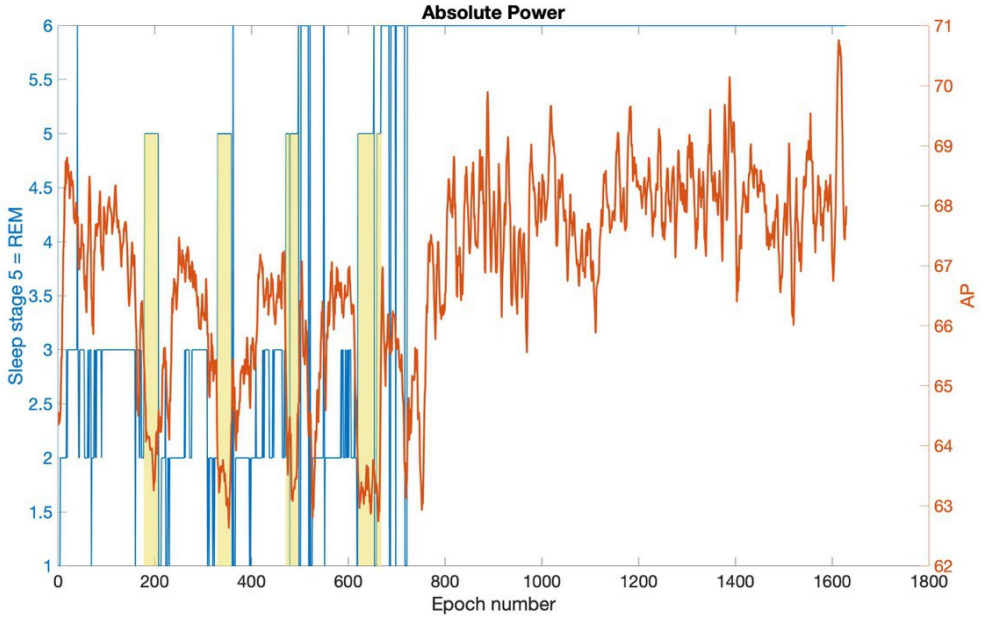


Figure 2.7: AP evolution during the epochs in the first patient

The last used characteristic is the relative power (RP), that is evaluated as:

$$RP = 20 \times \log\left(\frac{\sum_{i=n(f1)}^{n(f2)}(|mag_i|)}{\sum_{i=1}^{n}(|mag_i|)}\right) \tag{2.9}$$

Where *n(f1)* and *n(f2)* are again 8 and 16 Hz. RP does not have a particular behaviour in REM and NREM epochs but it is useful since it can stabilize the recognizing process [42].
All the data has been normalized before starting the machine learning process following the min-max scaling:

$$x' = \frac{x - min(x)}{max(x) - min(x)} \tag{2.10}$$

Part III:

After the feature extraction process, we obtained three long arrays containing all the SEFd, AP and RP calculations for every epoch of every patient. Then AP and RP are smoothed for removing variability.

The annotation array has been organized so that the NREM epochs has been set to 0 while the REM epochs to 5.
In order to perform machine learning algorithms, the MATLAB Machine Leaner App and the Statistics and Machine Learning Toolbox have been used. After a search on the best learning method proposed in literature and a test of which one gave the best performances, the *Ensemble Learning method* has been chosen.

The ensemble learning method is based on the idea that it uses multiple learning algorithms for reaching the best prediction level. It is a supervised method, which means that the input dataset is given with the column of the correspondent responses, in this case the belongings to the REM or NREM stage. The algorithm extracts a space of hypothesis in order to find the suitable one for making the best prediction. By constructing more decision trees from the same data set, every tree will follow a different reasoning method, therefore they could achieve a different prediction. For making a decision the algorithm consults every decision tree and then it chooses the answer that has resulted most times. The decision trees are made through the Boostrap aggregating: dataset is divided in sub-dataset randomly chosen (Figure 2.8). Every tree is estimated on each sub-dataset and yields to a classification. As said before the final prediction is chosen as the most represented between the trees. Following this path the stability and accuracy of the algorithm is promoted.

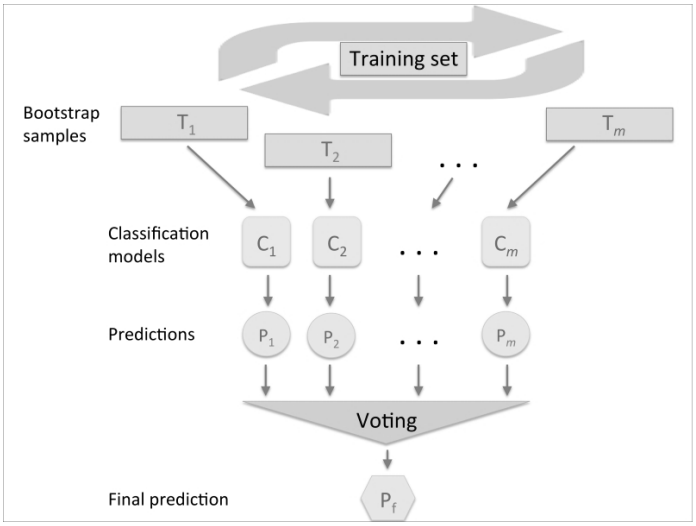


Figure 2.8: Bagging method, adapted from [43]

K-fold cross validation method is applied (Figure 2.9); it consists on a division of the total dataset in k equal parts and every loop each k-th part is used as validation dataset while the remaining part is used for training. This avoids the model overfitting.

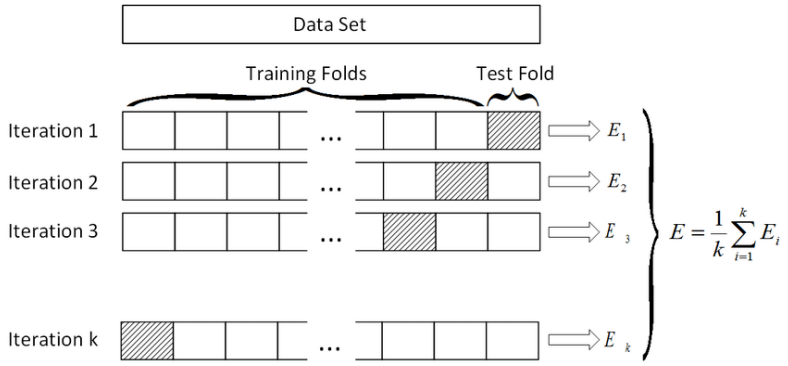


Figure 2.9: K-fold cross-validation method [44]

Therefore the main steps are:

1. Partition into k not intersected folds.
2. For each fold:
   - Train the ensemble model on the out-of-fold observations
   - Test the model on the in-fold data
3. Calculation of the average test error over all the chosen folds.

All stages recognition

Following the approach of the REM-NREM epoch recognition, all stages recognition is attempted.
In fact, the structure of the algorithm is essentially the same but new features have been extracted from the signal, since every sleep stage has its own characteristics.
SEFd, AP and RP are maintained since presented a good discriminant power for REM phases.
Below, a list of all the feature extracted is shown that have been studied in literature; each of them will be explained in detail later.

- SEFd

- AP
- RP
- Mean value
- Standard deviation (STD)
- Mean frequency (MNF)
- Spectral power in δ band
- Spectral power in α band
- Spectral power in θ band
- Spectral power in β band
- Fractal dimension
- Shannon entropy

Statistics features:

The main statistic features have been calculated. For every epoch the mean value of the EEG is extracted with the expression:

$$\mu = \frac{1}{n} \times \left(\sum_{i=1}^{n} x_i \right) \tag{2.11}$$

Then standard deviation is calculated; it yields the amount of dispersion or variation of a dataset. When the standard deviation is low, it means the data are close to the mean value, otherwise they are spread in a wider range. It is extracted as:

$$s = \sqrt{\frac{\left(\sum_{i=1}^{n} x_i - \mu \right)}{n}} \tag{2.12}$$

Feature extracted in frequency domain:

During the sub-epochs loop, after the Fourier transform, some feature are calculated. Mean frequency: since different EEG stages are characterized by a bandwidth, in every epoch the mean frequency is extracted because it can

give a contribution in the machine learning phase.
It is calculated from the statistic definition which expresses that, given a random variable x with a variability distribution of $f_x(x)$ then:

$$m^i = \frac{\int_{-\infty}^{+\infty} x^i f_x(x) dx}{1} \tag{2.13}$$

where m^i is the statistic moment of the i_{th} order and it is divided by the area of the distribution which is 1.
From this the equation used in the algorithm is directly obtainable:

$$MNF = \frac{\int_0^{f_N} f \cdot P(f df)}{\int_0^{f_N} P(f) df} \tag{2.14}$$

Where P is the spectral power, f is the frequency array and f_N is the Nyquist frequency.

Then, the power in each EEG distinctive bandwidth has been extracted; more in specific, δ bandwidth has been selected from $0.5Hz$ to 3.5Hz and it represents the state of a very deep sleep, θ bandwidth between 3.5Hz and 7Hz and it shows up deep sleep, α bandwidth between 7Hz and 14Hz when the subject is relaxed and the last β bandwidth between 14Hz and 21Hz in the other states.
Each of them has been calculated on every sub-epoch and then averaged on the entire epoch.

Entropy and complexity features:

Fractal dimension (FD) and Shannon entropy have been calculated on every epoch and every sub-epoch respectively: Shannon entropy has been averaged then on the entire epoch.

In fractal geometry, a fractal dimension is a ratio that provides an index of the signal complexity; in fact, it statistically measures how well a fractal matches the data at different scales [45]. A fractal is a geometrical object that presents homothety, that is a geometrical transformation in space that dilates or contracts objects maintaining the angles unchanged. FD is very used in the EEG studies in order to distinguish physiological state phases.
There are many methods for estimating the fractal dimension and one of them is the Higuchi's method that exhibits high accuracy and robustness

[46]. In fact, Higuchi's method can be suitable for estimating FD of short signals or for signals that are not stationary because of its moving window approach; therefore, it divides the long signal into short parts that can be considered stationary [47].
The signal epoch is given by a sequence of samples $y(1), y(2), y(3)...y(N)$.
In the Higuchi's algorithm a division in sub-epochs is computed generating k new sub-epochs, defined as [47]:

$$y_m^k = \{y(m) + y(m+k) + y(m+2k), ..., y(m+Mk)\}, m = 1,2,3, ..., k \quad (2.15)$$

where m is an integer that indicates the initial time and k is another integer that represents the interval time. Moreover $M = \frac{(N-m)}{k}$.
For each sub-epoch y_m^k an average length $L_m(k)$ as is computed:

$$L_m(k) = \frac{1}{k}\left[\frac{N-1}{Mk}\sum_{i=1}^{M}(|y(m+ik) - y(m+(i-1)k|)\right] \quad (2.16)$$

$\frac{N-1}{Mk}$ is a normalization factor. The epoch length $L(k)$ is extracted by doing the average on all the k for the different m values [47]:

$$L(k) = \sum_{m=1}^{k} L_m(k) \quad (2.17)$$

Therefore if $L(k)$ is proportional to k^{-D}, the curve which describes the epoch shape could be represented with fractal of dimension D.
By plotting in a double logarithmic scale what should appear is a straight line with slope D as shown in Figure 2.10

This feature is included under the hypothesis that the signal complexity is higher in the wake and REM stages while lower during deep sleep. In order to show this general behaviour the average FD for the sleep stages has been plotted in Figure 2.11

The trend of this feature is inferable from the picture even if the difference between the stages is not very consistent. This could be due to the fact that data are taken from patients that had an acquisition system with a specific sampling rate that can lead to uncertainty.

The last feature is the Shannon entropy (SE).
The study of the entropy is important in order to understand the signal complexity. If the signal has high entropy, it means that it has a low predictability.

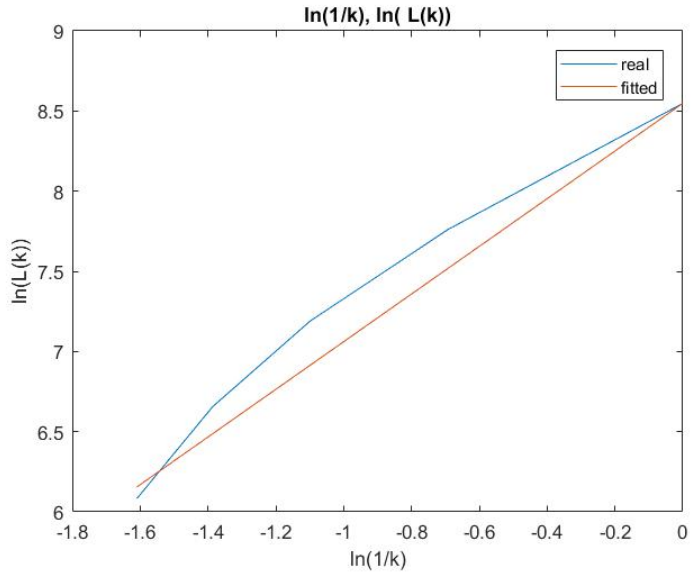


Figure 2.10: Fractal dimension of one epoch belonging to the last subject

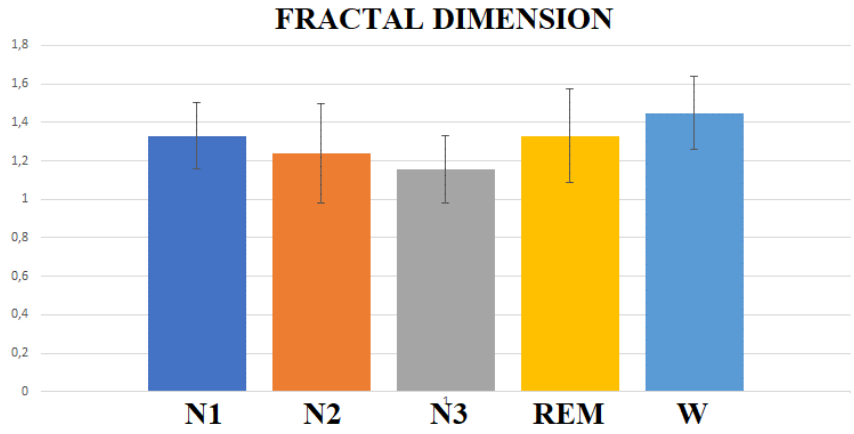


Figure 2.11: The average values of fractal dimension divided by stages

The spectral entropy has been adapted to the Shannon entropy definition; there are two main steps, first the spectral power is normalized on the total area, second the equation for calculating the entropy value is applied.

$$p[f] = \frac{P[f]}{\sum_1^N P[f]} \tag{2.18}$$

$$SE = \sum p[f] \times log\left(\frac{1}{p[f]}\right) \tag{2.19}$$

This feature is expected to have a similar behaviour as the previous one, since the complexity of the signal decreases with the sleep depth, as shown in Figure 2.12.

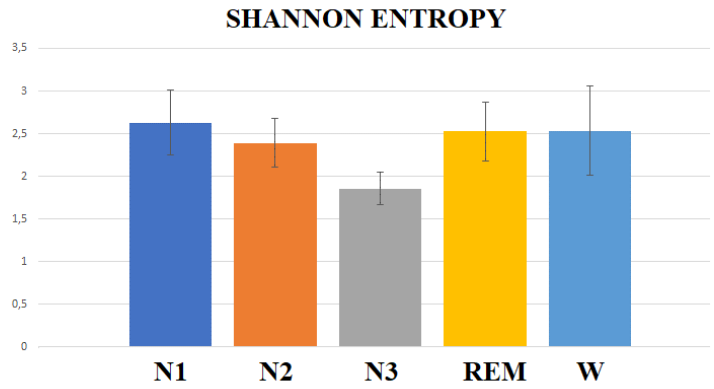


Figure 2.12: Average Shannon entropy during the stages

At the end of this features extraction process, before using features in the machine learning, they have been normalized through the min-max scaling.

The machine learning process is the same as the previous model, that is the *ensemble learning method*, with *K-fold cross validation*.

All stages recognition through wavelet transform

Another path has been followed in order to recognize all sleep stages. Since in literature the wavelet transform use is very common, this approach has been tested too. In fact, EEG is a dynamic signal, mostly non-stationary, and this technique can provide details that other techniques could not give.

Wavelet transforms represent the signal through a finite oscillating wave form that is scaled and shifted in order to adapt to the input signal. This transformation is based on the scalar product between the signal and a mother wavelet, that is allowed to change only in time extension but not in shape. The continuous wavelet transform (CWT) permit to take advantage of a larger window for analysing the low frequencies of the signal and

a narrower one for analysing the high frequencies. Therefore there are two important parameters: the scaling variable and the shifting variable.

$$[W_\psi f](a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\psi^* \left(\frac{t-b}{a} \right) \tag{2.20}$$

This is the scalar product between the signal and a set of wavelets scaled and shifted by the parameters a and b respectively, called daughter wavelets.

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right) \tag{2.21}$$

With the CWT, signal is analysed in a continuous way through functions that are a continuously modified versions of the mother wavelets; a certain level of redundancy is entailed, namely not all the results are useful for reconstructing the original signal. Since a and b are discrete, the level of sampling allows different degrees of redundancy. As a consequence, a discrete wavelet tranform is applied (DWT).
Through a dyadic sampling (Figure 2.13) all the basis functions are needed to reconstruct the signal. This means that every decomposition level is built by filtering the signal in a way that half of the input bandwidth is on the output.

$$a = 2^n \tag{2.22}$$

$$b = m2^n \tag{2.23}$$

$$\psi_{n,m}(t) = \frac{1}{2^{\frac{n}{2}}} \psi \left(\frac{t-2^n m}{2^n} \right) \tag{2.24}$$

Through this process the signal is analysed at different resolutions, therefore it is called a *Multiresolution analysis*.
In order to approach to this procedure and consequently understanding the algorithm that has been used, it is important to specify a chain of nested spaces (Figure 2.14) that are defined *approximation spaces*, as instance $V_{i+1} \subset V_i \subset V_{i+1}$.
Every space gives additional information with respect to the previous one of $W_{i+1} = V_i \backslash V_{i+1}$ where W_{i+1} is the *space of details*.
A function included in a space can be written as the direct sum of a detail

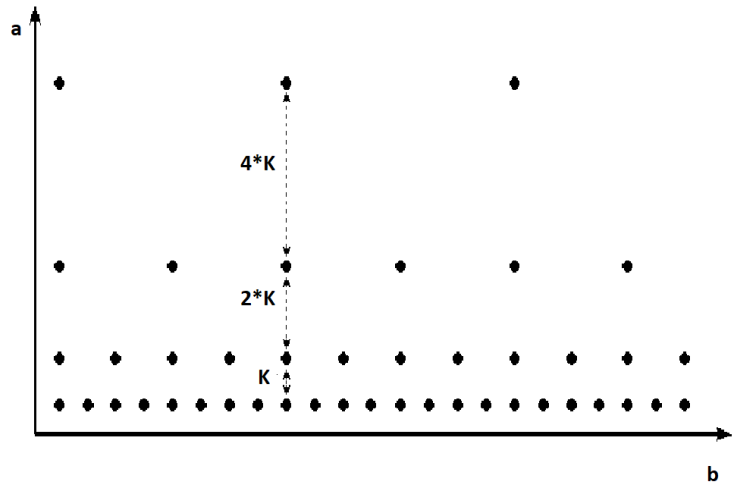


Figure 2.13: Dyadic sampling of a and b

space and an approximation one, for instance:

$$V_{-1} = V_0 \oplus W_0 = \overbrace{V_1 \oplus W_1}^{V_1} \oplus W_0 ... \tag{2.25}$$

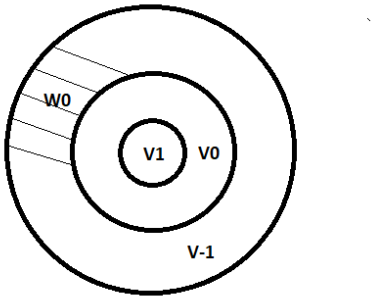


Figure 2.14: Nested spaces adapted from [48]

Therefore the approximation spaces are constructed on scaled and shifted versions of a wavelet called father wavelet, while the detail spaces are built on the versions of the mother wavelet. In particular:

$$V_j = 2^{-j/2}\phi(2^{-j}t - n) \tag{2.26}$$

and

$$W_j = 2^{-j/2}\psi(2^{-j}t - n) \tag{2.27}$$

Considering the father wavelet in the V_0 space, it is also included in the V_{-1} space; from this assumption it can be decomposed as a basis of V_{-1}:

$$\psi(t) = \sqrt{2} \sum_n h_n \psi(2t - n) \tag{2.28}$$

The mother wavelet is included in W_0 therefore can be seen as before, belonging to V_{-1}:

$$\phi(t) = \sqrt{2} \sum_n g_n \phi(2t - n) \tag{2.29}$$

It results that the coefficients h_n can be expressed as a low-pass filter while g_n as a high-pass filter.
Every decomposition level is therefore composed by a the application of a low pass filter, a high-pass filter followed by a down-sampling, Figure 2.15

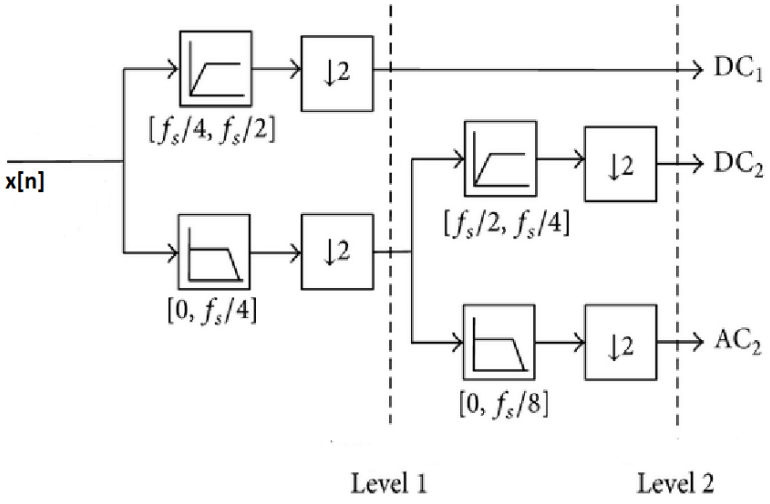


Figure 2.15: Bank of filters

Therefore, for each epoch and for every patient, some features directly extracted from the results of this transform are calculated.
In particular, the epochs are decomposed in seven levels and, by knowing the corresponding bands of each decomposition, the coefficients are consequently gathered in order to reproduce the classic EEG bandwidth splits (θ, δ, α, β, γ).
For this procedure MATLAB provide a wavelet packet and the "db2" wavelet is used, from the Daubechies wavelet family.
The feature extracted are:

- Energies in every band: after the division in band the associated energy is calculated through the coefficients (Ea_1, Ea_2, $Ea3$, Ea_4, Ea_5).

- Total energy: the sum of all the energies in the levels is considered ($E7$).

- Ratio between each band on the others: for example the first ratio calculated is energy in α band on θ and δ ($E8$, $E9$, $E10$).

- Mean value: it is calculated for every band ($m1$, $m2$, $m3$, $m4$, $m5$).

- Standard deviation: it is calculated for every every band too ($st1$, $st2$, $st3$, $st4$, $st5$).

- Spectral spread: this feature returns the spectral spread of the epoch over time ($spread$).

- Spectral flatness: this feature returns the spectral flatness of the epoch over time and can be used for understanding how noisy the epoch is [49] ($flatness$)

- Spectral slope: it returns the slope of the spectral shape and how fast the spectrum goes towards high frequencies ($slopet$).

- Spectral centroid: it returns the spectral centroid as the weighted mean of frequencies in the signal ($centroid$).

- Spectral entropy: it returns the spectral entropy derived from the calculated coefficients ($entropy$).

This features have been normalized through the min-max scaling technique and then used in the same machine learning model known as ensemble learning method, and then validated through the K-fold validation.

The obtained results has been similar to the previous algorithm, but after a feature selection the accuracy has been improved.
In fact, MATLAB provides several feature selection functions and the chosen one uses the neighbourhood component analysis in order to understand the input feature relevance with respect to the response vector.

2.2.2 Exploring inertial signals

After studying the EEG signal taken from an already existing dataset, a study on inertial signals is begun; this time we collected the data.
In fact, inertial signals during sleep can give further information about the subject health status. In particular, a combination of an EEG and an inertial system could give an idea about the sleep phases the subject goes through and the related movement he does.
We will present some features that can be studied by employing the SensorTile.box on healthy people and on a PD patient.

Healthy subject signals

In order to understand how PD patients sleep in comparison to controls, inertial signals from both have been collected.
A group of 6 healthy subjects, between 23 and 25 years old, was selected. The recordings started when the subject laid down in bed before sleeping and stopped in the morning, with an average of eight hours of sleep.
The SensorTile.box is inserted in an elastic band and applied on the chest. The axis are oriented as the Figure 2.16 represents.

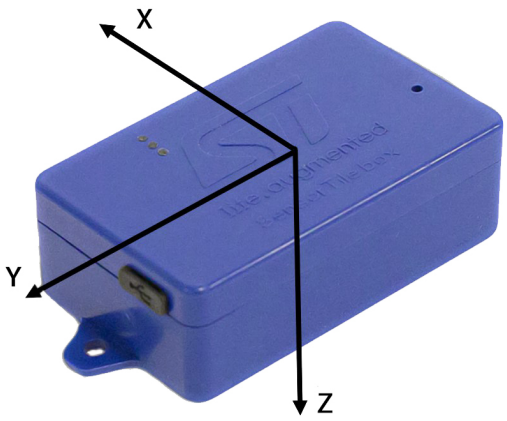


Figure 2.16: Sensor axis

The sensor is connected through app for smartphones (ST BLE sensor) via Bluetooth; for beginning the recordings a start button is pressed. At the end, the recorded data were transferred to a personal computer and processed using Matlab.

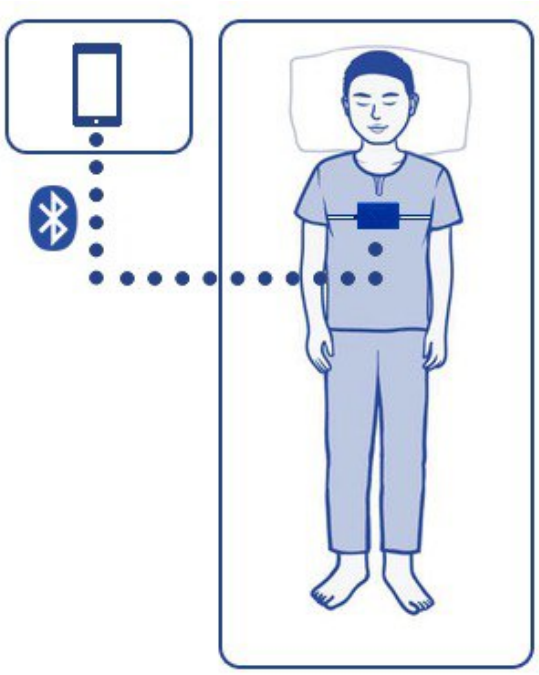


Figure 2.17: Sensor positioning

The obtained data are in .csv format: they include date, time, acceleration and angular speed on the three axis.
From these signals an activity index is calculated, a number of turning in bed and the velocity is extracted.

In Figure 2.18 there are the accelerations from an healthy subject, recorded at home through the SensorTile IoT app.

The first step was determining from the accelerometers the patient position in bed. This algorithm is based on thresholds.
The sensor is positioned with the x axis on the longitudinal axis of the body, therefore the z axis and the y axis will be taken into account for assessing the position.
The position is assigned depending on the acceleration each axis feels; for example, when the subject is supine, on the z axis there is all the contribution of the gravitational acceleration.
The accelerometer signal is divided into 60 seconds epochs and the z component is checked first, then the y component. If the z and the y has a similar

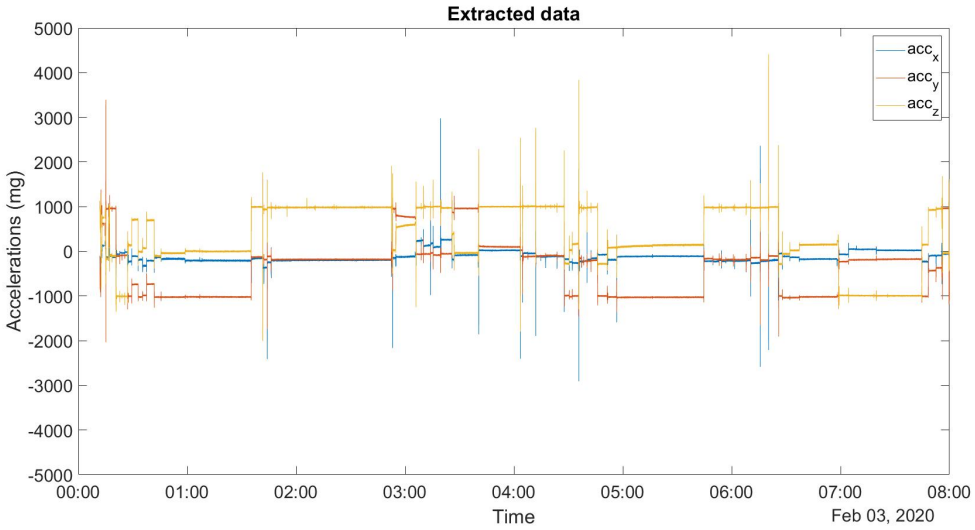


Figure 2.18: Accelerations from an healthy subject

value, the position is assigned based on which one has a bigger modulus.
In Figure 2.19 the implemented flowchart is represented.

At the end of this process, each epoch has its position, codified as a number.
The second step is to understand when the subject has turned in bed. By making the difference between the elements in the position vector, when the difference is not zero, the patient has turned. We only considered all the changes in position that has last, at least, five minutes, since it is reported in literature.
In order to make a double check on the turning in bed moments the angular velocity is extracted too. When the patient changes position there is a peak in the angular velocity: this value helps to confirm the event and to calculate time for turning and the maximum velocity the patient has done. In this case, since the subject rotates on his longitudinal axis, the x axis is taken into account (Figure 2.20).
The accelerometer is divided in one-minute epochs while the velocity in 5 seconds epochs, in order to have a better resolution. When the peak in the angular velocity is searched, this difference in epoch length has been taken into account.

The final value calculated is the activity index.
Since most of the activity indexes are not publicly available because are used

in commercial devices, we implemented a new method taken from literature [51].
It exploits the variance of the raw accelerometry data for calculating the proposed index.

Each velocity is divided into 60 seconds epochs, and for every epoch the variance is extracted; a systematic noise variance is calculated when the sensor is not moving, $\bar{\sigma}_i$. Then is applied the following equation for every epoch:

$$AI = \sqrt{\frac{1}{3}[(\sigma_x^2 - \bar{\sigma}_i^2) + (\sigma_y^2 - \bar{\sigma}_i^2) + (\sigma_z^2 - \bar{\sigma}_i^2)]} \tag{2.30}$$

The proposed AI is easy to implement and presents characteristics as additivity and rotational invariance, which means that it summarize the magnitude of movement over three axes regardless the device orientation.

Patient signal

The patient was hospitalized at Molinette hospital, in Turin, in order to undergo the DBS implantation and agreed to participate to this project. The subject suffered from PD since 2011 and wearing off symptoms were impairing. Other reported symptoms are limbs pain and difficulties in moving.
The SensorTile.box has been inserted in an elastic band and applied on the chest in the same way as on controls.
In order to have a validation from the data extracted, all patient's movements have been observed and annotated during the night. Moreover the patient has been also given a diary in which she could write when she woke up and got up.
The recordings begin at 22:26 in the night and are stopped at 3:57am.
The accelerometers data are plotted in Figure 2.21.

The z axis (yellow) measures a positive acceleration since it detects the gravitational acceleration (g); then the blue signal is the x axis while the red one is the y signal.
As can be seen from the graph, the patient basically does not move during the night remaining in a supine position. This is inferable also by the annotations that has been done. Since most of the PD patients show a difficulty in moving and turning in bed, this could be one of those cases.
More in specific, the first three oscillations in the plot corresponds to two coughing episodes and a moment in which the subject was drinking water,

always maintaining the same position.
Another remarkable event is at about 3:50am in which the patient stood up, as can be seen in the x axis, that perceived a negative gravitational acceleration.

In fact, the subject never turns in bed, but has just some brief moments in which slightly moves.
Therefore just the *activity index* can be extracted in order to evaluate the amount of movement the patient has done.

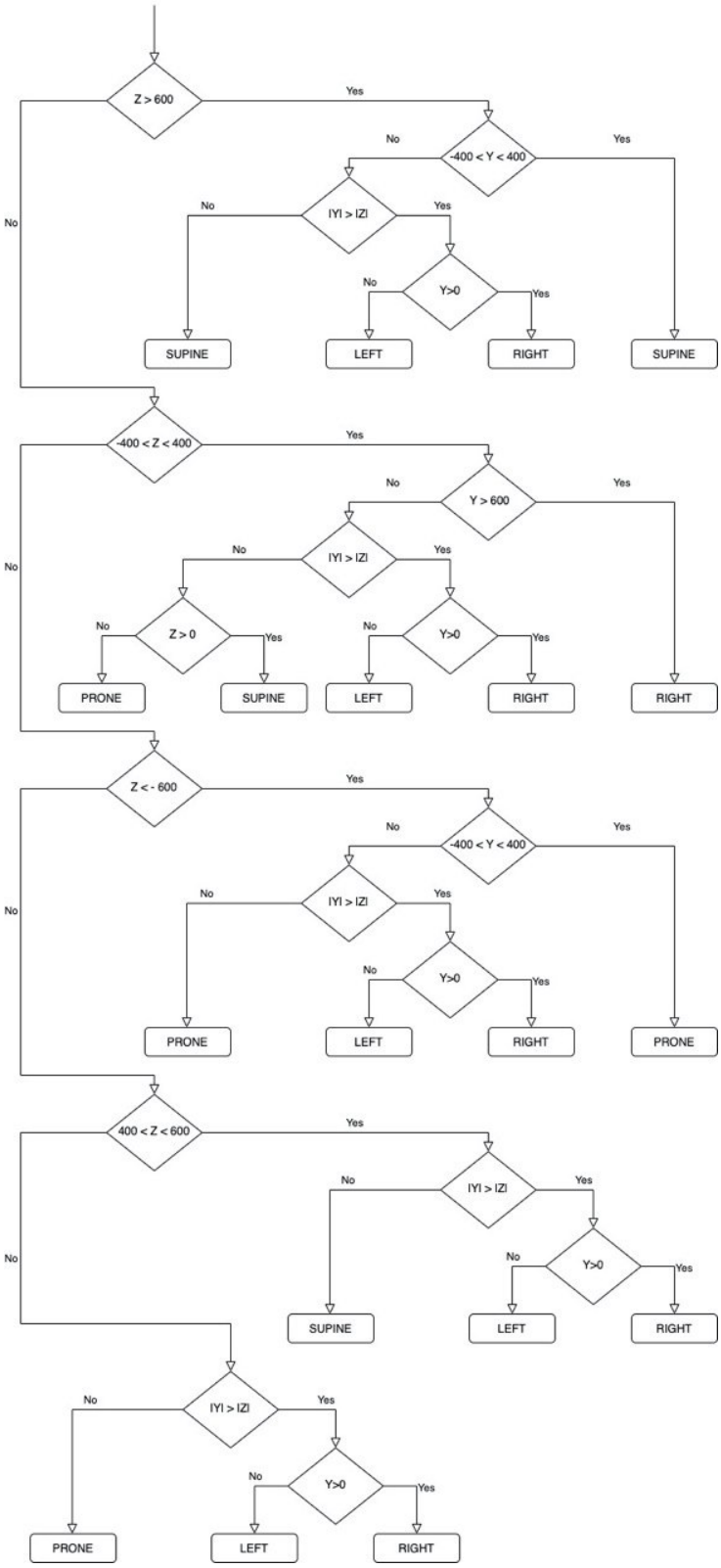


Figure 2.19: Flowchart for assessing the position

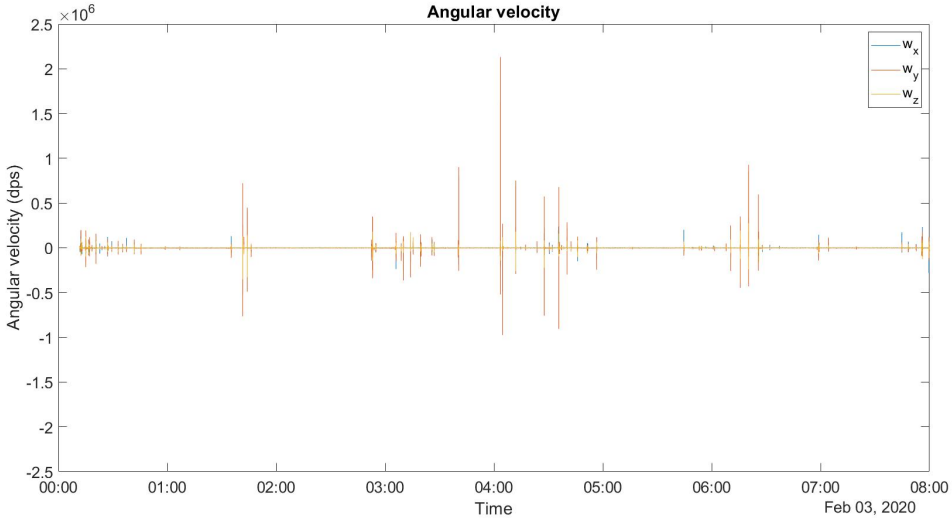


Figure 2.20: Subject angular velocity

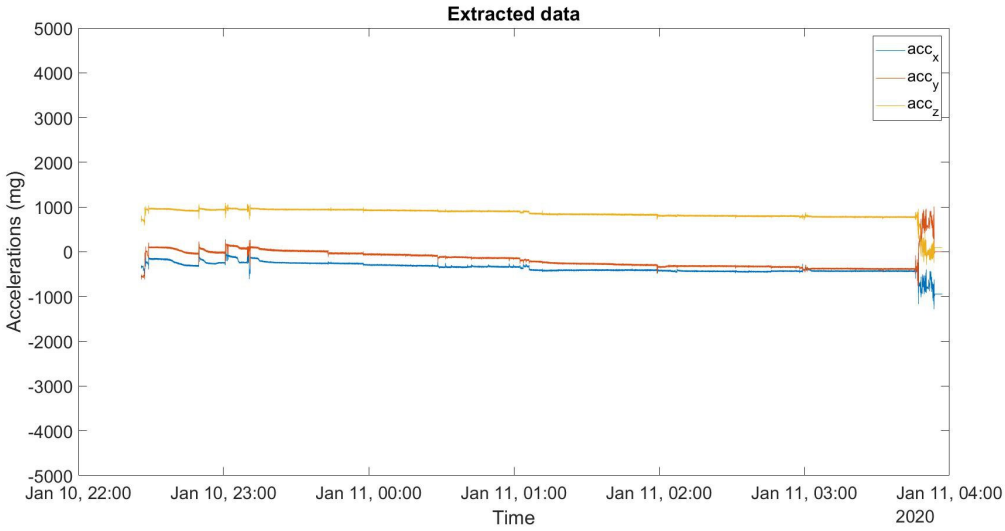


Figure 2.21: Patient's accelerometers data

Chapter 3

Results

In this section the results of the algorithms that has been tested are going to be presented.
The accuracies reached with the EEG algorithms have been promising; the study on inertial signals has been interesting too, since the obtained results match what read in literature and what observed during the night in the hospital.

3.1 REM-NREM epochs recognition results

The results of this algorithm have been satisfactory since the number of features used is restricted but very focused on the problem.
Here below in Figure 3.1 it is shown the confusion matrix after the k-fold cross validation, therefore the prediction is made on the test set.

The rows express the true class while the columns are the predicted class. Green cells are where the classifier has performed well, on the diagonal there are cells where the true class and the prediction match.
Percentages of true classification and misclassification are also calculated (Figure 3.2).

Less then 1% is incorrectly classified as NREM while 9% is misclassified as REM. This can be explainable also because the algorithm has more possibilities to train on NREM than on REM epochs.
Since it is a binary recognition, some values can be expressed in terms of sensibility, specificity and accuracy.

$$Sensibility = \frac{TP}{TP + FN} = \frac{7078}{7078 + 734} = 0.9060 \rightarrow 90.60\% \tag{3.1}$$

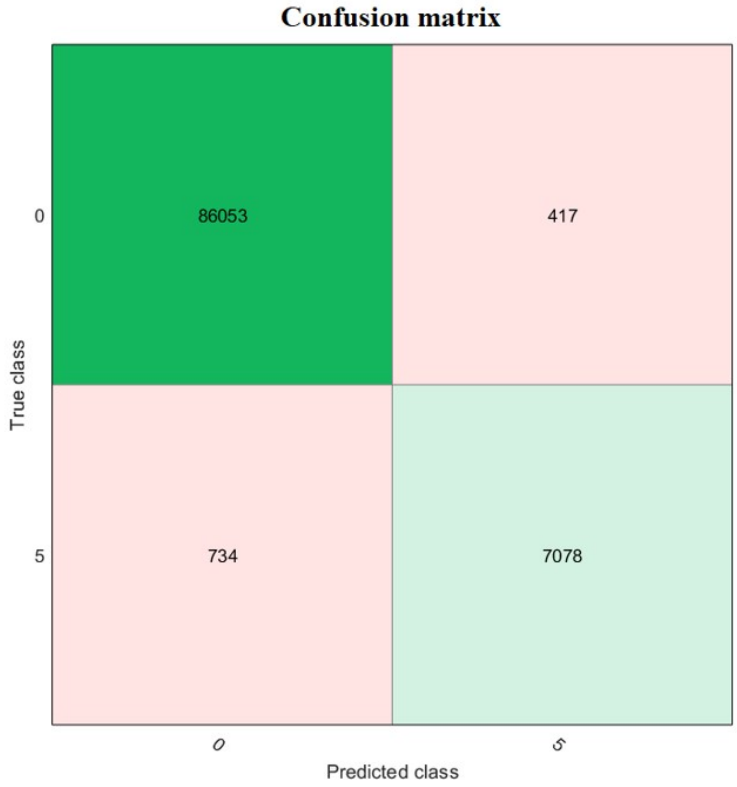


Figure 3.1: Confusion matrix in the REM and NREM recognition algorithm

$$Specificity = \frac{TN}{TN + FP} = \frac{86053}{86053 + 417} = 0.9951 \rightarrow 99.51\% \quad (3.2)$$

$$Accuracy = \frac{TP + TN}{TN + FN + FP + TP} = \frac{93131}{94282} = 0.9878 \rightarrow 98.78\% \quad (3.3)$$

Where:

- TP= A true positive is when the epoch is classified as REM and it actually is.
- FP= A false positive is when it is classified as REM when is not.
- TN= A true negative is when the epoch is classified as NREM and it actually is.
- FN= A false negative is when it is classified as NREM when is not.

Sensibility expresses the algorithm ability of recognizing the REM epochs while the specificity gives the ability in recognizing the NREM epochs. By

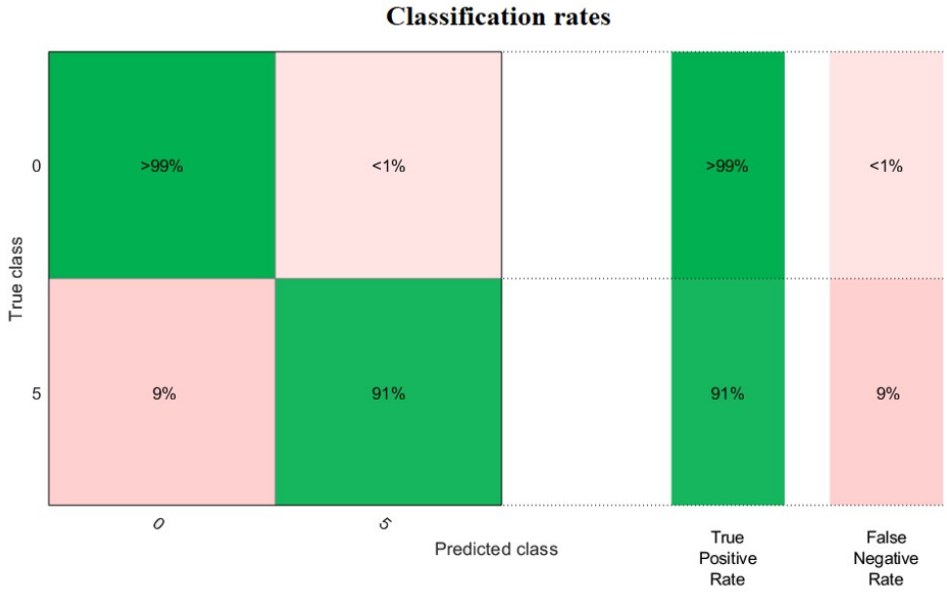


Figure 3.2: Confusion matrix with percentages

using this two values the receiver operating curve (ROC curve) can be extracted, Figure 3.3.

The AUC is the area under the curve that gives an overall quality of the classifier.

3.2 All stages epochs recognition results

The results of this second algorithm have been encouraging too; in fact, a high accuracy has been reached. The confusion matrix is reported in Figure 3.4.

Here again there are the true classifications on the rows and the predicted classifications on the columns. The percentages of true classification are showed too in Figure 3.5.

As can be seen from the values the worst classification is given on the N1 stage. This can be explained because the frequencies of this first phase are similar to the N2 and also can be misunderstood as REM. In fact, a measure of movement could be essential for given a precise classification for the REM phase, since the presence of muscle atonia.
Moreover the N1 epoch are less than all the other stages, therefore the model has a lower number of epochs to be trained on (Table 3.1)

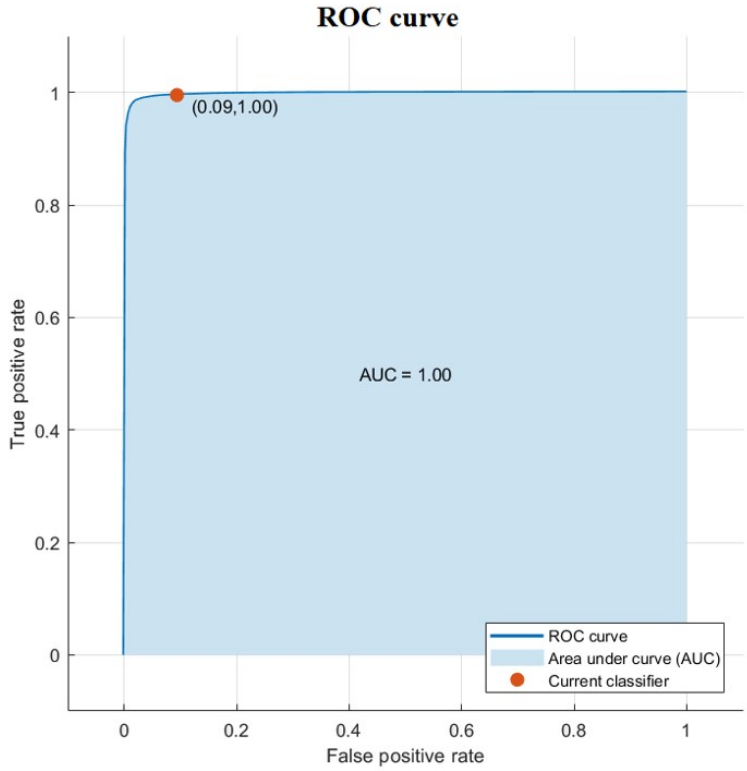


Figure 3.3: ROC curve in the REM-NREM algorithm

	N1	**N2**	**N3**	**REM**	**AWAKE**
Epochs Number	3165	18870	5695	7812	58740

Table 3.1: Number of epochs per stage

Under these hypothesis, a measure of movements and an higher number of N1 epochs could increase the recognition rate.
Also in this case, the sensibilities and specificities for every stage can be calculated (Table 3.2).

The total accuracy has been calculated as the corrected classification divided by the sum of the entire matrix.

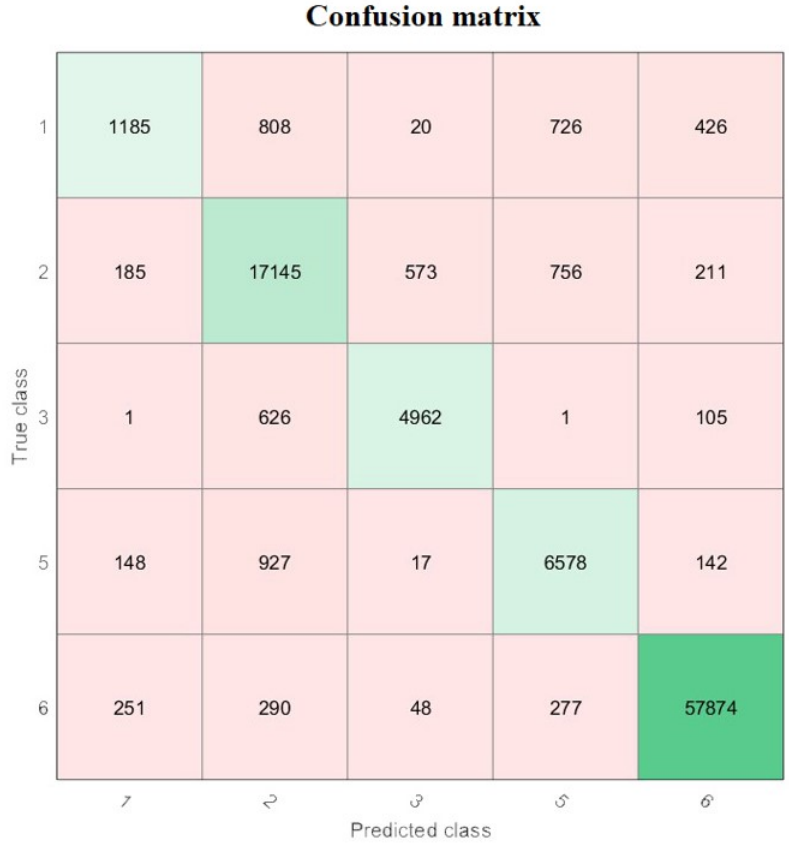


Figure 3.4: Confusion matrix of the all stages recognition algorithm

	N1	**N2**	**N3**	**REM**	**AWAKE**
Sensibility	37,441%	90,859%	87,129%	84,204%	98,526%
Specificity	66,949%	86,608%	88,292%	78,892%	98,496%

Table 3.2: Sensibilities and specificities for every stage

$$Accuracy = \frac{87744}{94282} = 0.9307 \rightarrow 93.1\% \tag{3.4}$$

The ROC curves for each stage are shown below, in a unique figure (Figure 3.6).
The best AUC is obtained for the awake stage as expected.

The worst classification is on the N1 stage and it reflects on the ROC curve, in fact it is the one that is closer to the bisector. All the AUC are presented in the Table 3.3.

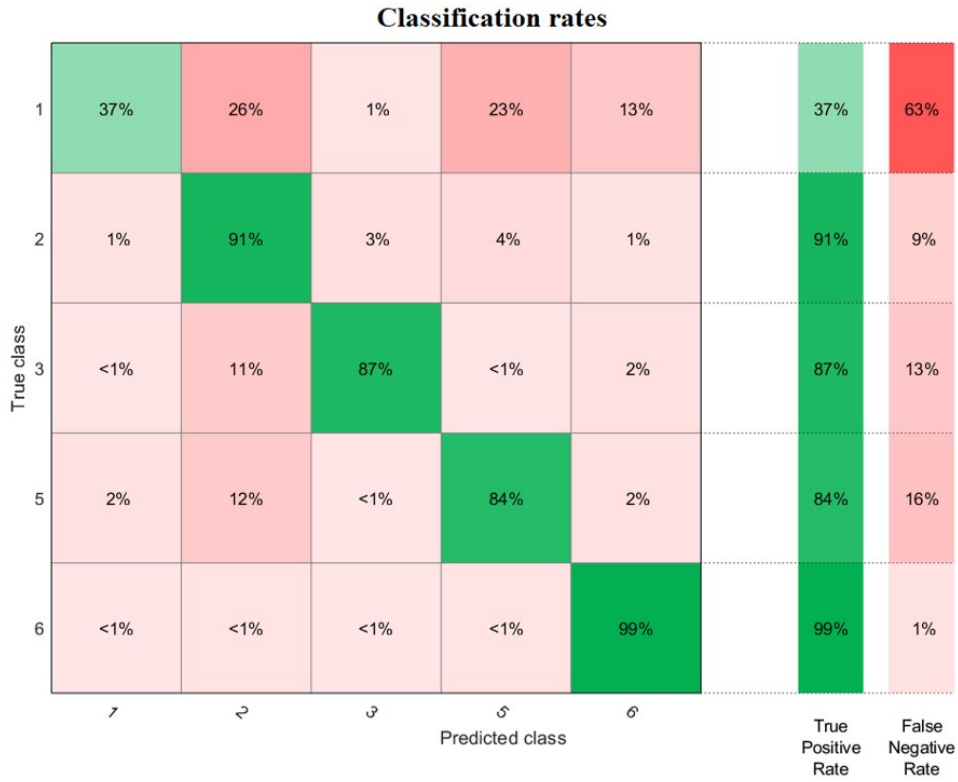


Figure 3.5: Confusion matrix with classification rates in the REM-NREM recognition algorithm

	AUC
N1	0.9443
N2	0.9867
N3	0.9934
REM	0.9862
AWAKE	0.9977

Table 3.3: Area under the curves for every stage

3.3 All stages epochs recognition through wavelet transform results

From the ensemble learning method, by including all the features explained in the 2.2.2 section, good results have been achieved. The correspondent

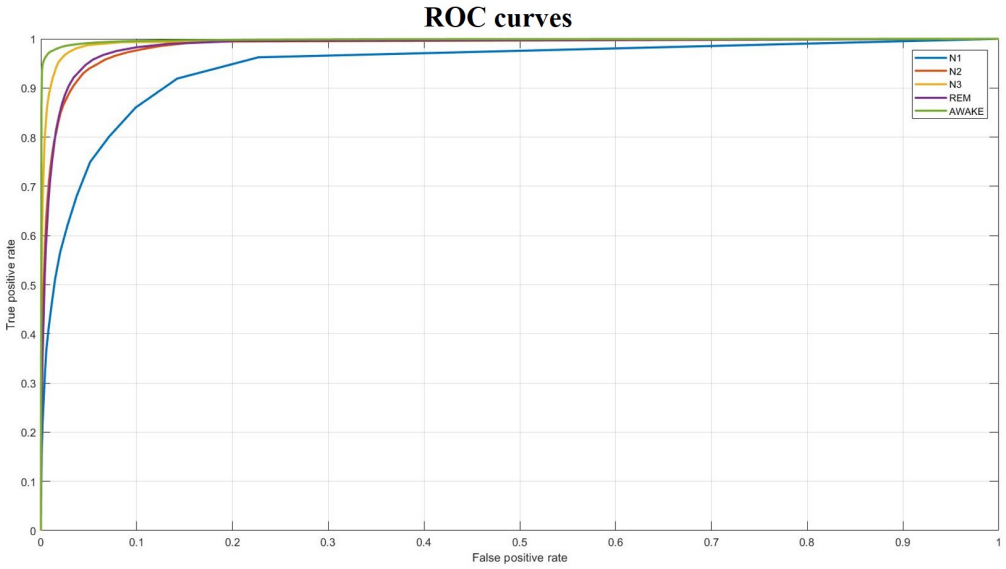


Figure 3.6: ROC curves for every stage

confusion matrix is reported in Figure 3.7.

Here again the most problematic stage to be classified is N1, since it is misclassified as N2 or REM.
The accuracy of the model reaches the 91.6% and the correct classification rates for each stage are in Table 3.4.

N1	N2	N3	REM	AWAKE
32%	89%	86%	75%	98%

Table 3.4: Correct classification rate

Since it has been noticed that the relevance of SEFd, AP and RP have a high weight in the model performances, they have been included again in the final model.
The ensemble learning method, with all the features, yielded a model accuracy of 93.2%. The confusion matrix with the number of classifications and the percentages are shown in Figure 3.8 and in Table 3.5.

It is important to underline that, as suspected, the correct classification of REM phases has increased from 75% to 83%.

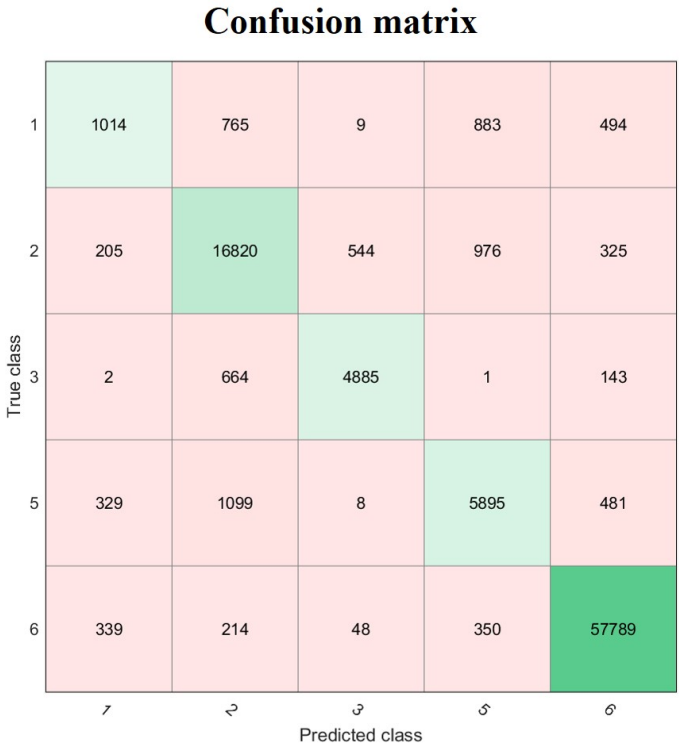


Figure 3.7: Confusion matrix of the all stages recognition algorithm through wavelet transform

N1	N2	N3	REM	AWAKE
40%	91%	87%	83%	99%

Table 3.5: Correct classification rate

As anticipated, the last step has been understanding the feature relevance in relation to the response and, potentially, removing some of them in order to achieve the best performance. The neighbourhood components analysis has been used and the feature relevance has been plotted in Figure 3.9.

The relevance of SEFd, AP and RP has been pointed out to remark their importance.
Therefore, after this identification, the feature that has weight in amount of zero has been removed, in particular: the energies and the total energy, the ratio $\theta/(\delta+\theta)$, the standard deviation of the bandwidth containing noise and

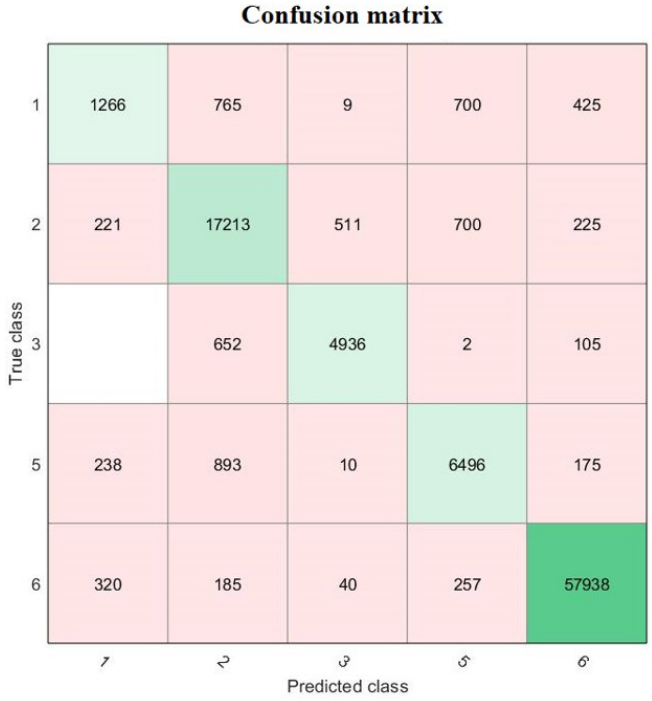


Figure 3.8: Confusion matrix with the addiction of SEFd, AP and RP

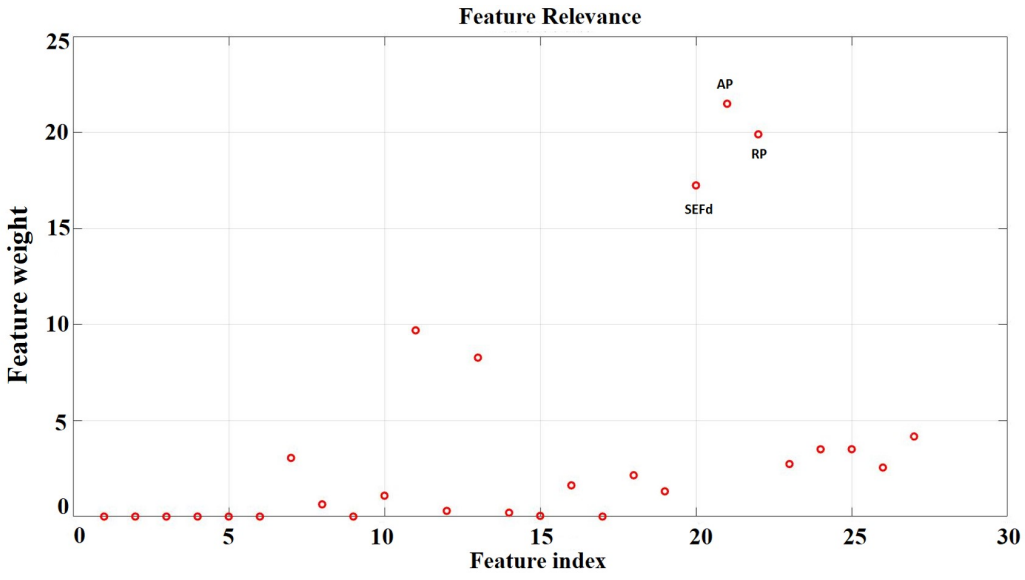


Figure 3.9: Feature Relevance

γ and the α rhythm standard deviation.
The accuracy of the ensemble model has increased up to 93.5% after this procedure (Figure 3.10).

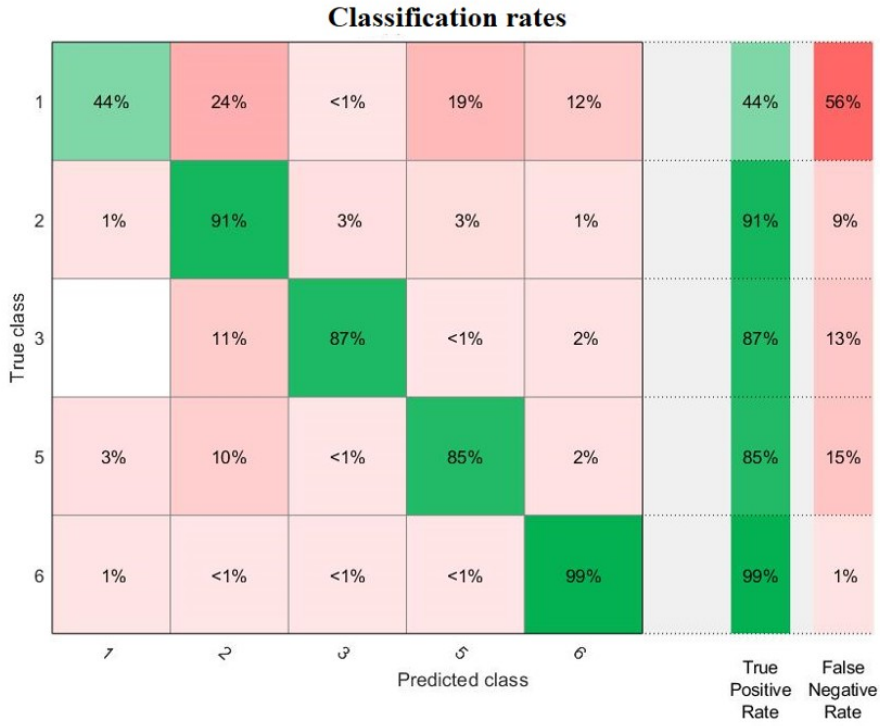


Figure 3.10: Confusion matrix with percentages after the feature selection

In this case the best results in the N1 recognition is reached with a 44% of corrected classified, but also the other stages identification improve.
The ROC curves are represented below too in Figure 3.11.

3.4 Inertial signals algorithm results

Six subjects has been recorded during six nights in order to understand their movements during sleep. In Table 3.6 the main features are presented: minutes in each position, number of turns, average time for turning in bed, average maximum velocity and the maximum activity index reached during the night.
In Table 3.7 there are the patient data. By comparing the two tables, it is possible to understand the patient barely moved during the night, with only one turn in bed. Healthy subjects turn, on average, twelve times per night

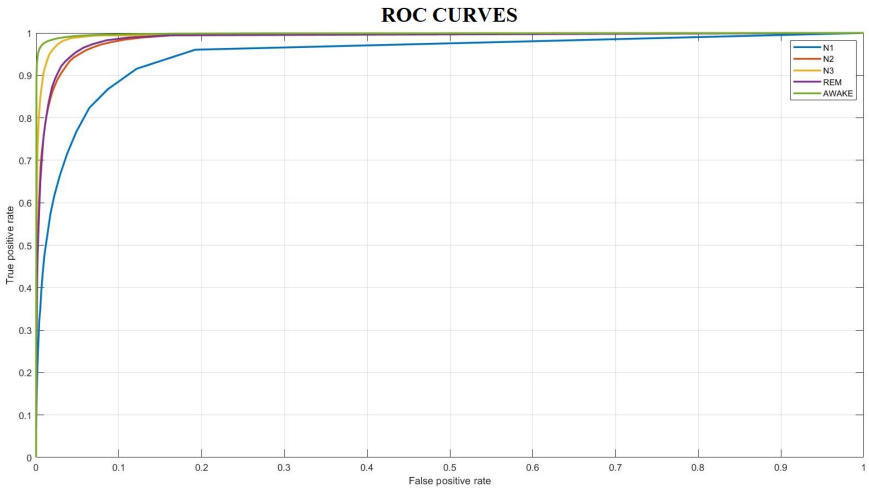


Figure 3.11: ROC curves after the features selection

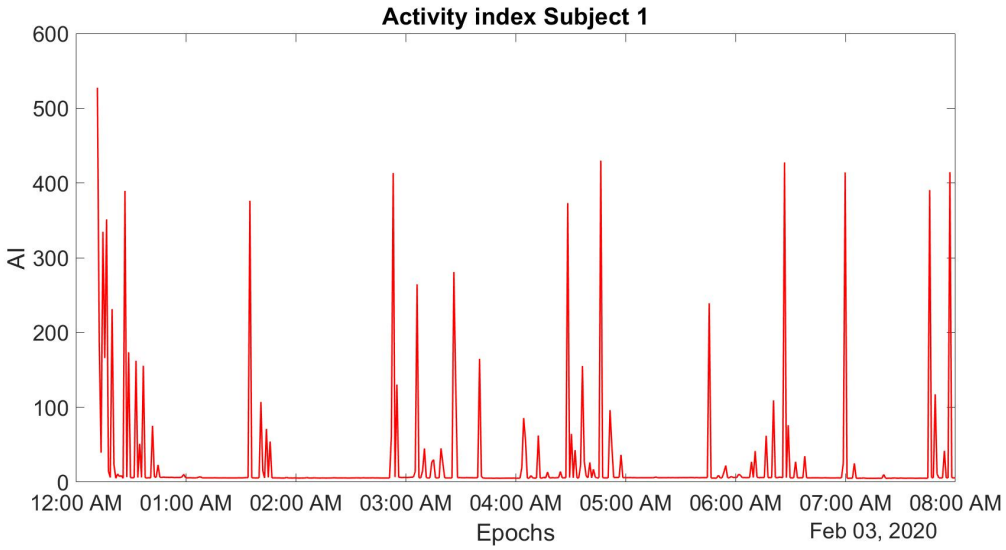
and they take half the time for doing it.
As we observed during the night in the hospital, the patient remained almost all the time in a supine position; the proposed algorithm recognizes 325 supine one-minute-epochs as well.
The only turn in bed results at the end of the recording, when the subject is waking up.
The activity index confirms this behaviour too, since the maximum AI reached by the patient is 200 while the other subjects reached, at least, a double value.
A comparison between the activity indexes is proposed in Figure 3.12(b) and 3.12(a); even if the duration of the recordings is different the amount of movement in the patients is clearly reduced. This impairment can be seen also in the trend shown in Figure 3.13 that demonstrate how the patient slowly stops moving as the hours pass, probably confirming the theory of hypokinesia. The trend has been extracted by taking thirty minutes epochs in the AI calculation.
Moreover the total sleep duration of the patient is lower then the healthy subjects: PD patients have difficulties in maintaining sleep for many hours.
The third subject reported back pain during the day before the recording: this information can be observed in the activity index that results lower than the other healthy subjects.

Subject	**R** *min*	**L** *min*	**SUP** *min*	**PRO** *min*	**n TURNS**	**TIME** *s*	**VEL** *rad/s*	**AI**
1	37	173	206	52	13	10.69	0.27	527
2	190	61	92	99	9	13.61	0.26	566
3	75	201	170	13	11	13.44	0.24	533
4	198	18	198	25	13	12.84	0.19	417
5	159	76	128	102	9	12.03	0.26	572
6	130	227	139	52	14	12.94	0.29	575

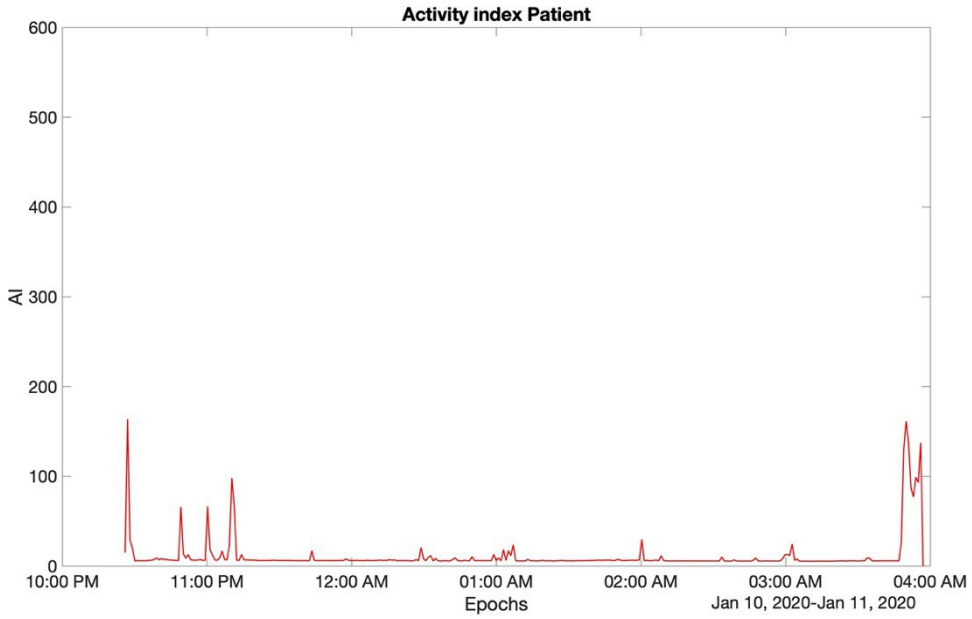
Table 3.6: Values from healthy subjects

Subject	**R** *min*	**L** *min*	**SUP** *min*	**PRO** *min*	**n TURNS**	**TIME** *s*	**VEL** *rad/s*	**AI**
1	6	0	325	0	1	23.58	0.076	164

Table 3.7: Values from the patient



(a) Subject 1 activity index



(b) Patient activity index

Figure 3.12: Comparison between activity indexes

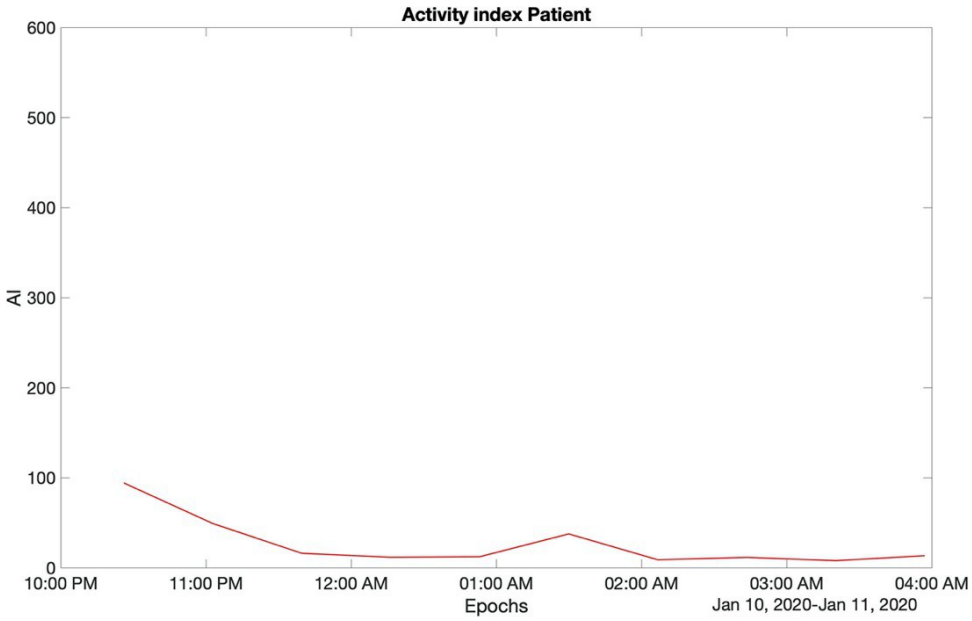
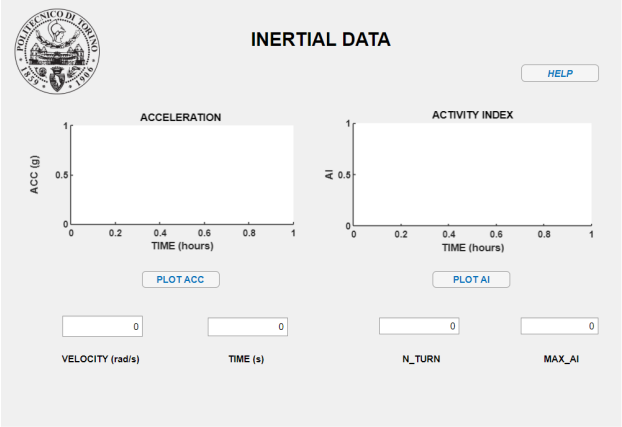


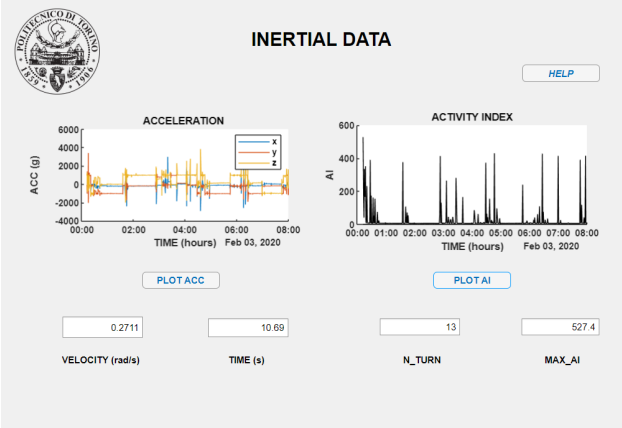
Figure 3.13: Evolution of the patient's movement every 30 min

3.5 User interface

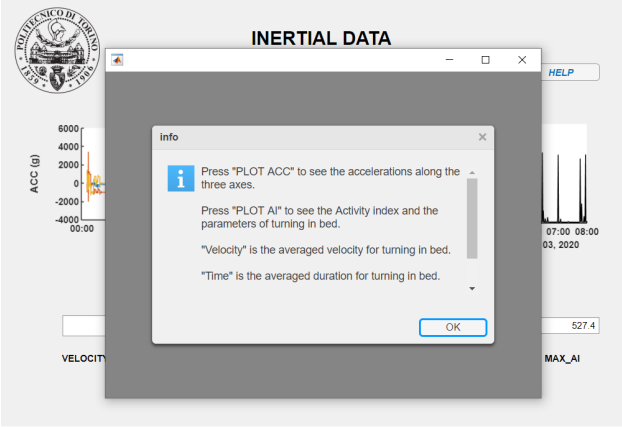
This work is aimed to help and make PD patients' clinical picture more complete. Therefore an user interface is often a tool for visualizing data in a easier and faster way, especially for doctors and clinicians.
MATLAB allows to design interfaces for presenting results.
An interface is proposed for displaying inertial data (Figure 3.14(a)): accelerometer data plot can be obtained by pressing the correspondent button. Another button, if pressed, shows the activity index plot and the main features already reported in Table 3.6 (Figure 3.14(b)).
A help button is available in order to explain briefly the depicted data and the values (Figure 3.14(c)).



(a) User interface layout



(b) Data shown after pressing both buttons



(c) Help message

Figure 3.14: Comparison between activity indexes

Chapter 4

Conclusions

4.1 Future developments

The study we pursued lays the foundation for a more detailed research, in order to have a complete vision about sleep disorders in Parkinson's disease. Some improvements will be presented in this section.
First of all, a higher number of patients has to be investigated: we presented only one patient. Some features can already be seen from the only patient we had but it could be interesting to enlarge the study and generalize results. Moreover a statistical study requires a high number of data to reach meaningful conclusions.
In fact, the study on EEG is centred on discriminating sleep stages in healthy subjects: it does not take into account disorders and medication effects on sleep. For this reason, collecting EEG and inertial signals from the same subjects during the night could complete the research. In this way, more features could be added for assessing a correct sleep stage and eventually disorders can be detected. In particular, we focused on hypokinesia but disorder as RBD or PLMD are likely to be recognized.
Independently from PD, other disorders can be observed with this system: insomnia and somnambulism are two of them.
It is important to underline that normally more than one EEG channel is used: many channels allow to have signals from different parts of the scalp and consequently more information. Since the system is wanted to be maintained minimally invasive and easily usable, even at home, EEG electrodes can be inserted in a flexible band to promote comfort.
The last proposal in terms of comfortability involves the Sensortile.box casing: a flexible case instead of a plastic one could adapt better to the patient's

body.
Here again, a user interface is proposed for making data visualization immediate and it is reported in Figure 4.1.

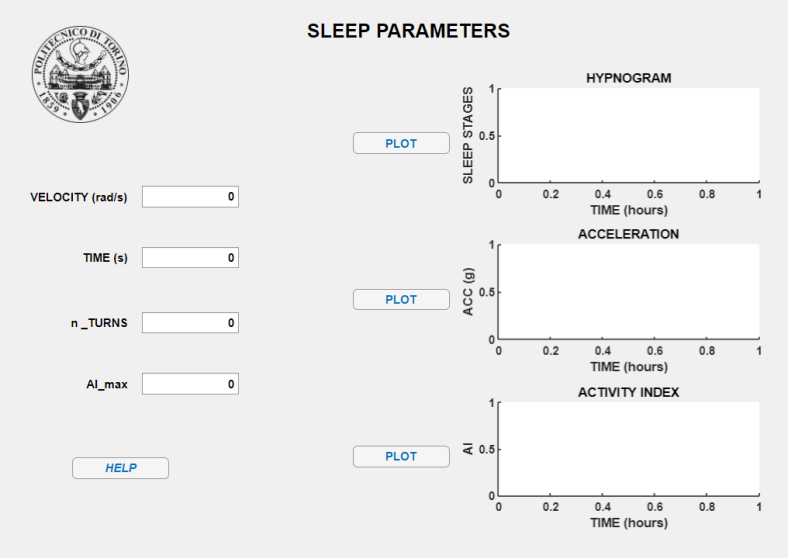


Figure 4.1: User interface layout for all the collected signals

4.2 Conclusions

Behind this work, an intense research on the state of art has been conducted: Parkinson's disease is a widely spread neurodegenerative disease that is still not completely comprehended on every aspect. Non-motor symptoms such as pain, olfactory dysfunction and sleep disorders are relevant in patient's life as well as the most visible and commonly recognized motor-symptoms. Sleep is particularly disturbed and, for this reason, it is the main focus of this thesis; for deepening this topic, a further research has been pursued.

Since sleep can be observed through changes in brainwaves, EEG signal is investigated. The first step has been to distinguish sleep stages because some PD sleep disorders are related to the stage the patient is going through. An already existing database is used: the Sleep-EDF-Database-Expanded contains PSG recordings with relative hypnograms. In the first place, a REM-NREM discrimination has been investigated; specific features are extracted

from the signal and a supervised machine learning algorithm is implemented. The accuracy showed by this algorithm is promising (98.7%): the designated features are based on the Fourier coefficient extracted in every epoch. The algorithm works better in finding the NREM epochs; this is explainable because the dataset presents more NREM epochs than REM, therefore the algorithm has more possibilities on training on NREM epochs.
In the second place, an all-stages recognition is attempted: two path have been explored. One algorithm exploits characteristics found in literature that comprehend statistics, frequency, entropy and complexity features. The reached accuracy is 93.1%. The other, through the wavelet transform, extracts time-frequency features. The final accuracy is 93.5%. The common thread of all these algorithm is the use of three features that allowed a significant increase in accuracy: SEFd, AP and RP.
The used machine learning technique is the ensemble learning with a K-fold validation.
In literature sleep is explored from different point of views: a technique which recently is taking hold is the inertial signals study. SensorTile box has been used to collect data from a PD patient who was hospitalized. To quantify the amount of movements some specific values are calculated: firstly, an algorithm for understanding the patient's position is written. From the position vector it has been extracted the number of turns in bed, the turning maximum velocity and the time for turning. Healthy subjects has been recorded too in order to make a comparison. The results suggest an evident difference between the two categories: the patient turns in bed just one time while healthy subjects do it on average twelve times. This is a clear sign that patient barely moves during the night as can be seen also by the activity index. The activity index allows to have a general idea on how much the subject has moved in every epoch.
By cross-checking EEG and inertial signals, information on the patient's night can be obtained; this is further step that can be realized by collecting inertial and EEG signals from the same patient.

The results obtained in this thesis are encouraging and can be a starting point for future studies projected to improve Parkinson's Disease patients life.

Acronyms

AASM American Academy of Sleep Medicine
ADL Activities of Daily living
AI Activity index
AP Absolute power
AUC Area Under the Curve
CBT Core Body Temperature
CWT Continuous Wavelet Transform
DBS Deep Brain Stimulation
DIST Distal Body Temperture
DTFT Discrete Time Fourier Transform
DWT Discrete Wavelet Tranform
ECG Electrocardiogram
EDS Excessive daytime sleepiness
EEG Electroencephalogram
EMG Electromiogram
EOG ELectrooculogram
ESS Epworth Sleepiness Scale
FD Fractal Dimension
FFT Fast Fourier Transform
FN False Negative
FNE First Night Effect
FOG Freezing of gait
FP False Positive
IMU Inertial Mearsurement Unit
LOC Left Outer Canthus
MNF Mean Frequency
N1, N2, N3 Sleep stages according to AASM
NREM Non-Rapid Eye Movement
OSA Ostructive Sleep Apnea
PD Parkinson's Disease
PET Positron Emission Tomography
PGO Ponto-Geniculo-Occipital
PLMD Periodic Limb Movement Disorder
PPN Pedunculopontine nucleus
PRX Proximate Body Temperature
PSG Polysomnography
R&K Rechtschaffen & Kales

RBD REM Behaviour Disorder
REM Rapid Eye Movement
RFNE Reverse first eye Effect
RLS Restless Leg Movement
ROC Right Outer Canthus
ROC Receiver Operating Characteristic
RP Relative Power
S1, S2, S3, S4 Sleep Stages accordig to *R&K*
SE Shannon Entropy
SEF Spectral Edge Frequency
SNpc Substantia Nigra pars compacta
SOL Sleep Onset Latency
SPECT Single Positron Emission Tomography
ST Standard Deviation
TN True Negative
TP True Positive
TST Total Sleep Time
UPDRS Unified Parkinson's Disease Rating Scale
WSS Wide Sense Stationary

Acknowledgements

We would like to thank our Professor, Gabriella Olmo, that followed us during this path.
She guided and gave us advices to reach the best results, making us comprehend the actual patients' situation affected by Parkinson's disease.
Above all she instilled us passion for the research in this field.

Thanks also to Irene Rechichi that suggested us the right changes to be made to this thesis. She reassured us in our worst moments and we worked as a team during these months.

Bibliography

[1] OB Tysnes, A Storstein, *Epidemiology of Parkinson's disease*, National Istitutes of Health, 2017.

[2] M. Steriade, *Sleep oscillation and PGO waves*, Encyclopedia of Neuroscience, 2009.

[3] M. Steriade, *Brainstem control of wakefulness and sleep* ,Plenum, 1990.

[4] EF. Pace-Shott, *The neurobiology of sleep: genetics, cellular physiology and subcortical network.*,Nat Rev Neuurosci, 2002.

[5] D. Barone, *The Function of sleep*, AIMS Neuroscience, 2015

[6] MV. Sanchez*A social conflict increases EEG slow-wave activity during subsequent sleep.*, Physiol Behav, 2001.

[7] A. Rechtshaffen *Current perspectives on the function of sleep*, Perspect Biol Med, 1998.

[8] MH. Schmidt*the energy allocation function of sleep: A unifying theory of sleep, torpor, and continuous wakefulness*, Neurosci Biobehav Rev, 2014.

[9] M Mackiewicz, KR Shockley*Macromolecule biosynthesis: a key function of sleep.*Physiol Genomics, 2007.

[10] TC Roth, NC Rottenborg *The ecological relevance of sleep: the trade-off between sleep, memory and energy conservation.*Philos Trans R Soc Lond B Biol Sci, 2010.

[11] J Born, B Rasch *Sleep to remember*, Neuroscientist, 2006.

[12] JC Chen, RL Brunner *Sleep duration and risk of ischemic stroke in postmenopausal women*, Stroke, 2008.

[13] KL Chien, PC Chen*Habitual sleep duration and insomnia and the risk of cardiovascular events and all-cause death: report from a community-based cohort*, Sleep, 2010.

[14] JE Gangwisch, D Malaspina *Inadequate sleep as a risk factor for obesity: analysies of the NHANES*, Sleep, 2005.

[15] FP Cappuccio, L D'Elia*Sleep duration and all-cause mortality: a systematic review and a meta-analysis of prospective studies*, Sleep, 2010.

[16] D Moser, P Anderer *Sleep Classification According to AASM and Rechtschaffen & Kales: Effects on Sleep Scoring Parameters*,Sleep, 2009.

[17] RK Malhotra, AY Avidan Sleep Stages and Scoring Technique, Atlas of Sleep Medicine

[18] J Newell, O Mairesse *Is a one-night stay in the lab really enough to conclude? First-night effect and night-to-night variability in polysomnographic recordings among different clinical population samples*,Psychiatry Res, 2012

[19] A Roebuck, V Monasterio, E Gederi, M Osipov, J Behar, A Malhotra, T Penzel, and GD Clifford,*A review of signals used in sleep analysis*, National Istitutes of Health, 2014.

[20] V Lorraine, E Anthony,*Parkinson's disease*, The Lancet, 2015.

[21] Groiss S. J., Wojtecki L., Schnitzler A., *Deep Brain Stimulation in Parkinson's Disease*, Neurological Disorders, 2009.

[22] J Jankovic, M McDermott *Variable expression of Parkinson's disease: a base-line analysis of the DATATOP cohort*, Neurology, 1990

[23] P Martinez-Martin, C Rodriguez-Blazquez*The impact of non-motor symptoms on health-related quality of life of patients with Parkinson's disease* Mov Disord, 2011

[24] DW Dickson*Parkinson's disease and parkinsonism: neuropathology*, Harb Perspect Med, 2012.

[25] MH Polymeropoulos,C Lavedan,E Leroy *Mutation in the alpha-synuclein gene identifi ed in families with Parkinson's disease*, Science, 1997

[26] H Braak, K Del Tredici, U Rüb, *Staging of brain pathology related to sporadic Parkinson's disease*,Neurobiol Aging, 2003

[27] MG Tansey, MS Goldberg, *Neuroinflammation in Parkinson's disease: its role in neuronal death and implications for therapeutic intervention*,Neurobiol Dis, 2010

[28] S Phani, JD Loike, S Przedborski *Neurodegeneration and inflammation in Parkinson's disease*, Parkinsonism Relat Disord,2012

[29] S Lesage, A Brice, *Parkinson's disease: from monogenic forms to genetic susceptibility factors*, Hum. Mol. Genet., 2009

[30] LML De Lau, MMB Breteler, *Epidemiology of Parkinson's disease*, Lancet Neurol, 2006

[31] SK Van Den Eeden, *Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity*, Am J Epidemiol, 2003

[32] JA Driver, G Logroscino, *Incidence and remaining lifetime risk of Parkinson disease in advanced age*, Neurology, 2009

[33] CM Tanner,*Rotenone, Paraquat and Parkinson's Disease*, Environ Health Perspect, 20

[34] AJ Noyce, JP Bestwick, *Meta-analysis of early nonmotor features and risk factors for Parkinson disease*, Ann Neurol, 2012

[35] The National Collaborating Centre for Chronic Conditions, *Symptomatic pharmacological therapy in Parkinson's disease". Parkinson's Disease*, Royal College of Physicians, 2010

[36] H George, H Otto LuÈders, H Jasper, *The ten-twenty electrode system of the International Federation*, International Federation of Clinical Neurophysiology, 1999

[37] R Shamir, A Noecker, C McIntyre, *Deep Brain Stimulation*, Front Young Minds, 2014

[38] Mariana H.G. Mongie, Guglielmo Foffani, Josè Obeso, and Alvaro Sanchez-Ferro,*New Sensor and Weatable Technologies to Aid in the Diagnosis and Treatment Monitoring of Parkinson's Disease*, Annual Review of Biomedical Engineering, 2019.

[39] CH Schenk, SR Bundlie, MW Mahowald*Delayed emergence of a parkinsonian disorder* 38% *of* 29 *older man diagnosed with idiopathicrapid eye movement sleep behaviour disorder*, Neurology, 1996

[40] D Aygun, *Sleep disorders in Parkinson's disease*, from the book "Parkinson's Disease - Understanding Pathophysiology and Developing Therapeutic Strategies", 2018

[41] C Cabrera, Z Munoz-Torres, Y del Rio-Portilla, *Power and coherent oscillations distinguish rem sleep, stage 1 and wakefulness*,Int. J. Psychophysiol.,2006

[42] SA Imitaz, E Rodriguez-Villegas, *A Low Computational Cost Algorithm for REM Sleep Detection Using Single Channel EEG*, Annals of Biomedical Engineering, 2014

[43] S Raschka, *Python Machine Learning*, Packt Publishing, 2015

[44] R De Oliveira, RC Fernandes Araújo, F Barros, *A System Based on Artificial Neural Networks for Automatic Classification of Hydrogenerator Stator Windings Partial Discharges*, Journal of Microwaves, Optoelectronics and Electromagnetic Applications, 2017

[45] JL Rodríguez-Sotelo, A Osorio-Forero, A Jiménez-Rodríguez, *Automatic Sleep Stages Classification Using EEG Entropy Features and Unsupervised Pattern Analysis Techniques*, Entropy, 2014

[46] A Accardo, M Affinito, M Carrozzi, *Use of the fractal dimension for the analysis of electroencephalographic time series*, U Biol Cyber, 1997

[47] BS Raghavendra, D Narayana Dutt, *Computing Fractal Dimension of Signals using Multiresolution Box-counting Method*, International Journal of Electronics and Communication Engineering, 2010

[48] L Mesin. *Introduction to biomedical signal processing*, self-publishing-ilmiolibro, 2017

[49] AR Hassan, MI Hassan Bhuiyan,*A decision support system for automatic sleep staging from EEG signals using tunable Q-factor wavelet transform and spectral features*, Journal of neuroscience methods, 2016

[50] H Matsumoto, R Saito, Y Kakuta, *Sudden death in Parkinso's disease: a retrospective autopsy study*, J. Neurol. Sci., 2014

[51] J Bai, C Di, L Xiao, K R Evenson, *An Activity Index for Raw Accelerometry Data and Its Comparison with Other Activity Metrics*, Journal Plos One, 2016