FACULTY OF ENGINEERING OF THE UNIVERSITY OF PORTO

Sleep at the Wheel: Wearable Sensors for Detection of Drowsy Driving

Duarte de Lima Oliveira Dinis Pereira

DISSERTATION

INTEGRATED MASTER IN BIOENGINEERING

Supervisor: Prof. Luís Paulo Reis, PhD

Second Supervisor: Brígida Mónica Faria, PhD

July 30, 2021

Sleep at the Wheel: Wearable Sensors for Detection of Drowsy Driving

Duarte de Lima Oliveira Dinis Pereira

INTEGRATED MASTER IN BIOENGINEERING

July 30, 2021

Abstract

Drowsy driving is one of the main causes of traffic accidents, with some solutions addressing this problem by giving feedback when the driver is drowsy. However, few solutions tackle the problem in a way that allows for portability and early prevision, through monitoring of the circadian rhythm. This dissertation focuses on solving the portability issue, by developing a system for drowsiness detection during driving. Wearable sensors and intelligent algorithms are used, for a low-cost, portable, automated, and non-intrusive solution. Circadian rhythm and sleep stage concepts are explored, as well as the clinically standardized methods for sleep analysis (polysomnography and sleep questionnaires), and current methods for drowsy driving detection.

With these concepts in mind, the wearable sensors chosen for biosignal acquisition are Empatica's E4 wristband for heart activity acquisition, and Brainlink Pro for brain activity acquisition. Recurring signal processing techniques and evaluation metrics used in similar studies are explored. The proposed solution follows an algorithm flow of signal acquisition, pre-processing, feature extraction, feature selection, and classification. Features used were mainly time-domain-based (e.g., statistics) and time-frequency-domain-based (e.g., wavelet transforms).

The algorithms were trained and validated through the use of a database developed for this specific study, with 13 young adult participants (11 with normal last-night sleep, and 2 without any last-night sleep). Participants were asked to answer to Pittsburgh and SATED questionnaires, after which photoplethysmography and electroencephalhography physiological signals were acquired during driving in a simulation environment. Some limitations of the use of wearable sensors were found, nominally the reduced number of leads/electrodes reducing signal quality, the high initial costs, unstable wireless connections, and the reduced real-time feedback provided by the current iteration of the proposed solution. The proposed solution far surpasses existing ones in its portability, with future work promising improved automation and speed.

Resumo

A sonolência ao volante é uma das principais causas de acidentes de viação, com algumas soluções resolvendo o problema através de feedback quando o condutor está sonolento. Contudo, poucas soluções permitem portabilidade e previsão antecipada, através da monitorisação do ritmo circadiano. Esta dissertação foca-se em resolver o problema da portabilidade, desenvolvendo um sistema para deteção da sonolência na condução. Sensores vestíveis (do inglês wearable sensors) e algoritmos inteligentes são usados, para uma solução de baixo custo, portátil, automática, e não-intrusiva. Conceitos relativos ao ritmo circadiano e estados do sono são explorados, assim como os métodos clinicamente estandardizados para análise do sono (polisomnografia e questionários do sono) e métodos atuais para deteção de sonolência ao volante.

Com estes conceitos em mente, os sensores vestíveis escolhidos para aquisição dos biosinais são a banda de pulso E4 da Empatica para aquisição de atividade cardíaca, e Brainlink Pro para aquisição de atividade cerebral. Tecnologias recorrentes de processamento de sinal e métricas de avaliação utilizadas em estudos prévios são exploradas. A solução proposta segue o fluxo de pré-processamento, extração e seleção de atributos, e classificação. Os atributos usados foram principalmente nos domínios do tempo (por exemplo estatísticas) e tempo-frequência (por exemplo transformadas de wavelets).

Os algoritmos são treinados e validados pelo uso de uma base de dados criada expecificamente para esta dissertação, com 13 participantes jovens adultos (11 com sono na noite passada normal, e 2 de direta). Foi pedido aos participantes que respondessem aos questionários de Pittsburgh e textitSatisfaction, Alertness, Timing, Efficiency and Duration, depois dos quais sinais fisiológicos de fotopletismografia e eletroencefalografia foram adquiridos durante a condução em ambiente simulado. Algumas limitações no uso de sensores wearable foram encontradas, nomeadamente o número de elétrodos reduzido limitar a qualidade do sinal, o alto custo inicial, ligações wireless instáveis, e o reduzido feedback em tempo real providenciado na iteração atual da solução proposta. A solução proposta consegue distinguir-se pela sua portabilidade, com maior automação e velocidade possíveis em trabalho futuro.

Contents

List of Figures

List of Tables

Abbreviations

AC	Alternating Current
AASM	American Academy of Sleep Medicine
AUC/ROC AUC	Area Under Receiver Operating Characteristic Curve
BAC	Blood Alcohol Content
BCI	Brain-Computer Interface
bpm	beats per minute
BT	Body Temperature
BVP	Blood Volume Pressure
CSV	Comma-Separated Values
CSV	Convolutional Neural Networks
CWT	Continuous Wavelet Transform
DC	Direct Current
DTC	Decision Tree Classifier
DWT	Discrete Wavelet Transform
ECG	Electrocardiography
EDA	Electrodermal Activity
EEG	Electroencephalography
EMG	Electromyography
EOG	Electrooculography
HR	Heart Rate
HRV	Heart Rate Variability
IBI	Interbeat Interval
IHR	Instantaneous Heart Rate
KNN	K Nearest Neighbors
LIACC	*Laboratório de Inteligência Artificial e Ciência de Computadores*
ML	Machine Learning
MLP	Multi-Layer Perceptron
PPG	Photoplethysmography
PSD	Power Spectral Density
PSG	Polysomnography
PSQI	Pittsburgh Sleep Quality Index
RBF	Radial Basis Function
REM	Rapid Eye Movement
RR	Respiration Rate
R&K	Rechtschaffen and Kales
SATED	Satisfaction, Alertness, Timing, Efficiency and Duration Questionnaire
SD	Sleep Deprivation
SEM	Slow Eye Movement
SVC	Support Vector Classifier
SVM	Support Vector Machine

Chapter 1

Introduction

This chapter includes the Motivation for the proposed solution, the general and focused dissertation Objectives, and the Document Structure.

1.1 Motivation

In a growing society, sleep restrictions have a negative impact and risks from multiple factors. The driving activity places highly complex perceptual, physical, and cognitive demands on the driver (Sawyer et al., 2012). According to the AASM, being awake for at least 18 hours is the same as someone having a blood alcohol content (BAC) of 0.05%, while being awake for at least 24 hours is equal to having a blood alcohol content of 0.10%. This is higher than the legal limit (0.08% BAC) in the USA. Therefore, methods for detecting sleepiness in driving are under investigation, with promising results.

It is widely known that monotonous or nighttime driving for long periods often lowers driving performance significantly. This contributes to it being one of the leading causes of injuries and deaths from traffic accidents each year (Lin et al., 2014).

Even though a third of our life is spent sleeping (Mancia, 1993), sleeping disorders are very common. 15 to 35% of the adult population complain of sleep quality disturbance (Breslau et al., 1996). Sleep disturbances are also related to higher rates of depression, anxiety disorders, alcohol abuse, or drug abuse.

To measure sleep quality, subjective methodologies can be used, predominately through questionnaires. However, these methods are not enough, since they rely on the self-awareness and honesty of the subject. Then, objective measurements of sleep are required and thus enter polysomnography (PSG). These tests tend to be made in a specialized facility overnight.

To find a response to drowsy driving, a change of paradigms is necessary, in which methods for sleep evaluation need to be substituted for the automatic detection of sleep disturbances or chronic sleep deprivation. This can be achieved by integrating intelligent algorithms that also classify the circadian rhythm from a subject. A low-cost, portable, and non-intrusive solution is ideal, to facilitate everyday usage.

1.2 Objectives

The project "Sono ao Volante 2.0", or Sleep at the Wheel 2.0, proceeds preliminary works (Nogueira et al., 2018) (Oliveira et al., 2020), with the main objective of developing the first prototype of an integrated data system that is non-intrusive and low-cost, which allows sleep prevision while driving and detection of disturbances or chronic sleep deprivation. The investigation aims to advance science and technology in the resolution of the problem "sleep at the wheel".

The contribution from this dissertation focuses on the use of wearable sensors and intelligent algorithms, to conceive in detail the functional and technical architecture of a low-cost, non-intrusive and portable system for sleep quality assessment in a preliminary proposal, and for detection of drowsy driving episodes in a more advanced methodology. To do so, concepts in sleep evaluation, driving monitoring, driving simulation, and signal processing practices must be reviewed.

1.3 Document Structure

The document is divided into six chapters: Introduction, Methods for Sleep Evaluation, Signal Acquisition and Processing, Methodology, Results and Discussion, and Conclusions. Excluding the Introduction and Conclusions chapter, all chapter contain a final Lessons Learned and Limitations section, in which the most important concepts are referenced, as well as current work limitations and suggestions to solve them.

The 1st chapter, Introduction, contains a general overview of the Motivation, with sleep statistics and concepts, as well as the project Objectives and Document Structure.

The 2nd chapter, Methods for Sleep Evaluation, is subdivided into Sleep and Circadian Rhythm, Sleep Stages, Subjective Methods, Polysomnography, and Drowsy Driving Detection. The Subjective Methods section includes a background for the Pittsburgh; and the SATED questionnaires. The Polysomnography section includes the several types of signals used in sleep exams, as well as a detailed overview of the most relevant signals for the dissertation plan.

The Drowsy Driving Detection section originates from the need for a safe and close-to-reality driving environment to test the practicality of wearable sensors. It starts by studying the generalized Marketed Solutions, and then specifies further into BCI-based Studies, and Driving Simulations that can be used in tandem with these studies.

The 3rd chapter, Signal Acquisition and Processing, includes a review on Wearable Sensors of interest for the proposed applications as well as some concepts needed for signal acquisition. This section is subdivided into two: Brain Activity Acquisition; and Heart Activity, Skin Conductivity, and Temperature Acquisition, since the latter list of signals can be measured using similar devices/wristbands. This chapter also includes Signal Processing, which includes a general overview of the algorithm concepts referenced in related literature, and is subdivided in Pre-processing, Feature Extraction, Feature Selection, Classification, and Evaluation Metrics.

The 4th chapter, Methodology, includes the Preliminary Work - Sensor Trial and Sleep Studies, Proposed Solution, Experimental Setup and Procedure, Dataset, Signal Processing Method, Classification, and Evaluation Metrics. The first section explains how the first contact with the sensors was performed, and a proposed solution more focused on signal acquisition during sleep, being subdivided in Preliminary Solution Proposal, Preliminary Methodology and Preliminary Proposed Evaluation. The Dataset section is subdivided in Acquisition and Study Subjects. The Signal Processing Method section is subdivided in Pre-processing, Feature Extraction, and Feature Selection. The Classification section focuses on the classifiers used and their hyperparameters. The Evaluation Metrics section presents some usual metrics used for classification evaluation.

The 5th chapter, Results and Discussion, includes Preliminary Results, Study Participants and Sleep Questionnaires, Practice Run Discrimination, Individual Classification, and Sleep Deprivation Detection. The Preliminary Results section presents some examples of the biosignals acquired during preliminary work, being subdivided into two sections: E4 Wristband and Neurosky Brainlink Pro. The Study Participants and Sleep Questionnaires section provides statistics on the participants of the study, as well as the results obtained from sleep questionnaires. The final three sections explore the results obtained for the different classification methods employed.

The 6th chapter, Conclusions, is subdivided in Current Achievements and Limitations, and Future Work. The first section focuses on the successes and limitations found during the project, while the last section addresses proposes solutions to the current shortcomings.

Chapter 2

Methods for Sleep Evaluation

This chapter provides background for Sleep and Circadian Rhythm, and Sleep Stages, as more theoretical sleep-centered concepts. Subjective and objective methods for sleep evaluation are then tackled, with subjective methods focusing on sleep questionnaires, and objective methods on polysomnography. Drowsy Driving Detection methods are then presented, which include Marketed Solution, BCI-based Studies and Driving Simulations.

2.1 Sleep and Circadian Rhythm

Alertness and reaction time vary according to the circadian rhythm, which makes it an important factor in this study. Living cells in animals have rhythmic variations in their function on a circadian cycle (gan, 2019). If they are entrained, this process usually coincides with day-night light cycles in the environment. If they are not entrained, they become asynchronous from the light-dark cycle. The entrainment process is dependent on the suprachiasmatic nuclei, located above the optic chiasm, bilaterally.

The sleep-wake cycle and the secretion of the pineal hormone melatonin are reliant on neurohormonal signals that participate in this entrainment. According to (Leung and Martinez, 2020), circadian rhythm biomarkers include cortisol levels, peak expiratory flow, blood lipids, DNA damage, lipid peroxidation, protein oxidation, antioxidants, white blood cell counts, estradiol, progesterone, follicle-stimulating hormone, body temperature, blood pressure, and muscle strength. Cellular responses include inflammatory response and cellular trafficking, while some affected molecular processes include oxidative stress responses, DNA methylation, and histone modification. Only signals which are measured during polysomnography exams are considered, as referred in Section 2.4.

2.2 Sleep Stages

Two sleep stage sets of rules can be considered: the ones proposed by Rechtschaffen and Kales (R&K), and the ones proposed by the American Academy of Sleep Medicine (AASM) (Moser

et al., 2009). It is relevant to note that sleep stages are more spectral-like than discrete in their evolution through the sleep cycle, so both rulings can be applied with the proper normative data to make them practicable. With this in mind, the AASM method is the most recent (the R&K rules have been the standard for approximately 40 years), so new normative data for the AASM is still being established.

According to the R&K scoring rules, sleep can be divided in 7 distinct stages: wake, stage 1, stage 2, stage 3, stage 4, stage of Rapid Eye Movement (REM), and movement time. Although this standard can be of interest, it has also been under critics for leaving too much room for subjective interpretation, leading to excess variability between exams due to the reliance on visual analysis of sleep stages. These rules were also developed for a closed demographic (young healthy adults), and do not always apply to certain patients and age groups.

The AASM method iterated the standard guidelines proposed by R&K, with a new guideline for terminology, recording methods, and scoring rules. The review addresses the topics of digital analysis and report parameters, visual scoring, arousal, cardiac and respiratory events, movements, and children and young teenager specific scoring. In the AASM classification, stages S1 to S4 are substituted by non-REM (NREM) N1 to N3, with N3 involving slow-wave sleep (corresponding to stages S3 and S4 in the R&K terminology); and stage REM being presented as stage R.

Nevertheless, normal sleep is widely accepted to be characterized by a succession of four to six cycles lasting around 90 minutes. The early cycles tend to be characterized by deep, slow sleep with slow-wave activity, while the latter corresponds to REM, during which rapid eye movements are observed and muscle tone ceases. The usual duration of sleep tends to vary between six and ten hours, depending on several factors, with the most notable being age and genetics (Guillodo et al., 2018). An example of a typical sleep cycle classified according to the AASM rules is presented in Figure 1.

2.3 Subjective Methods

Subjective methods for sleep assessment are those that, contrarily to physiological signal acquisition methodologies, are reliant on the respondents own testimony. A clinical interview constitutes the most critical first step in the evaluation of sleep related complaints, alongside the physical examination by a professional, which aids the generation of a multi-faceted diagnosis (Chokroverty and Bhat, 2014). This chapter will focus on some existent self-administrated questionnaires used for sleep quality assessment.

2.3.1 Pittsburgh Questionnaire

The Pittsburgh Sleep Quality Index (PSQI) is arguably the most utilized questionnaire for sleep quality assessment (Mollayeva et al., 2016).

The PSQI insides on sleep quality during the previous month (Buysse et al., 1989). This provides information about the night-to-night variations occurring in sleep quality, as well as the

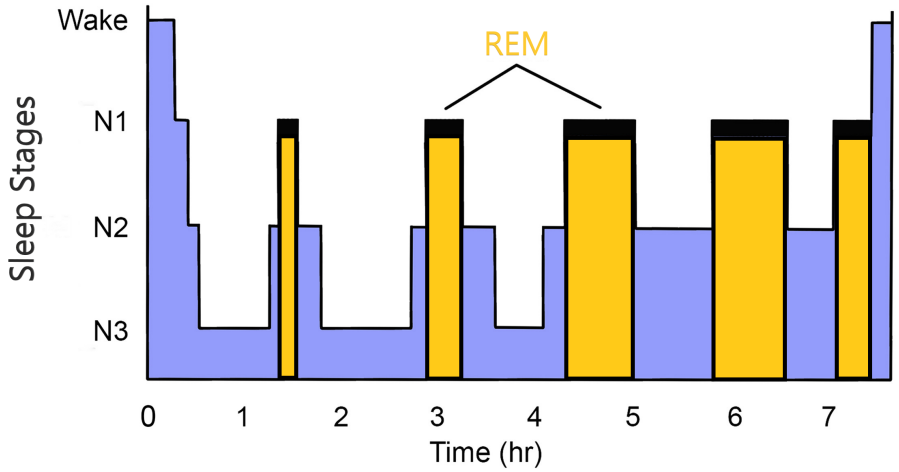


Figure 1: A typical example of a night's sleep for a young adult. The night begins with deep NREM sleep (N3) and then switches between NREM and REM sleep every 90 min approximately. NREM sleep becomes lighter through the night and REM sleep episodes become longer. Modified from Scammell et al. (2017).

duration, frequency, and severity of abnormal behavior duration and frequency over a long period of time.

The PSQI is constituted by 19 self-rated questions and 5 questions rated by the bed partner or roommate. The self-rated questions focus on a vast quantity of factors relating to sleep quality, such as sleep duration, latency, frequency, and severity estimated for each specific sleep issue. The 19 items are grouped into 7 component scores, each weighted from 0 to 3. The seven scores are then added to each other to obtain the global PSQI score, with a range 0-21. Higher scores are associated with worse sleep quality.

The 7 components of the PSQI are subjective sleep quality, sleep duration, sleep latency, usual sleep efficiency, sleep disturbances, use of sleeping medications, and daytime dysfunction. The asked questions and scoring practice are shown in Figure B.1, in the Appendix.

2.3.2 Satisfaction, Alertness, Timing, Efficiency and Duration Questionnaire

A more recent and reliable approach to subjective sleep quality assessment is the Satisfaction, Alertness, Timing, Efficiency and Duration Questionnaire (SATED) (Benítez et al., 2020). SATED evaluates five dimensions of sleep health: 1) satisfaction, 2) alertness while awake, 3) timing, 4) efficiency, and 5) duration. Objective measures can be obtained from every level, excluding satisfaction.

There is a question relating to each of these five dimensions. The subject is asked to provide the frequency of meeting each dimension. Each question can be answered with a level ranging from 0 to 2, with 0 for "never" or "rarely",1 for "sometimes", and 2 for "usually" or "always". Some versions of the questionnaire also divide 0 and 2 into two separate sub-levels.

The total score ranges from 0 to 10 points, from worst to best sleep quality, respectively.

2.4 Polysomnography

Polysomnography (PSG) plays a critical role in confirming suspicions found in more subjective exams and helps guide further diagnosis of sleep disorders (Chokroverty and Bhat, 2014). PSG consists of the overnight recording of various physiological characteristics during sleep. These recordings allow evaluation of sleep stages, alertness, cardiocirculatory functions, respiration, and body movements. Electroencephalography (EEG), electrooculography (EOG), and electromyography (EMG) applied to the chin area are of particular interest for sleep staging (Berry, 2012). EMG measurements of the tibialis anterior may be requested in cases of sleepless leg syndrome (Spiegelhalder et al., 2008). PSG also records single-channel electrocardiography (ECG), snoring (via a sound recorder), movements, and behavior (via sound and video recording) typically. Oronasal airflow is measure in the case of sleep apnea or obstructive lung disease, as well as respiratory effort and chest movement. It is worth noting that sleep apnea can also be detected and classified via EEG and ECG analysis (Koçak et al., 2012) (Lee et al., 2002). Finger or pulse oximetry is also useful in this regard, being helpful in the general identification of hypopneic events in general. Other less widespread techniques may be used, but they are not relevant for the focus of this study.

The biosignal inputs obtained through PSG measurements have very small amplitudes, so amplification is required to obtain interpretable records. Frequency filtering is also pivotal to exclude artifacts and analyze critical features. PSG equipment is then composed of both amplifiers, with the function of amplifying and recording activity, and tunable filters for signal-specific frequency sensitivity.

Low-pass (or high-frequency) filters allow the passage of lower frequencies while attenuating higher frequencies, while high-pass (or low-frequency) filters do the opposite. PSG amplifiers also contain a 60 Hz notch filter, which attenuates the main frequency. This is particularly valuable in the North American power grid, to remove the predominant artifacts, but may also remove important signal features from other components, so it is generally discouraged for use.

Differential amplifiers are used to augment the difference between two electrode inputs, which is valuable for unwanted noise removal. The common-mode rejection ratio represents the ability of noise suppression of an amplifier. In PSG amplifiers, this ratio must exceed 1000 to 1 ideally, but most updated PSG amplifiers use an excess ratio of 10 000 to 1. Alternating current (AC) amplifiers are used to record high frequencies, such as EEG, EOG, EMG, and ECG. Direct current (DC) amplifiers have no low-pass filters, is typically used for low-frequency signals, such as pulse oximetry and pressure readings. Both AC and DC amplifiers can be used for respiratory flow and effort readings.

To facilitate sleep staging, the standard recording speed is 10 mm/sec, so that each monitor corresponds to a 30 seconds epoch. EEG abnormalities, in the form of epileptiform activities, may be better analyzed by slowing the recording to 30 mm/sec, or even 5 mm/sec for a 60 seconds epoch for respiratory event visualization.

Sensitivity (microvolts per millimeter or millivolts per millimeter) regulation is necessary to

obtain the wanted amplitudes for interpretation. Therefore, similarly to and speed settings, the optimal sensitivity changes according to the physiological signal being focused.

2.4.1 Electroencephalography

EEG consists of the study of electrical fields from the brain recorded by amplifying voltage differentials between electrodes typically placed in the scalp (Misra, 2018). EEG readings are usually classified regarding abnormality by comparison with labeled data, and then further classified regarding epileptiform activity. Epileptiform activity in EEG signals may include spikes, sharp waves, or spike-and-wave complexes that can be present not only during a seizure but also shortly before it occurs or between seizures (Gajic et al., 2015).

EEG is mainly used in PSG recordings to distinguish between wakefulness and the various sleep stages (Malhotra and Avidan, 2014).

2.4.1.1 EEG Acquisition

Electrodes conduct electrical potentials from the patient's scalp to the EEG machine, through a conductive interface. This interface constitutes a paste that also stabilizes the electrode, diminishing movement artifacts. The electrode material should constitute a good conductor. Electrode placement rules for EEG follows the international 10-20 system, which assigns a number to each EEG electrode to specify the location in the left or right hemisphere, as seen in Figure 2.a).

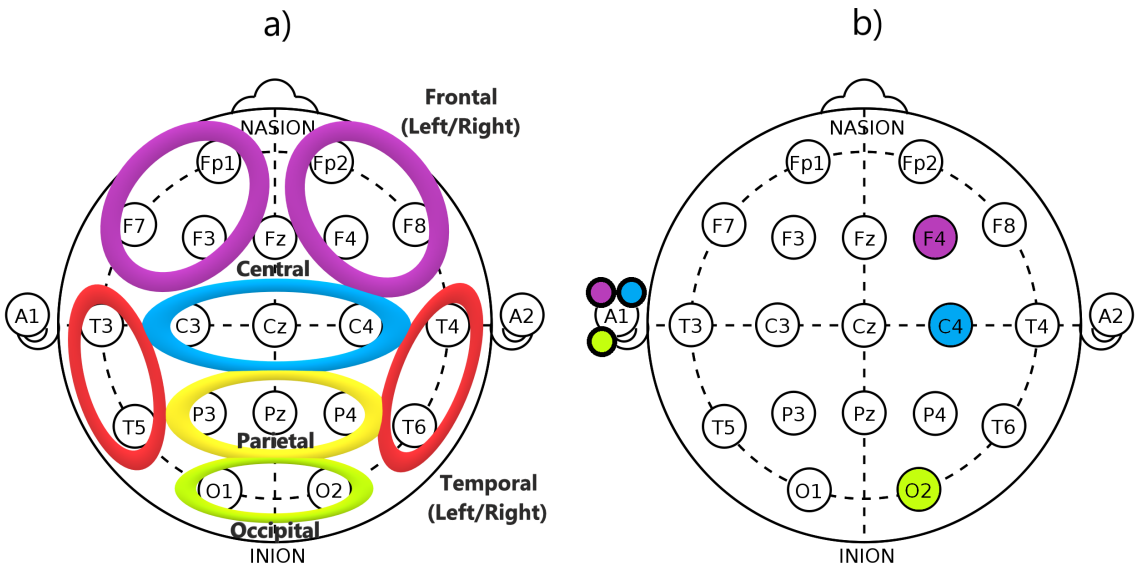


Figure 2: a) 10-20 Electrode Placement Map. The z location indicates the midline. The position Cz corresponds to the vertex, being at the top of the head or in a mid-central position, modified from Valiulis (2014). b) Electrode locations as recommended in 2007 by the AASM, modified from Malhotra and Avidan (2014).

The derivation for sleep studies recommended by the AASM (Ruehland et al., 2011) consists of a minimum of three channels (F4-M1, C4-M1 and O2-M1) representing the right frontal, central, and occipital electrodes, with corresponding backup electrodes over the left hemisphere (F3-M2, C3-M2, O1-M2), all six referenced by the corresponding side's contralateral mastoid electrode, as seen in Figure 2.b).

However, it is worth noting that there are serious limitations to using just this minimum recommended montage. Even though it is sufficient for detection of posterior dominant rhythm in wakefulness and major sleep phenomena (vertex sharp waves, steep spindles, and K complexes; further description available at section 2.4.1.2), a single-hemisphere recording may result in inaccuracies in lesion detection, or if a legion is present in the hemisphere of focus. The AASM also considers acceptable the use of a derivation that includes the midline and a center channel (Fz-Cz, C4-M1, Oz-Cz) (Duce et al., 2014), but this results in attenuation of important hemisphere features. Therefore, a montage recording over the hemispheres and the temporal line is recommended. A full seizure montage is recommended for nocturnal seizure studies, which includes analysis of the temporal and parasagittal regions.

2.4.1.2 EEG Activity during Wakefulness and Sleep

Six EEG wave patterns are conventionally identified for differentiation of wake and sleep states, as well as classification of the different sleep stages. These are alpha activity, theta activity, vertex sharp waves, sleep spindles, K complexes, and slow-wave activity. Further description presented in Table 1.

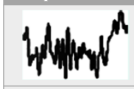
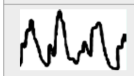
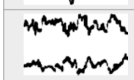
Sample	Definition	Label
	Alpha activity	8-13 Hz waves, tend to be more prominent in occipital leads. Possible origin in the cortex, via a dipole located in layers 4 and 5. Used as a marked for relaxation during wakefulness, as well as central nervous system arousal.
	Theta activity	4-8 Hz waves, most prominent in central and temporal leads. Sawtooth activity, as shown in the sample figure, is a unique variant of theta activity, in which waveforms contain a notched or sawtooth-shaped appearance. Usually prominent during REM sleep.
	Vertex sharp waves	Are most often recorded by central leads placed near the midline. Stand out from the background activity, with sharply countoured, negative bursts.
	Steep spindle	11-16 Hz activity burst, more strongly present in central scalp leads. Typical duration of 0.5 to 1.5 seconds. Presents scalp reflexion of thalamocortical discharges. The name comes from the spindlelike shape.
	K complex	Consists of a well-delineated negative sharp wave closely followed by a positive portion, with a total duration of at least 0.5 seconds. Usually, maximum amplitude is obtained through frontal leads.
	Slow waves	Less than 2 Hz frequencies, variants of delta activity (1-4 Hz), with high amplitudes of at least 75 µV. Appear characteristically during stage N3 sleep.
	REM	Rapid eye movements occur during REM sleep and are linked with the dreamer trying to look at a sensed dream image. Sharply peaked, with an initial deflection of at least 0.5 second duration.
	SEM	Slow eye movements are typically rhythmical rolling eye movements, with an initial deflection of less than 0.5 seconds in duration.

Table 1: Sleep Electroencephalographic Waveforms. (modified from Malhotra and Avidan (2014)).

Several EEG studies also suggest that delta (1-3 Hz), theta (4-7 Hz), and alpha (8-12 Hz) activities are highly related to fatigue, drowsiness, and reduction of task performance (Lin et al., 2014). The filter settings for PSG studies tend to be a low-pass filter of 0.3 Hz, a high-pass filter of 35/70 Hz, and a time constant of 0.4 seconds, with a sensitivity of 5-7 µV/mm (Chokroverty and Bhat, 2014).

2.4.2 Electrocardiography

The PSG often includes a single channel ECG, with one electrode placed over the sternum and the other at a lateral chest region. This montage detects most arrhythmias. Gold cup electrodes are used for the measurement. Sensitivity settings tend to begin at 1 mV/cm, with adjustments required throughout the PSG study. Regarding filter settings, the time constant is set to 0.12 seconds, and high-pass frequency filtering is required at 15 Hz and a low-pass filter at 1.0 Hz (Chokroverty and Bhat, 2014). Single-lead measurements are limited for a more specific diagnosis, so follow up PSG recordings with 12-lead ECGs may be required.

2.4.2.1 ECG Acquisition

ECG conventionally requires twelve leads, which are used to record the difference in potential between electrodes placed on the surface of the body (Loscalzo, 2017). These leads are divided into two distinct groups: six limb (extremity) leads and six chest (precordial) leads. The limb leads record potentials transmitted onto the frontal plane, as seen in Figure 3.a), and the chest leads record potentials transmitted onto the horizontal plane, as seen in Figure 3.b).

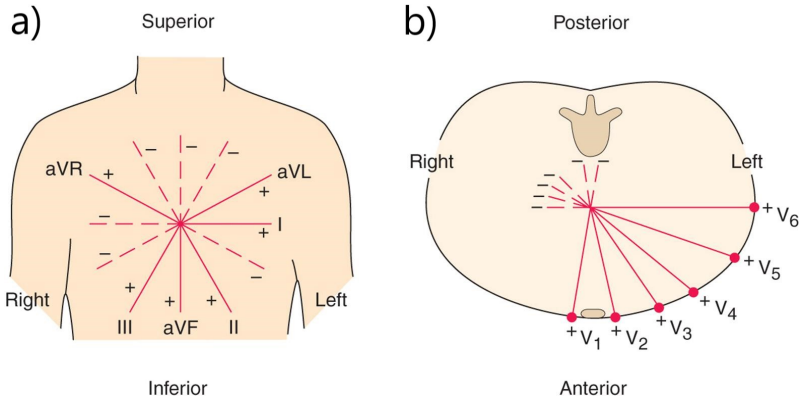


Figure 3: The six frontal planes (a) and six horizontal planes (b) leads allow for a 3D representation of electrical cardiac activity.(modified from Loscalzo (2017)).

2.4.2.2 ECG Waveforms

The ECG waveforms are named alphabetically (Loscalzo, 2017). The P wave represents atrial depolarization. The QRS complex corresponds to ventricular depolarization, and the ST-T-U complex (ST segment, T wave, and U wave) corresponds to ventricular repolarization. The J point

represents the junction between the end of the QRS complex and the start of the ST segment. A standard ECG wave is presented in Figure 4.

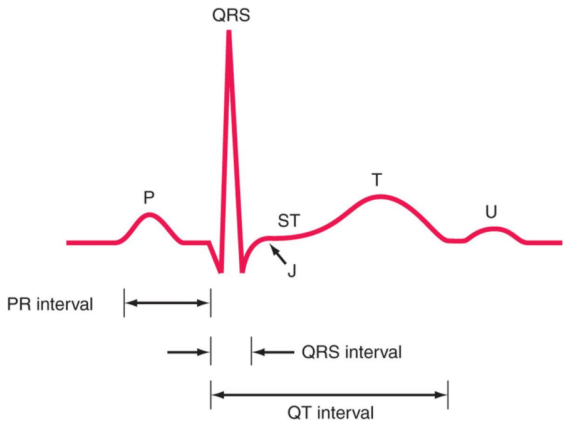


Figure 4: Classic ECG waveform. The R-R interval, corresponding to the time between consecutive QRS complexes, is not shown. (extracted from Loscalzo (2017)).

There are four major ECG intervals: R-R, PR, QRS, and QT. For sleep studies, only R-R and QRS are considered, as further discussed in Section 3.2.2.4. The heart rate (HR) in beats per minute (bpm) can be quickly computed from the interbeat (R-R) interval by dividing 300 by the number of large time units (0.20 s) between consecutive R waves, or 1500 by the number of small units (0.04 s). The QRS interval has a usual duration of 100–110 ms or less and reflects the duration of ventricular depolarization.

2.4.2.3 Important Cardiac Arrhythmias

According to the AASM (Caples et al., 2007), bradycardia during sleep should be defined in people aged over 6 years old as a continual HR < 40 bpm. Furthermore, aystoles (characterized by an absence of ventricular contraction) of duration $\leq$ 2 seconds during PSG can be detected in normal adults. Consequently, longer asystoles (duration $\geq$ 3 seconds) correspond to a pathological indicator in symptomatic individuals and can be an indicator of sleep apnea.

Regarding tachycardias, sinus tachycardia during sleep can be defined in adults as a consistent HR > 90 bpm. Wide complex tachycardia corresponds to a sustained rhythm lasting more than 3 cardiac cycles, with QRS duration $\geq$ 120 ms and HR > 100 bpm. Narrow complex tachycardia is a sustained rhythm that lasts more than 3 cardiac cycles with QRS duration < 120 ms and HR > 100 bpm. Atrial fibrillation can be described as an irregular ventricular rhythm, associated with the substitution of P waves with quick oscillations or waves with an irregular shape, size, and timing.

It is also important to take into account that the average heart rate is lower during sleep than wakefulness, and generally higher in females than males during sleep (Caples et al., 2007).

2.4.3 Photoplethysmography

Plethysmography (PPG) or Blood Volume Pressure (BVP) measures changes in blood volume over time in a body segment. The blood volume increases during systole and decreases during diastole, with arterial inflow and venous outflow, respectively (Graham, 1996).

Finger or pulse PPG can be a good indicator for the prediction of the occurrence of sleep apnea (Karmakar et al., 2014) and sleep stage prediction (Walch et al., 2019).

PPG can be used to determine the interbeat interval (IBI), which is the PPG equivalent to the R-R interval in ECG (Tremper, 1989) (Sinex, 1999). Similarly, the IBI can then be converted to HR and HRV.

For low-movement acquisitions such as during sleep, artifacts that are most often associated with PPG can simply be filtered out with bandpass filters, with ranges in literature around 0.5-5Hz (Ram et al., 2012) (Bagha and Shaw, 2011). Therefore, following the Nyquist theorem, measurements should have a sampling rate of at least 10Hz. For high-movement acquisitions, movement-related noise may need to be removed, with some studies relying on accelerometers to do so accurately (Santos et al., 2012).

2.4.4 Electrodermal Activity

Electrodermal Activity (EDA) consists of the variation in skin conductance in response to sweating. It has been shown to have interesting applications for sleep stage scoring and daily sleep monitoring systems (Hwang et al., 2017), respectively because high-frequency EDA patterns can be ascertained during slow-wave sleep compared with other sleep stages, and skin conductance progressively increases as a subject relaxes.

2.4.5 Temperature

It is well documented that, due to the effect of body temperature on the circadian rhythm, lower environmental temperatures increase drowsiness, with higher temperatures also diminishing the length of slow-wave and REM sleep stages (Sagot et al., 1987) (OKAMOTOMIZUNO et al., 2005) (Muzet et al., 1983). Therefore, body temperature (BP) can be a method for sleep stage classification and drowsiness prevision in this regard, considering that higher BP can be linked to increased wakefulness, while lower BP can be indicators of increased drowsiness.

2.5 Drowsy Driving Detection

2.5.1 Marketed Solutions

Measures for drowsiness detection include vehicle based measures, such as lane position, steering wheel movements, and pressure on acceleration pedal (Sikander and Anwar, 2019). Others are

based on driver behavior, such as eye blink rate, eye closure, yawning, or head position. Physiological based measures can be detected through heart beat (ECG or PPG) or brain activity (EEG). These approaches are shown in Figure 5.

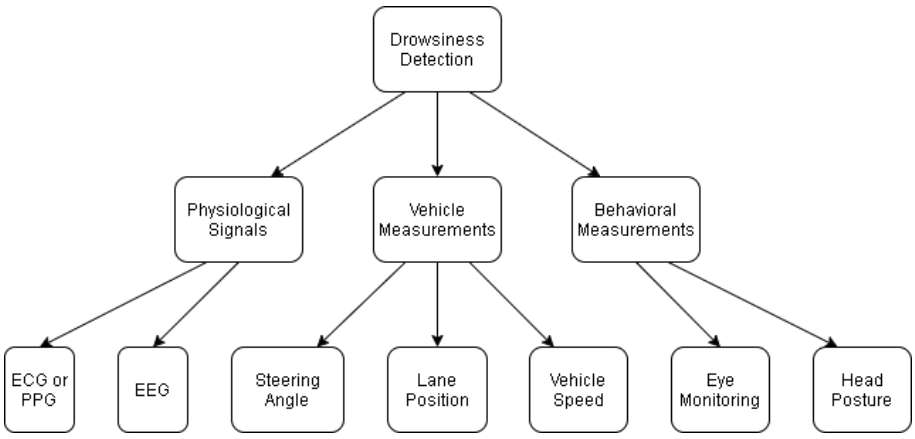


Figure 5: Approaches for Drowsiness Detection. (modified from Sikander and Anwar (2019)).

Multinational automobile companies are currently performing research on driver inattention systems (Sikander and Anwar, 2019).

Toyota's 2019 model Crown implemented a fatigue detection module, which was based on eyelid activity. Toyota Safety Sense P is a newer system, which includes lane deviation.

Nissan implemented a system in the 2016 model Maxima that tracks the driver's steering data and alerts the driver when detecting any unusual deviation.

Volkswagen's Rest Assist system has lane tracking, pedal, and steering wheel pattern recognition. When fatigue is detected, the system warns the driver by visual, auditory, or vibration (in the steering wheel) feedback.

A study using Google Glasses (He et al., 2017) also had promising results, by measuring increased eye blink frequency measurements (through a proximity sensor) in drowsy drivers.

HealthyRoad is a 2016 Portuguese startup in the automotive sector, that focuses on drivers' facial features. The software evaluates drowsiness, fatigue, distraction, and stress. It is also able to identify different drivers, including their sex and age, in order to create a driver's profile. The software's main goals are to alert the driver of dangerous driving behavior, to give a better user experience, and finally to support the other sensors in the vehicle to take action, if necessary.

2.5.2 BCI-based Studies

Alhola and Polo-Kantola (2007) documents a significant cognitive impairment effect of sleep deprivation (SD), and distinguishes between chronic partial sleep restriction (with fewer hours of sleep per night for 4 days or more) and full sleep deprivation (corresponding to 16h-18h-24h-36h without sleeping with increasing SD cognitive impairment). It also puts and emphasis on self-evaluation of cognitive performance with attention tests (i.e., reaction time, vigilance, addition

or subtraction tasks, visual and auditory attention tasks), working/long-term memory tests (recalling words, spaces, numbers), visual-motor performance, decision-making, and response inhibition among other measures.

Regarding driver's cognitive response to turning left or right, Taghizadeh-Sarabi et al. (2013) presents a methodology using EEG and a Hopfield neural network. The signal was acquired by a 19-channel montage during a simulator experiment, in which a driving pathway is presented without obstacles and indications are given to the subject to turn left or right by means of keyboard input. The study achieved a 97.6% accuracy.

When it comes to solutions for emergency braking prevention, EEG usage as shown promising results in reading driver's breaking intention prior to the action being performed (Lee et al., 2017). The acquisition of driving data had a duration of 120 minutes, with two emergency situations and a normal driving situation in a virtual driving environment. The EEG signal was recorded through 64 Ag/AgCl electrodes based on the modified International 10-20. A recurrent convolutional neural network was employed to better capture contextual and spatial patterns of brain signal. The solution was able to recognize the braking intention up to 380 ms prior to the action, with AUC scores range between 0.61 and 0.81.

Although not the same as drowsy driving, driving fatigue also has a tremendous impact on road safety. Hu (2017) develops an EEG-based solution for driving fatigue detection. The proposed solution uses entropy-based features and an AdaBoost classifier. 28 young adults volunteered for the experiment, with an unspecified number of minutes to practice in a driving simulation. The driving procedures were divided according to duration in 20 minutes to 60-120 minutes, with the final 5-minute signal acquisitions being labeled as normal or fatigued, respectively. The driving environment set up for the experiment was a low-density traffic highway, to induce fatigued driving in subjects. EOG response was used to confirm the existence of driving fatigue. The EEG was acquired from 30 channels, with a bandpass filter of 0.15-45Hz during pre-processing. The study achieved improved area under the receiver operating curve of 0.984, recall of 0.984, F1 score of 0.976 and Mathews correlation coefficient of 0.952.

Hernández et al. (2018) also focuses on anticipatory braking recognition to prevent traffic accidents. Seven EEG signals were recorded from the frontal, central and parietal lobes, as well as EMG from the anterior tibial muscle from the right leg. The signals were recorded during a simulation in which participants were asked to avoid collisions with obstacles by performing emergency braking. Stress and fatigue were also factored in the study. The classification between braking intention and normal driving was performed through support vector machines and convolutional neural networks. A recognition of braking intention of 71% was achieved with both classifiers, in addition to 80 and 88% accuracy of classification for SVM and CNN respectively.

2.5.3 Driving Simulations

The need for a safe and close-to-reality driving environment for this dissertation lead to the following review of driving simulations that can be used for dataset acquisition and testing of the practicality of wearable sensors.

Driving simulators present a promising environment for assessment of driving conditions, that facilitates the measurement and recording of driver parameters (Taghizadeh-Sarabi et al., 2013) (Imhoff et al., 2016) (Lee et al., 2017) (Hu, 2017) (Hernández et al., 2018). These simulations have the upside of being safer than real life environments for drowsiness studies.

Studies show a significant impact of drowsiness in driving performance, with these effects being more significant in monotonous or long-duration situations and landscapes (Soares et al., 2020). Some documents include several metrics for the detection and prediction of sleepiness, as mentioned in Section 2.5.1.

Some city driving simulators compatible with portable computer and offering compatibility with simulation hardware (such as *Logitech G* series wheel, pedals, and gearbox) include Forward Development's **City Car Driving** and some Kunos Simulazioni's **Assetto Corsa** mods. These simulations provide replicability through predefined courses and different route and environment settings (such as traffic level, traffic aggressiveness, climate, road type, or pedestrian density).

2.6 Lessons Learned and Limitations

Sleep questionnaires have the disadvantage of being too reliant on subject memory and honesty. Besides, some questions related to sleep (e.g., if the person snores) can often only be answered accurately if a bed partner/roommate is present. Therefore, in order to correctly assess subject baseline sleep quality in the future, there are a few options: bed partners/roommates must be included in the questionnaires, SATED objective measure may be employed, tests may be taken from patients undergoing polysomnography exams in a sleep institution, or subjects may be subject to drowsy driving detection methods available in the market (for example, through EOG signals).

Polysomnography signal acquisition is the current gold standard for circadian rhythm and sleep stage monitoring and evaluation. The monitored indicators include electroencephalography, electrocardiography, photoplethysmography, electrodermal activity, temperature, electromyography, oronasal airflow, and electrooculography. Among these signals, electrodermal activity and temperature have the lowest frequencies, which makes them better suited for long duration studies. EEG, ECG/PPG, EMG, and EOG also appear quite frequently in literature and marketed solutions for drowsy driving detection.

Marketed solutions for the issue of drowsy driving are mostly integrated within automobiles. Therefore, the development of portable solutions, nominally those that measure physiological signals or eye movement, can be of interest for solutions that work in all kinds of vehicles. As of now, no marketed solution provides the portability and simple biosignal acquisition that characterizes wearable sensors. However, a more thorough market analysis is necessary, in order to guarantee that no overwhelming competitors appear in the future.

Overall, research shows promise in EEG usage for anticipatory braking recognition, while methodologies involving turning being relatively absent in literature. Sleep deprivations appear through a chronic restriction (with fewer hours of sleep per night during 4+ days) or full deprivation (18+ hours without sleep during a single night), with both having a negative impact on driving

performance. Furthermore, the usage of wearable devices is still lacking, with most studies utilizing equipment that lack portability and usability but provide more accurate results.

Driving simulations are often used in EEG-based studies, since they provide a safe environment for repeatable and fast data acquisition. These offer repeatability and safety. The need for a safe and close-to-reality driving environment for this dissertation lead to the review of driving simulations that can be used for dataset acquisition and testing of the practicality of wearable sensors. Some promissing simulations include some Kunos Simulazioni's **Assetto Corsa** mods, and Forward Development's **City Car Driving**. The latter presents compatibility with the **Logitech G27** wheel, pedals, and gear box.

More specific detection and/or prevention systems related to drowsy driving, such as braking intention detection, can be of interest in future works.

Chapter 3

Signal Acquisition and Processing

This chapter focuses on background directed to the Wearable Sensors of focus (Brainlink Pro and E4 wristband), and Signal Processing algorithms, including Pre-processing, Feature Extraction, Feature Selection, Classification, and Evaluation Metrics.

3.1 Wearable Sensors

3.1.1 Brain Activity Acquisition

User acceptable and reliable EEG devices for real-time monitoring are still a challenging proposition (Lin et al., 2014). Data acquisition from most EEG recording techniques requires skin preparation and conductive gel to reach optimal electrical conductivity at the interface. These procedures can be slow at the time of application and uncomfortable. Also, the gel may have to be reapplied, since the reading decays in quality as the gel dries out. Therefore, the EEG system needed must be a dry-electrode, both wearable and wireless, facilitating prolonged and portable use. The system must also be able to capture the required brain signals for assessing wakefulness and sleep stages.

According to LaRocco et al. (2020), some promising consumer EEG wearable headsets with Bluetooth compatibility include InteraXon Muse, Neurosky Mindwave, OpenBCI, and Emotiv Epoc and Insight. In Table 2, a comparison between these products is presented, in which Neurosky Brainlink Pro and Muse S were also included.

Even though there is a decent amount of commercial headsets available in the market, a large portion of them lacks in the number of electrodes, since they are more targeted for focus, relaxation, or gaming purposes. The ones with the better characteristics are Emotiv EPOC and Open BCI, but the price of the latter exceeds most consumer capabilities. Therefore, Emotiv EPOC seems to be the best candidate, as it offers a wide range of electrodes, which allow the recording of signals from different brain regions, all at an accessible price counting that the shipping taxes do not increase the cost too much.

Regarding Brainlink Pro, it can be of interest for the proposed solution, since the Fp1-Fp2 channel has been used in literature for sleep stage scoring, nominally by Lucey et al. (2016). This

Device	Electrodes	Sampling Rate	Other Information
InteraXon	Rigid electrode placement	256 Hz	Research Tools for Windows and Linux
Muse v2, S	4 channels: Af7, Af8, Tp9, Tp10	12 bits	Source Developer Kit (SDK) for Android, IOS, Windows **Cost**: 270 € for v2, 380 € for S
Neurosky Mindwave, Brainlink Pro	Rigid electrode placement 1 channel (Mindwave): Afz 2 channels (Brainlink Pro): Fp1, Fp2	512 Hz 12 bits	SDK Available **Cost**: 150 € for Mindwave, 250 € for Brainlink Pro
Open BCI	Up to 16 channels Flexible electrode placement at 35 locations	256 Hz 24 bits	Open-source software, firmware, and hardware **Cost** (in the UK store only, so without shipping included): 450 € for 16 channel, 225 € for 8 channels
Emotiv Epoc, Flex and Insight	Rigid electrode placement Epoc: 14 channels (Af3, F7, F3, Fc5, T7, P7, O1, P8, T8, Fc6, F4, F8, Af4) Insight: 5 channels (Af4, Af4, T7, T8, Pz)	128 Hz 14 bits	Research Tools for Windoes, Mac and Linux **Cost**: 700 € (Epoc), 250 € (Insight)

Table 2: Comparison of consumer EEG headsets (modified from LaRocco et al. (2020)).

study obtained a poor sensitivity of 0.2 for stage N1 due to the lack of occipital electrodes. The study also found that sleep latency and REM onset latency readings were compromised relatively to the PSG diagnosis, as well as sleep disturbance detection (e.g., sleep apnea). However, a strong and substantial agreement ratio with PSG measurements — of 67 % — was verified overall, having particularly found that REM assessment, combined with N2 and N3 sleep and frontal slow wave activity can be well assessed through single-channel means. However, this study did not use automatic means for sleep classification, which introduced subjective factors, due to the use of a limited number of human EEG scorers, biased for standard PSG analysis.

3.1.2 Heart Activity, Skin Conductivity, and Temperature Acquisition

The Empatica E4 wristband has been the target of study for ECG applications that depend on heart rate (HR) measurements, with promising results (Ollander et al., 2016) (Milstein and Gordon, 2020) (McCarthy et al., 2016). The wristband derives heart rate variability (HRV) from Blood Volume Pressure (BVP), which is another designation for PPG. These studies also include electrical conductivity in the skin, peripheral skin temperature, and motion-based activity. Additionally, the E4 possesses internal memory that allows for recording of up to 36 hours, with USB connection to a device needed to recover the data; or a Bluetooth streaming mode that allows for visualization of data in real-time. After recording, the data can be uploaded to the Empatica cloud service and visualized or imported through a web dashboard. The device costs approximately 1 400 €.

3.2 Signal Processing

The general structure found in ML systems for automatic sleep staging follows the phases of pre-processing, feature extraction, and classification. Initially, the user's biosignals are recorded, followed by, a pre-processing stage, which includes filtering and artifact removal for signal enhancement. The resulting signals finally suffer feature extraction to return useful attributes for the classification stage (Aboalayon et al., 2016) (van Wouwe et al., 2011) (Guillodo et al., 2018).Some systems also include dimension reduction and feature selection, to generate new features with low-dimensions derived from the input features.

Numerous single- or multi-channel techniques for automated sleep stage scoring appear in the literature. Some of these studies are reviewed here. According to the detailed review of automatic sleep stage classification, sleep stage characteristics' extraction and sleep disorder detection systems performed by Aboalayon et al. (2016), the most often used features, feature extraction, classifiers, dimension reduction, and feature selection techniques in EEG-based signal processing of sleep studies are listed in Table 3.

a)

Technique	Features
Time Domain	Standard statistics
	Zero crossing
	Hjorth parameters
	Renyi Entropy
	Shannon Entropy
Frequency Domain	Non-parametric analysis
	Parametric analysis
	Spectral entropy
	Itakura distance
	Harmonic parameter
Time-frequency Domain	WT
	STFT
	EMD
	WVD
Complexity measures and non-linear parameters	Lempel-Ziv
	Fractal dimension
	Approximate entropy
	Sample entropy
	Autoregressive
	Hurst exponent
	Permutation entropy

b)

Technique	Technique Variations
ANN	-
Statistical	LDA
	SVM
	Bayesian
Instance base	KNN
Decision tree	DT
Ensemble	Adaboost
	Bagging RF
Clustering	K-means

c)

Technique Variations
Minimum redundancy maximum-relevance (Mrmr)
Sequential methods
Principal component analysis
Linear discriminant analysis

Table 3: Most often used EEG signal processing ML techniques in sleep studies. a) Features and feature extraction techniques b) Classification techniques c) Feature selection and dimension reduction techniques (modified from Aboalayon et al. (2016)).

Recent approaches employ Convolutional Neural Networks (CNNs) as classifiers, with promising results (Mousavi et al., 2019) (Supratak et al., 2017).

3.2.1 Pre-processing

The raw EEG signal has low signal-to-noise ratio, as well as surrounding artifacts that contaminate measurements, nominally blinks and facial movements (EMG). As such, it is important to take

signal processing steps that guarantee cleaning of the signal, so it presents mostly brain activity data. With this is mind, a Butterworth Bandpass filter can be applied. In this filter, the frequency response is very flat, leading to less signal distortion overall, as shown in Figure 6.

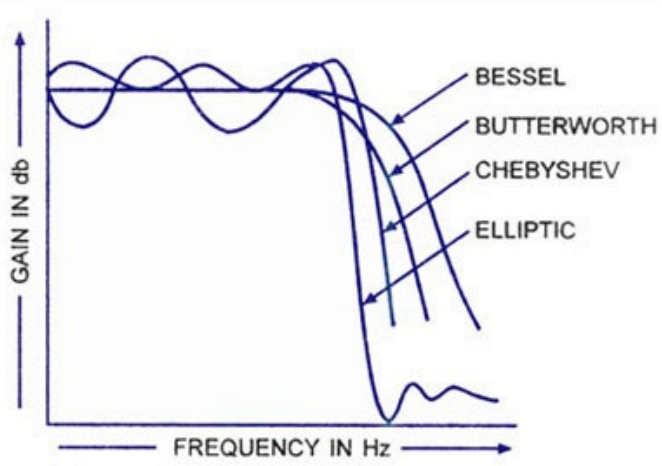


Figure 6: Butterworth bandpass filter response, compared to other filters. (modified from www.CircuitsToday.com, last accessed 21/06/2021)

As mentioned in section 2.4.1.2, the activities of interest are mostly focused in the frequency band between 0.5 and 50 Hz (Lin et al., 2014) (Chokroverty and Bhat, 2014), with 1Hz-30Hz also being a valid interval, depending on the wave components of interest (nominally, alpha and beta waves).

The attenuation of low frequencies allows for the removal of eye blink-related artifacts. For the high cut-off frequency, the upper 50Hz limit allows for power line noise and muscle artifact mitigation, while the 30Hz limit reduces some EEG frequency bands of interest.

However, it is worth noting that EEG wearables are better suited for brain activity related to meditation and attention (Hunkin et al., 2020), which coincides with alpha and beta waves, in the 8-13Hz and 13-30Hz frequency ranges, respectively. Therefore, the 30Hz limit imposed on the signal can better isolate this low-frequency EEG activity.

Regarding the BVP/PPG signal, frequency and shape are the main attributes to be mindful of during processing, with signal amplitude being less relevant. The frequency range of the typical PPG signal is 0.5-5Hz (Bagha and Shaw, 2011) (Islam et al., 2017), so Butterworth Bandpass filtering can be performed with 0.5-10Hz cut-off frequency range, in order to obey to the Nyquist principle. Movement related noise (i.e., abrupt movement of the wrist to which the sensor is attached) can difficult significantly the extraction of relevant information, since the typical signal frequency coincides with movement artifacts. Some strategies available for noise removal include the discrimination of signal sections in which very high peaks appear (due to movement noise), or by using accelerometer data to minimize PPG signal oscillations correlated with high movement.

3.2.2 Feature Extraction

After pre-processing, extraction of relevant information from the signals must be performed. Features can be associated to the Time Domain (recorded signal over time), the Frequency Domain (recorded signal plotted along its frequency components) or the Time-Frequency Domain (non-stationary frequency behavior).

3.2.2.1 Time-Domain Features

Some statistical features can be applied to 1-dimensional arrays. These include mean, standard deviation, kurtosis, quartiles, and range, respectively shown in Equations 1 to 5.

$$\overline{X} = \frac{\sum X_i}{N} \tag{1}$$

$$\sigma = \sqrt{\frac{\sum (X_i - \overline{X})^2}{N}} \tag{2}$$

$$k = \frac{\sum (X_i - \overline{X})^4}{\sigma^4} \tag{3}$$

$$Q_q = \frac{q}{4}(N+1) \tag{4}$$

$$R = max(X) - min(X) \tag{5}$$

where X is the 1-dimensional array, X_i the i^{th} element of X, N the number of elements in X, $overlineX$ the mean, σ the standard deviation, k the kurtosis, Q_q de q-order quartile (q can be 1,2, or 3, and Q_2 is equivalent to the median), and R the range.

Other time-domain features employed include zero-crossings, which can be determined by the number of sign changes in the signal, and entropy analysis, which can be extrapolated from 1D signals, namely Shannon entropy, sample entropy, and spectral entropy (Acharya et al., 2015).

3.2.2.2 Frequency-Domain Features

Regarding frequency-domain features, Power Spectral Density (PSD) measures power over the entire frequency range of the signal. There are several methods to perform this, such as the Welch method (Welch, 1967), the multitaper method (Thomson et al., 2000) and the periodogram (Schuster, 1906). After power spectrum calculation, the spectrum is averaged so the PSD is obtained.

3.2.2.3 Time-Frequency-Domain Features

Time-frequency domain analysis can be performed through the Discrete Wavelet Transform (DWT) or Continuous Wavelet Transform (CWT). Fundamentally, the wavelet transforms consist on shifting a wavelet along a signal and computing the impulse response. The operation should allow only changes in time extension, but not shape. When choosing wavelet parameters, there is a trade-off between time resolution and frequency resolution (Percival and Mondal, 2012) (Samar et al., 1999) (Sweeney, 2013). An example is shown in Figure 7.

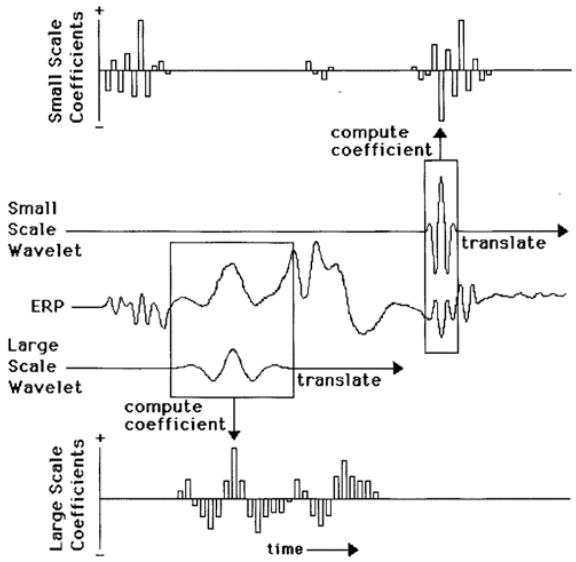


Figure 7: Example of discrete wavelet transform, with different scales. the smaller scale wavelet is good at originating a response from low-frequency regions of the raw signal (ERP), but with less time extension, while the larger scale wavelet detects high-frequency regions of the raw signal, in a larger extension. (sourced from Samar et al. (1999))

Some types of wavelets used for feature extraction in BCI-based studies are presented in Table 4.

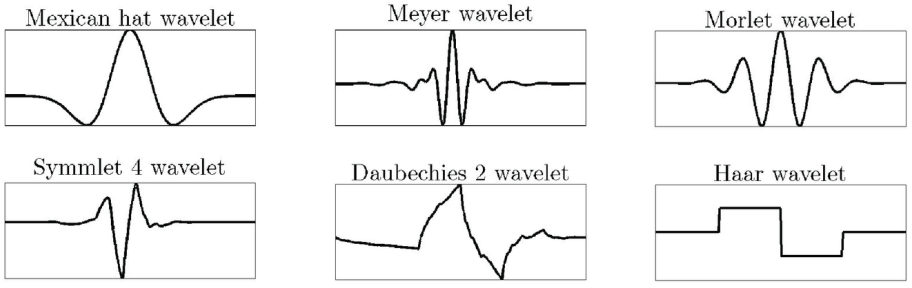


Table 4: Common wavelet types in BCI-based studies. (modified from Sweeney (2013)).

3.2.2.4 Heart Rate

The respiration rate (RR) is a pivotal sign for the detection of sleep-disordered breathing such as apnea. ECG-derived respiration techniques are an alternative to more invasive types of measurements, such as pressure and airflow sensors. The RSA method uses instantaneous heart rate variability to determine a respiratory signal. As described in Helfenbein et al. (2014), this method consists of using instantaneous heart rate (IHR) variability, also known as respiratory sinus arrhythmia. The phenomenon is linked with slight artery responses to pressure resultant from breathing. The IHR wave can be obtained through the frequency domain of the heart rate variability (HRV) frequency-domain. In general, IHR responds rapidly to respiration, increasing during inspiration and drops during expiration.

As illustrated in Figure 8, the QRS complex R-R interval in the ECG is inverted for IHR extrapolation, and then the IHR value is used as the respiration wave amplitude at each beat, as shown by the height of the arrows in Figure 8.b). Using HR instead of R-R inverts the signal, so the signal presents peaks after inspiration, and valleys after expiration. Cubic spline interpolation is then performed to produce the final wave-form.

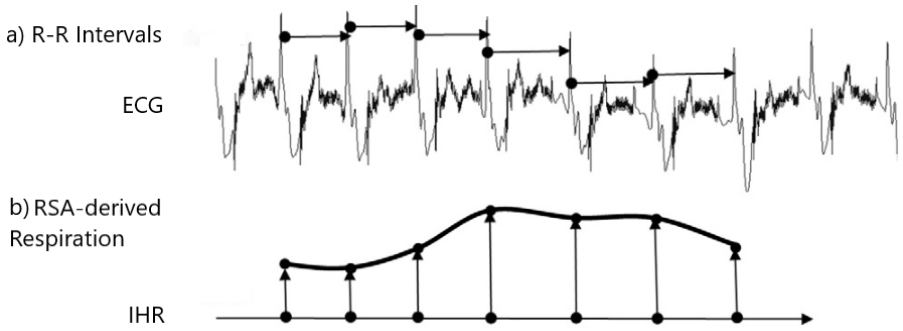


Figure 8: RSA-derived respiration. a) Varying R-R intervals (horizontal arrows), resulting from QRS detection on the ECG. b) RSA-derived respiration wave computed from cubic spline interpolation using the IHR (which was acquired from the inverse of the RRs) as amplitude knots. (modified from Helfenbein et al. (2014))

Similarly, the interbeat interval (IBI) can be obtained from the PPG signal, and extrapolated to a Heart Rate (HR) signal. This is shown in Figure 9. The figure also shows the impact of motion artifacts have on PPG acquisition.

3.2.3 Feature Selection

3.2.3.1 Correlation Coefficient

Correlation shows how related two variables, being expressed numerically by a coefficient (valued between -1.0 and 1.0). A 1.0 coefficient means a perfect match between the two variables, while a -1.0 coefficient means that the two variables have opposite behavior. A zero-valued correlation implies no linear relationship between the variables (Schober et al., 2018).

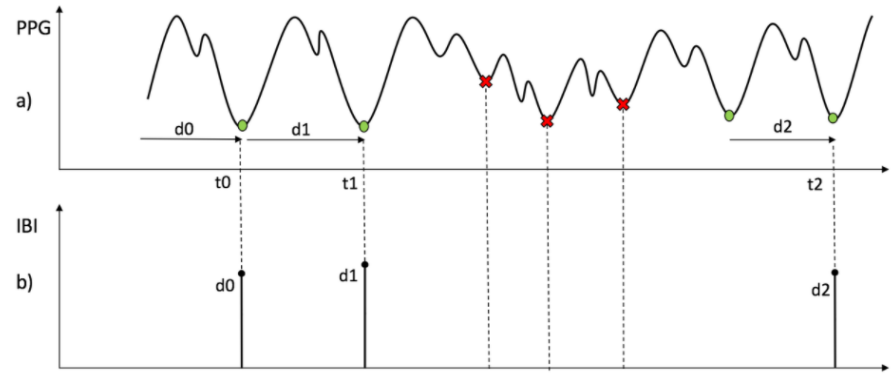


Figure 9: Derived interbeat interval from PPG signal. a) PPG/BVP signal, in which the lowest peaks corresponding to a period of intense movement (red crosses) are discarded, while the good heartbeats (green points) are selected. b) Derived interbeat signal consists of the time duration between beats (d_i) in function of the final time (t_i). (from Empatica web support, last accessed 21/06/2021)

In the feature matrix, the correlation coefficient is used as a threshold for the relationship between the extracted features, allowing for the removal of closely-related ones. The Pearson's correlation coefficient, r, is calculated as follows:

$$r = \frac{\sum(X-X)(Y-Y)}{\sqrt{(X-\overline{X})^2}\sqrt{(Y-\overline{Y})^2}} \tag{6}$$

where $\overline{X}$ is the average of observations of variable X, and $\overline{Y}$ is the average of observations of variable Y.

3.2.3.2 Chi-square Test

The chi-square distribution, χ^2, with k degrees of freedom, is the distribution of a sum of the squares of k independent random variables (McHugh, 2013).

The chi-square test is used in statistics to test the independence between two events, particularly, how the expected count and the observed count deviate from each other. Calculation proceeds as follows:

$$\chi_k^2 = \sum \frac{(O_i - E_i)^2}{E_i} \tag{7}$$

where O corresponds to the observed count, and E corresponds to the expected count. In feature selection, the aim is to select features highly dependent on the response.

Larger χ^2 values correlate with less independent features. Therefore, the higher the χ^2, the more dependent the feature is on the response, so it can be selected for model training.

3.2.4 Classifiers

Some machine learning classification algorithms are listed and briefly explained in this section, followed by background on hyperparameter tuning.

3.2.4.1 K Nearest Neighbors (KNN)

This model tries to find the most separated clusters of points in the feature dimension space, giving each cluster a label. The value K corresponds to the number of labels to be predicted. Nearer points are grouped in the same category (Zhang, 2016).

Weights are assigned to the contributions of the neighbors. A common weighting method consists in giving each neighbor a weight inverse to the distance to the input, so that the nearer neighbors contribute more to the average of the weights than the distant ones.

3.2.4.2 Support Vector Classifier (SVC)

SVC, also called Support Vector Machine (SVM) in literature, is an algorithm that finds the best possible decision boundary between K clusters of data in a number of dimensions equal to the number of features. The boundary is found according to the maximization of the margin, which is the minimum distance between data instances and the boundary (Ben-Hur and Weston, 2009).

Besides being able to perform linear classification, SVMs can be applied in non-linear classification by mapping their inputs into a higher-dimension space (i.e., a kernel).

3.2.4.3 Decision Tree Classifier (DTC)

It is an algorithm that tries to classify the different labels by setting simple rules determined from the input training data. Every rule corresponds to a node in the tree, with the uppermost node being the root node. Every test leads to a different branch. The end result for each sample is the labeled output (Song and Lu, 2015), also called the leaf node. Decision trees can easily be translated to human-level thinking through a flowchart diagram, as shown in the hypothetical decision tree shown in Figure 10.

This is a white box type ML algorithm, meaning that it shares internal decision-making logic, contrary to black box algorithms such as Neural Networks. The decision tree is a non-parametric method, meaning that it does not rely on any probability distribution assumptions.

3.2.4.4 Random Forest Classifier

It comprises an ensemble method for the DTC algorithm (i.e., multiple computation of the DTC). The final output chosen corresponds to the one most frequent in the several DTC previsions, which is found through averaging to improve the predictive accuracy and control over-fitting. Low correlation between the trees is important to test the data according to variable conditions. Each tree can either be trained with different sub-samples, or with the entire dataset.

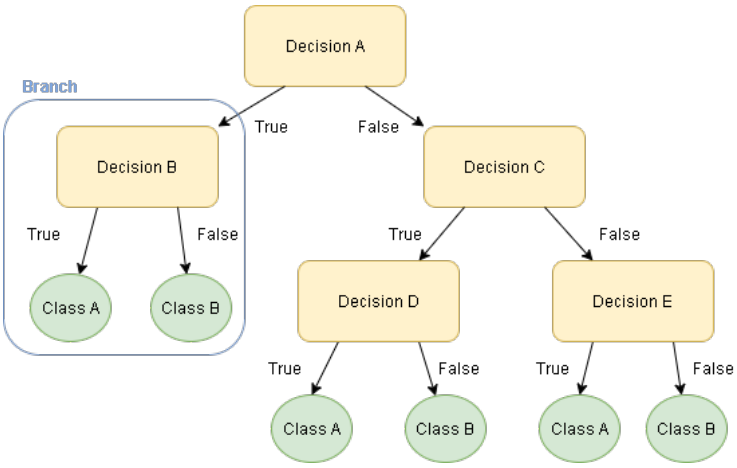


Figure 10: Hypothetical DTC flowchart diagram.

3.2.4.5 AdaBoost

Short for Adaptive Boosting, this algorithm is proposed by Freund and Schapire (1997), being a meta-estimator that begins by fitting a classifier on the original dataset and then fits additional copies of the classifier on the same dataset. This addition is performed with the weight of wrongly classified instances being adjusted, so following classifiers focus more on hard-to-distinguish cases.

AdaBoost is often considered one of the best out-of-the-box classifiers (Kégl, 2013), with decision trees as the sub-learners, since it can detect the specificities of each training data, and adapt itself to it in its sublevels.

3.2.4.6 Gaussian Naive Bayes Classifier

This algorithm utilizes the Bayes Theorem, described as follows:

$$P(C_K|x) = P(C_K)\frac{P(x|C_K)}{P(x)} \tag{8}$$

where C_K is class K, and $x = (x_1, ..., x_n)$ the instance to be classified (vector of length n being the number of features). Therefore, $P(C_K)$ is the probability of any instance being of class C_K, $P(x|C_K)$ is the probability of the instance being x knowing that the class is C_K, $P(x)$ the probability of any instance being x, and $P(C_K|x)$ is the probability of instance x belonging to class C_K.

The Gaussian variant of the Naive Bayes is intended to be applied to continuous data that follows a Gaussian distribution. The likelihood of the features follows a Gaussian distribution. The method also considers independence between classes (Rish, 2001).

3.2.4.7 Gaussian Process Classifier

Gaussian processes are a generalization of the Gaussian probability distribution and can be used as the basis for non-parametric classifications.

Like SVMs, they are a type of kernel model, and unlike SVMs, they are capable of predicting highly calibrated class membership probabilities, although choosing the kernel used as the basis of the method can be challenging. The Gaussian process also requires a link function that interprets the internal kernel representation and predicts class membership probability. The logistic function can be applied, allowing for binary classification.

3.2.4.8 Quadratic Discriminant Analysis (QDA)

Being a more general version of the linear discriminant analysis, in which normal distribution is assumed, but covariance is not considered identical between classes. The classifier uses a quadratic surface boundary to discriminate between classes (Tharwat, 2016).

If in the QDA model, the covariance matrices are assumed to be diagonal and the inputs are assumed as conditionally independent in each class. The resulting classifier is equivalent to the Gaussian Naive Bayes classifier.

3.2.4.9 Logistic Regression

This algorithm determines a threshold to distinguish between different labels (most often 1 or 0) by minimizing the cost function. The sigmoid or logistic function maps any real value into another value between 0 and 1. In machine learning, the sigmoid to map predictions to probabilities. These probabilities can then be compared to a threshold value for classification, as shown in Figure 11.

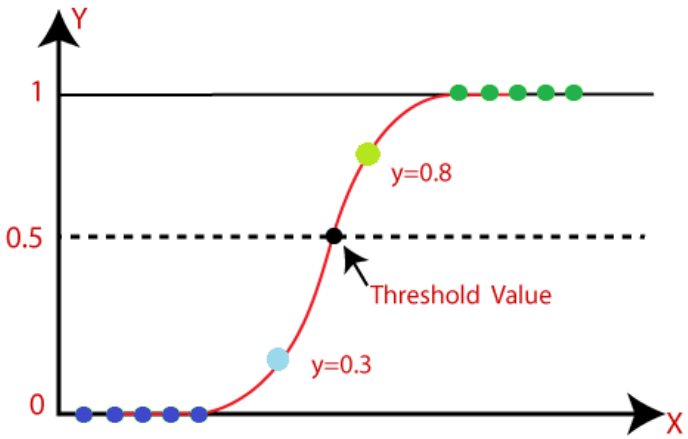


Figure 11: Sigmoid function threshold example. A value between 0 and 1 is given to each data point. The data points above the threshold correspond to one class (green), and the data points below the threshold correspond to another (blue).

3.2.4.10 Multi-Layer Perceptron (MLP)

This model corresponds to a neural network, in which each neuron corresponds to a coefficient determined through the training data. In MLP, neurons are separated by at least three layers: an input layer, an hidden layer, and an output layer. The input layer corresponds to the feature vectors, and the output layer is the label vector. In between these there are the hidden layers, with multiple architectures and nonlinear activation functions, that can be changed to better adapt to the dataset (Murtagh, 1991). A hypothetical MLP structure is presented in Figure 12.

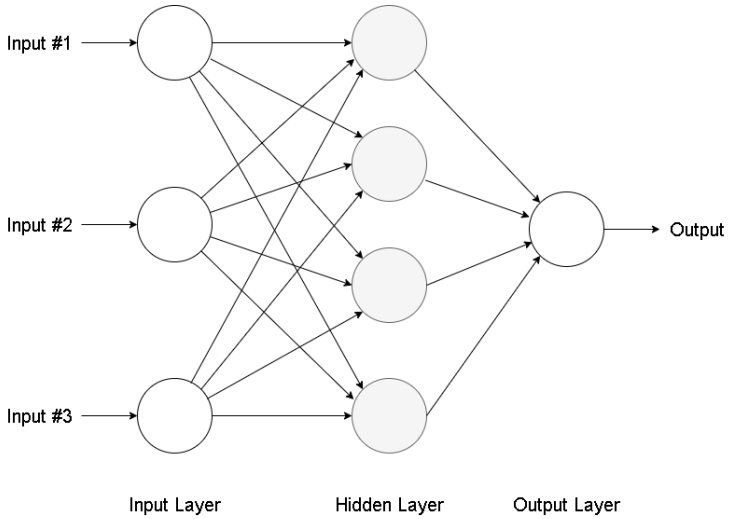


Figure 12: Hypothetical MLP neural network schematic.

In MLP and other neural networks, backpropagation training uses gradient descent of the error function with respect to the neural network's weights, in order to adjust them.

3.2.4.11 Hyperparameter Tuning

Machine learning algorithms tend to allow for flexibility, due to the various scenarios they can be used on. The variation of several hyperparameters shape the model to fit the data as good as possible. A grid search tunes these hyperparameters by iterating over a given list of possible parameters and applying cross validation. The model with the best performance is then chosen.

3.2.5 Evaluation Metrics

Regarding evaluation metrics (Mousavi et al., 2019) (Supratak et al., 2017), some include overall accuracy, precision, recall (sensitivity), specificity, and F1-score. These metrics can be defined as follows:

$$Accuracy = \frac{TP+TN}{TN+FP+FP+FN} \tag{9}$$

$$Precision = \frac{TP}{TN+FP} \tag{10}$$

$$Recall = \frac{TP}{TN+FN} \tag{11}$$

$$Specificity = \frac{TN}{TN+FP} \tag{12}$$

$$F_1 score = 2\frac{Precision \times Recall}{Precision + Recall} \tag{13}$$

where TP indicates the number of correctly labeled data points (e.g., sleep stages), TN (True Negatives) indicates the number of data points correctly classified as not corresponding to other class, FP (False Positives) the number of data points that were incorrectly labeled, and FN (False Negatives) corresponding to the number of data points classified as a different class. Cohen's Kappa coefficient is another useful performance evaluation metric, which measures the agreement between the decision of two entities (human or algorithmic) in the evaluation of the same data.

A receiver operating characteristic curve (ROC) is a graph showing the performance of a classification model at all classification thresholds. The curve plots recall (or true positive rate, TPR) over false positive rate. The false positive rate, FPR, can be defined as follows:

$$FPR = \frac{FP}{FN+TN} \tag{14}$$

A ROC curve plots TPR and FPR at different classification thresholds. Then, the area under the curve (AUC) cam be computed. This consists of the integral calculus between (0,0) and (1,1) of the curve, as shown in Figure 13.

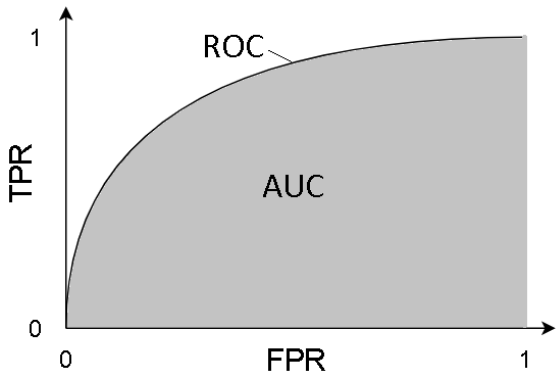


Figure 13: AUC (area under ROC curve).

The AUC value ranges from 0.0 to 1.0. A model that predicts correctly 0% of the time has AUC 0.0, while a model that predicts correctly 100% of the time has AUC 1.0.

3.3 Lessons Learned and Limitations

The E4 wristband costs approximately 1 400 €, which can make it a difficult investment to undertake for stakeholders.

The heart rate is one of the most relevant features that can be extracted from the PPG signal. The removal of movement artifact could be explored in future work, in order to facilitate the extraction of this feature.

Wearable devices for EEG acquisition are more suited to frontal-related activity acquisition, nominally frontal and beta waves, which are highly activated during meditative or attentive states. Further methods (such as pre-processing) have this limitation in mind.

Numerous features can easily be extracted from these methods, particularly the wavelet transforms. This implies not only a considerable computation cost and extraction time, but also a great number of features too similar to each other. Therefore, feature selection is a crucial step for removing correlated features. Plus, through the chi-squared method, the features that are more associated to the given labels can be selected and, therefore, improve the data given to the classifier.

Regarding the classification stage, classifiers are made to be adaptive to each specific situation. Therefore, hyperparameter tuning is an important step in improving classifier compatibility to the dataset.

Chapter 4

Methodology

In this chapter, the preliminary envisioned solution, methodology, and evaluation are presented in Section 4.1, followed by the final Proposed Solution. The latter differs from the system envisioned in the preliminary work, switching its focus from sleep monitoring to drowsy driving detection. This was primarily motivated by the lack of sleep studies occurring during the COVID-19 pandemic. The proposed solution's methodology is divided in Experimental Setup and Procedure, Dataset (including both acquisition and study subject gathering), Signal Processing Method (including pre-processing, feature extraction, and feature selection), Classification and Evaluation Metrics.

4.1 Preliminary Work - Sensor Trial and Sleep Studies

The following section relates to the work developed during the Monography writing.

4.1.1 Preliminary Solution Proposal

The proposed system can provide human interaction with digital devices by translating the different biosignals into a command or diagnosis. The system flow consists of signal acquisition, followed by data processing, which includes pre-processing, feature extraction, and classification, and ending in the output of a control command. The proposed system flow can be observed in Figure 14.

Feature selection and dimension reduction may also be applied to improve already extracted features. It is important to note that for an application that performs during driving (drowsiness prevention), immediate feedback is needed to alert the driver (in the form of visual, auditory, or proprioceptive stimuli). However, for the applications that perform during sleep exams (sleep stage classification and sleep disturbance detection), real-time classification and feedback are not a critical concern, while storing large quantities of raw data may be a requirement.

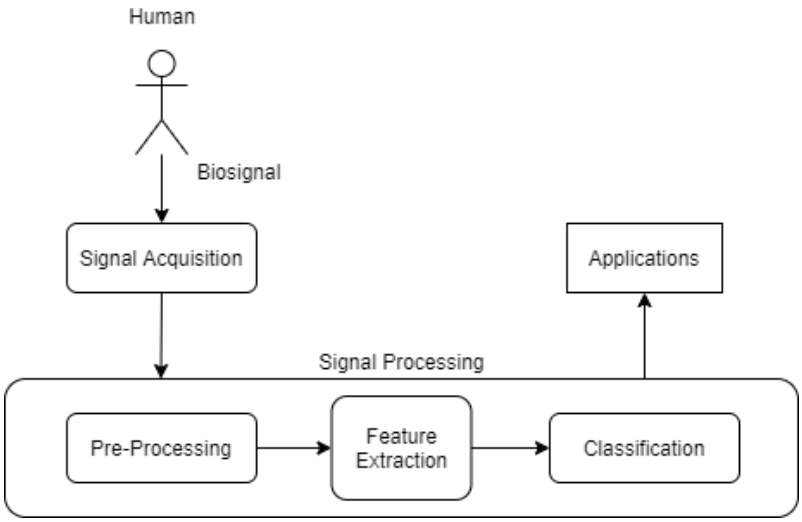


Figure 14: Flow diagram of the preliminary proposed system.

4.1.2 Preliminary Methodology

Regarding the equipment, a personal computer is used for data manipulation, writing, and running scripts. Further work with algorithms demanding more computational power may require the use of a server with access to a Graphical Power Unit (GPU). Considering the current social distance measures, a remote access connection with this machine may be needed.

Regarding the sensors, the dissertation will focus on the EEG headbands **Neurosky Brainwave and Brainlink Pro**, as they are some of the cheapest consumer-available wearable sensors for EEG measurements, and developing sleep analysis and prediction systems with these low-lead sensors could prove an important step for portability and comfort of use during driving or sleep. For the ECG, temperature, and skin conductivity readings, the **E4 wristband** will be used, since it has been previously used in literature for similar applications with promising results, is light and comfortable to use, as well as having an interface especially thought for scientific research. The Empatica cloud is also a valuable asset for this system, for dataset handling purposes.

The EEG signal is obtained via Bluetooth, with support from the Lucid Scribe software, which has been used in other studies using similar EEG wearable devices (Subramanian et al., 2018) (Bone et al., 2017) (Kadar and Iordachescu, 2017), for real-time visualization, recording and exportation of the raw data to a Comma-Separated Values (CSV) file. This file is then read in Python 3.9 through the *csvread* library. The signals from the E4 wristband, as disclosed in Section 3.1.2, can be uploaded directly to the E4 cloud via an intermediary mobile device with Bluetooth compatibility, and then exported through the web dashboard to a CSV file, that can also be read through the *csvread* library in Python for further analysis.

After reading the data and translating it to a Python list/dictionary variable, visualization through *matplotlib* can be performed. Then, pre-processing is performed, according to the methodology disclosed in Section 3.2.1, followed by feature extraction and classification, with possible

dimension reduction and feature selection, according to information shown in Table 3 from Section 3.2. Python libraries of interest for data processing include *pyhrv*, *scikit-learn*, *keras*, *pandas*, *seaborn*, and *numpy*.

4.1.3 Preliminary Proposed Evaluation

The systems for sleep classification and prevention of drowsy driving will be trained and validated with data obtained from a database of 20-30 sleep exams from different subjects (number limited by the current pandemic status), which will be built during development. These exams will be anonymized, to protect the privacy of the patients. The *Laboratório de Inteligência Artificial e Ciência de Computadores* (LIACC) team is discussing with the Porto CUF Hospital the possibility of obtaining more PSG exam data, with 5 being gathered so far, in .xls format, as well as sleep questionnaires targeting Transdev bus drivers. Public PSG datasets, such as the CAP sleep dataset (Terzano et al., 2002), may also be used. If obstacles arise in the gathering of these resources, a sufficient number of sleep exams may be gathered independently of CUF, from voluntary subjects with the proper methodologies and precautions in mind. For the preliminary work, signals extracted from and by the author were used for the development of the acquisition algorithms.

After gathering, processing, and classifying the data, evaluation metrics can be calculated for different feature and classifier groups, as described in Section 3.2.5.

4.2 Proposed Solution

This system can provide human driver interaction with digital devices in the vehicle by translating the different biosignals into a diagnosis of sleep deprivation. The system flow consists of signal acquisition, followed by data processing, which includes pre-processing, feature extraction, feature selection and classification. The proposed system flow can be observed in Figure 15.

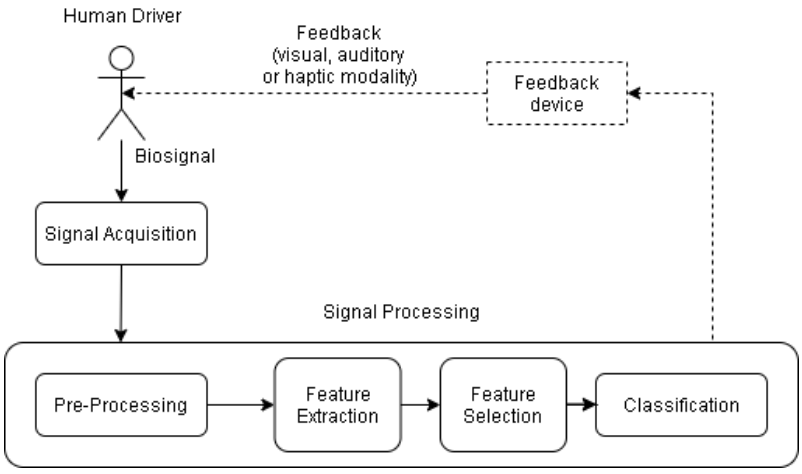


Figure 15: Flow diagram of the proposed system.

Upon detecting the drowsy driving state, an auditory, visual or haptic warning can be given to the driver. However, this isn't developed within this dissertation, with algorithm efficacy being prioritized over real-time feedback.

4.3 Experimental Setup and Procedure

Figures 16 and 17 show the experimental setup.

Figure 16: Experimental setup picture. The larger screen (top-left) displays the simulation, while the smaller screen (bottom-rigtht, corresponding to the main computer) serves as an interface with the acquisition script for Brainlink Pro. To the bottom-left, the Logitech G27 wheel, pedals, and gearbox can be seen. To the left, alcohol and tissue paper are used to clean the sensors, shown slightly at center-left. The technician sits to the right of the driving volunteer during the session.

Volunteers read and sign an agreement of consent (shown in the Figures B.2 and B.3 of the Appendix), and answer to a digital questionnaire including questions from the **Pittsburgh** and **SATED** questionnaires, as well as the amount of hours of slept during the previous night. The questions in the Pittsburgh questionnaire relating directed to the partners were not included for practicality reasons. The complete Google Forms questionnaire is shown in the Appendix C. Volunteers are also asked verbally if they had consumed coffee 12 hours previously to the experiment.

The sensors are placed and adjusted. In the case of **Brainlink Pro**, it is adjusted so as the two forehead electrodes remain pressed to the forehead region, slightly above the eyebrows, with the ground reference attached to the left ear. In the case of the **E4 wristband**, it is adjusted around the left wrist according to Empatica's guidelines. Both sensors are cleaned softly with a small amount of alcohol and tissue between volunteers.

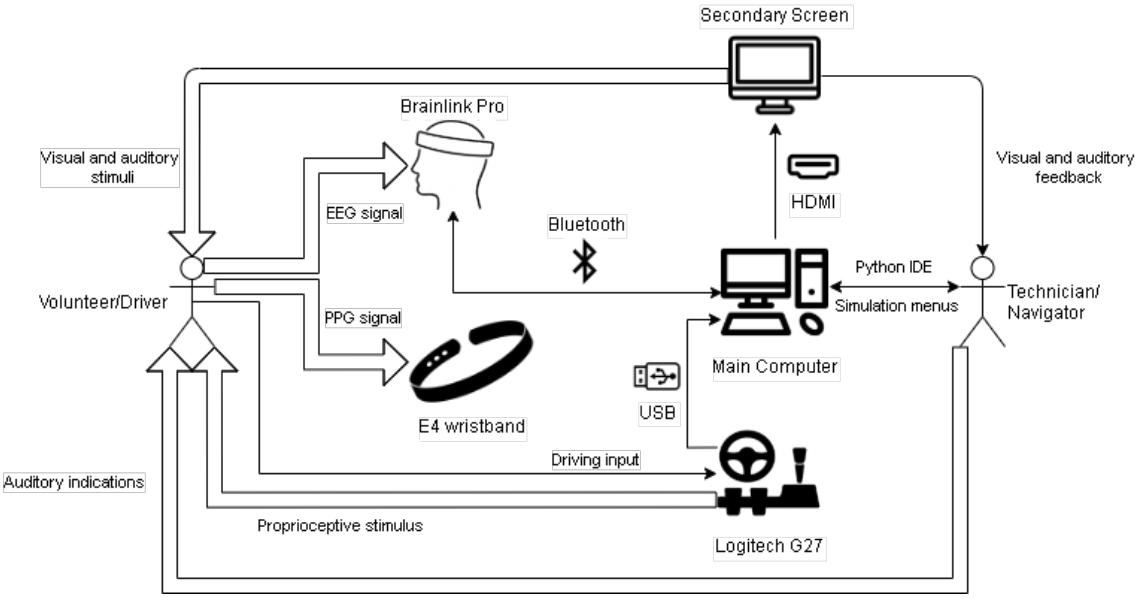


Figure 17: Experimental setup diagram.

The **Logitech G27** driving wheel, gearbox, and pedals are connected via USB to the main computer, with Logitech's **Gaming Software Profiler** running the default calibration sequence. The wheel provides proprioceptive feedback during simulation. The distance to the pedals is adjusted for each volunteer. The chair's height is constant during all acquisitions. A dual display screen is set up, with the larger one presenting the simulation environment to the volunteer and technician, and the smaller one presenting the script and a real-time raw EEG signal graph to the technician. The secondary screen is connected to the main computer via HDMI.

Sequentially, a practice run and an exam run are performed, both with 10-minute duration and within the same route. The practice run serves as a way for the volunteers to get used to the simulation specific conditions, such as controls, as well as the route itself. Signal acquisition is performed during both runs. The sequence for each run is presented in Figure 18.

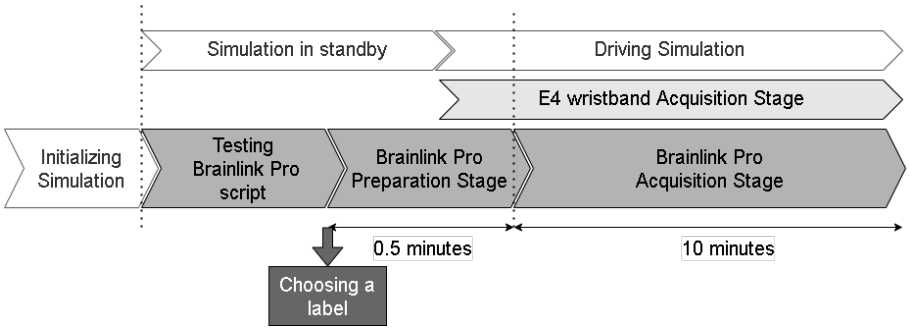


Figure 18: Acquisition sequence for each run (2 times, for practice and exam). The white blocks represent simulation-related activities, the dark-gray blocks represent Brainlink Pro script-based activities, and the light-gray block represents E4 wristband activity.

For both runs, the **City Car Driving** simulation software is run, and the free driving option is selected in the main menu. European Union traffic regulations are selected. In order to reduce the amount of stressful situations presented to the subjects as well as sources of distraction, the optional settings are set to: low traffic density (20%), quiet traffic, 0% pedestrian density, default vehicle, spring, clean weather, daytime, violation pop-ups disabled, fuel consumption disabled, radio disabled, and emergency situations disabled. The view is locked in first person, or in other words, the driver's perspective.

Upon initializing the simulation, Brainlink Pro's acquisition is tested (this could not be performed during loading due to the simulation bugging out), after which acquisition initiates for Brainlink Pro, followed by E4's acquisition (see Section 4.4.1). Finally, the volunteer starts up the vehicle, and follows the instructions provided by the technician. The technician directs the driver verbally according to a previously defined route, as shown in Figure 19.

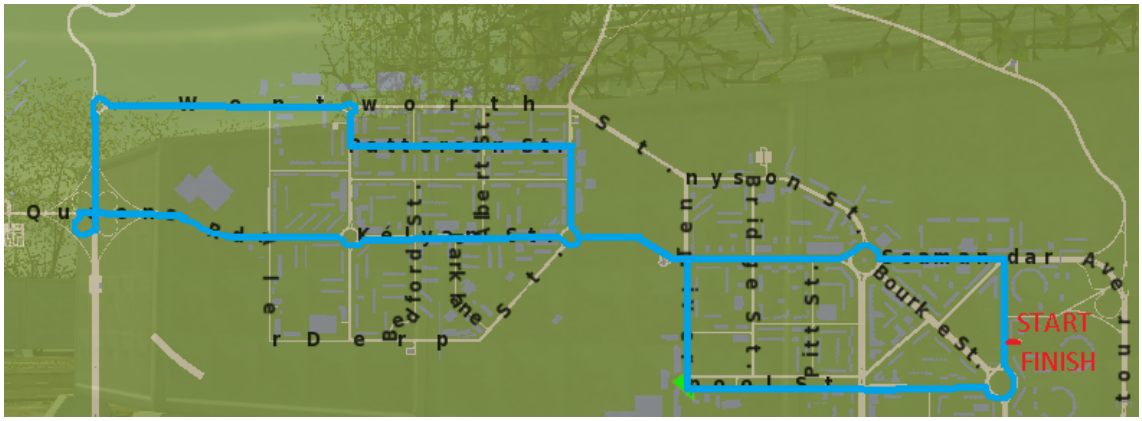


Figure 19: GPS view map within the simulation. The experiment route is marked in blue. The red mark corresponds to the starting location. The initial direction is given by START, following the route counterclockwise, and ending through FINISH, at the starting location. The volunteer does not see this map, with indications being given to him verbally.

The route is the same for every run, in order to limit route-dependent variables such as the number of turns and stops the volunteer would be required to make. It is also designed to last more than the acquisition's 10-minute duration, as well as to provide a wide range of driving situations to volunteers (e.g., roundabouts, traffic lights, highway segments).

The run finishes when Brainlink Pro's script ended, regardless of where the subject is in the route. E4 wristband's acquisition is finished manually soon after. The simulation is restarted between runs, to prevent it from bugging out.

After both runs finished, the volunteer removes the sensors, and answers to a final questionnaire, with questions regarding control, perception and overall difficulties found during the two runs. Some issues found are commented in Section 4.8.

4.4 Dataset

4.4.1 Acquisition

Before initializing the exams, **Brainlink Pro**'s acquisition is tested, in order to guarantee connection and stable acquisition. The sensor is turned on by pressing the side button, which initiates Bluetooth pairing mode. Connection with the main computer is established through the previously configured COM door 3.

After testing, if the acquisition was not working as intended, the test is repeated. Otherwise, a Python 2.7 acquisition script for the Brainlink Pro using the *mindwave* library for Python 2.7 is initiated. The technician inputs the desired label, with the following 30 seconds corresponding to the preparation stage, in which no signal is acquired nor recorded. Afterwards, the 10-minute acquisition is initiated, with a sampling rate of 60Hz, in which the timestamp, raw EEG, blink, attention, meditation, delta, high-alpha, high-beta, low-alpha, low-beta, low-gamma, mid-gamma and theta are recorded into a .csv file with the name of the number of the acquisition (for example, 00.csv for the first exam, 01.csv for the second and so on). For the purposes of this dissertation, only the raw EEG signal is used, due to the low frequency found in other possibly useful signals.

Upon the end of the 10 minutes, the new label is appended into a Labels.csv file, with both the filename and the corresponding label. The script then ends.

Regarding **E4 wristband** acquisition, this is performed via E4's inbuilt recording feature. Acquisition is initiated during Brainlink Pro script's preparation stage. At the end of Brainlink Pro's script, the acquisition is stopped manually. Later, the E4 wristband is connected to a computer via USB, where recorded sessions is uploaded to Empatica's cloud via the *E4 manager* software. Synced sessions can be searched by date, time and duration, as well as visualized and imported from the *E4 connect* website. The imported .zip file, corresponding to the desired session, contains a .csv for each of Empatica's recorded signals: ACC.csv for accelerometer, BVP.csv for blood volume pressure, EDA.csv for electrodermal activity, HR.csv for heart rate, IBI.csv for interbeat interval and TEMP.csv for temperature. Each file also includes the sampling rate for the respective signal in the first line. For the purposes of this dissertation, only the BVP.csv file, with the PPG signal, is used.

The libraries used were: for input/output file handling, *csv*; for path handling, *sys* and *global*; and for data visualization, *matplotlib* (Hunter, 2007).

4.4.2 Study Subjects

The study is being conducted in faculty ground during the morning, during the course of two work weeks. Due to this, subjects are expected to be mostly young adults, aged 21 to 24 years old. Driver's license is mandatory. Regarding education level, most participants are graduated or have completed high school. Participants are acquired through social network advertising.

In the first session, subjects are asked how many hours of sleep they had on the previous night. Non-sleep deprived participants are asked if they want to leave their contact information for the

scheduling of a non-mandatory second session, in which they are sleep-deprived. Participants answer to both the Pittsburgh and SATED questionnaires during the first session through Google Forms. No contact information is associated to the digital forms, in order to maintain subject privacy. They also give feedback about difficulties found during the session. The complete Google Forms questionnaire is shown in the Appendix C.

4.5 Signal Processing Method

For this dissertation, the signal processing algorithms were performed through Python 3.7, with the *PyCharm* IDE. The libraries used were: *pandas* (Wes McKinney, 2010) and *NumPy* (Harris et al., 2020) for data conversion and handling; *sys* and *global* for path handling; *dateutil* and *datetime* for date and time-related operations; *biosppy* (Carreiras, 2015–) for pre-processing; *scipy* (Virtanen et al., 2020), *PyWavelets* (Lee et al., 2019), *mne* (Gramfort et al., 2013), and *pyEntropy* for feature extraction; and *scikit-learn* (Pedregosa et al., 2011) for data normalization, set splitting, feature selection, classification and evaluation metrics.

The files and labels used for signal processing were changed according to one of the three optional classification experiment modes selected: Practice Run Discrimination (in which the labels are "Practice" or "Exam"), Individual Classification (in which the labels are "Individual" or "Other") and Sleep Deprivation Detection (in which the labels are "Exam SD", for sleep-deprived exams, or "Exam", for non-sleep deprived exams).

4.5.1 Pre-processing

The sequence for pre-processing preparation and execution is presented in Figure 20.

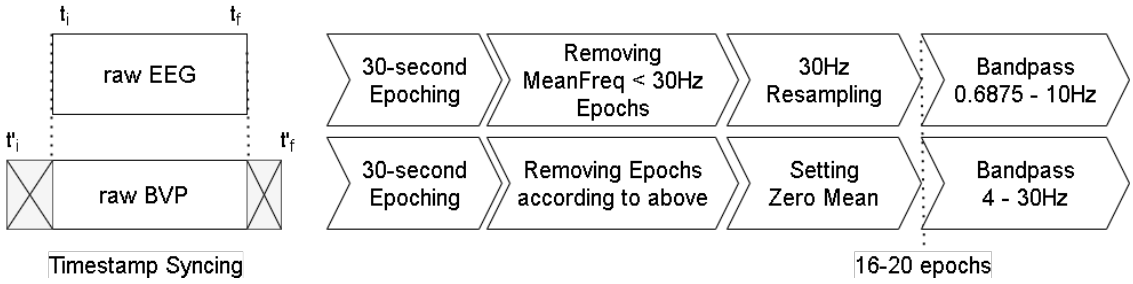


Figure 20: Pre-processing operation sequence. The blocks above represent operations performed to the raw EEG signal, and the blocks below operations performed to the BVP signal. In the first operation, Timestamp syncing, initial and final times for each signal are compared, the non-intersecting signal regions are discarded (crossed out regions). The 16-20 epoch number range presented relates to a single acquisition session.

Before pre-processing can be applied, the data from obtained from the two sensors (BVP/PPG from the E4 wristband and raw EEG from Brainlink Pro) must be synced timewise. To do so, the initial and final timestamps of each signal are compared, and only the intersecting signals are used for the remaining of signal processing.

Following time syncing, the signals are divided into 30-second duration epochs, in order to obtain more samples from the limited dataset, as well as samples that are more manageable for analysis. In a preliminary state, 10-second duration epochs were applied, but this would limit wavelet and heart rate feature extraction further on. Therefore, from an intersecting pair of signals with roughly 9.5-minute duration, roughly 19 (9.5×2) epochs can be obtained. It is worth noting that labels are given to each epoch according to the file from which the epoch originates.

Due to unexpected acquisition issues (low sampling frequency) for some Brainlink Pro sessions, pairs of epochs in which raw EEG frequencies below 30Hz are dominant are discarded. Then, the remaining raw EEG epochs are resampled to 30Hz. After syncing and epoching, the actual pre-processing can be performed to each epoch. Firstly, the raw EEG mean is set to zero. Bandpass filters are applied to both the BVP and EEG signals, with ranges 0.6875-10Hz and 4-30Hz, respectively. No movement-noise filtering is performed at the current iteration.

4.5.2 Feature Extraction

The feature matrix is a list of lists, in which each row corresponds to a sample (or epoch) and each column corresponds to a feature. Feature names are recorded in a separate 1D vector. After extraction, the complete feature matrix was normalized column-wise, to the range 0-1.

4.5.2.1 Statistics

A custom-made function for 1-valued statistic extraction from a 1D array is implemented is several stages of data processing. This function includes: sum of all values, value closest to the mean, values closest to the quartiles (25, 50 and 75%), zero-crossings, standard deviation, kurtosis, range (maximum minus minimum) and entropy. The function also returns the feature names, consisting of the input signal name and the statistic as a suffix. The statistics function is firstly applied to the pre-processed EEG and BVP.

4.5.2.2 Wavelets

Wavelet feature extraction is also performed, using the *PyWavelets* library functions, in both pre-processed epochs. Two main types of wavelet transform are computed: a three-level DWT and a morlet CWT. For the three-level DWT, types cycle between Daubechies 4, Daubechies 20, Coiflet 3, Haar, Symlet 4, and Discrete Meyer. Boundry conditions cycle between zero-padding, symmetrization and smooth padding. For the morlet CWT, widths cycle between the values 10, 15, 48, 72, 80 and 120. The wavelet transforms output coefficient arrays, which are then passed through the 1D statistics function described in Section 4.5.2.1.

4.5.2.3 Heart Rate

For the extraction of the heart rate from the pre-processed BVP, the signal is initially inverted, and peaks are found through a *scipy* library function, with a distance higher than 37. Peaks are

then counted for the 30-second interval. To compute the heart rate for each epoch, Equation 15 is performed.

$$HR = \frac{N_{peaks}}{\Delta t} \tag{15}$$

where HR is the heart rate in beats per minute, Δt the epoch duration in seconds (30 seconds), and N_{peaks} the number of peaks found in the epoch.

4.5.2.4 Power Spectral Density

PSD feature extraction is performed in the pre-processed BVP and EEG signals. Using the *scipy* library, welch and periodogram are extracted, and using the *mne* library, multitaper is extracted, all using the default function parameters. Afterwards, statistics are extracted from the array of powers, using the statistic extraction function described in Section 4.5.2.1.

4.5.2.5 Entropy

For the entropy feature extraction, *pyEntropy* functions are applied to the pre-processed epochs, which allows for extraction of sample, Shannon, and spectral entropy, with default function parameters.

4.5.3 Feature Selection

The selected feature matrix is converted to a *pandas* data frame. The correlation matrix (with dimensions number of features x number of features) is then obtained, consisting of a matrix in which each value is the absolute correlation between the row feature and the column feature. Only the upper triangle (from the diagonal line) of the correlation matrix is used, since values are symmetric to the lower triangle. If any correlation value in a column is inferior to 0.95, that column's index is added to a list. The corresponding columns are dropped from the initial data frame. The resulting data frame of selected features is reconverted to a list of lists.

Afterwards, from this matrix, the 30 best features are selected through the qui-squared test. Thus, the 1893 extracted features were reduced to the 30 best.

4.6 Classification

The classifiers used are Nearest Neighbors, Radial Basis Function (RBF) SVM, Gaussian Process, Decision Tree, Random Forest, Multi-layer Perceptron, AdaBoost, Naive Bayes, QDA, and Logist Regression. The models are initialized through *scikit-learn* functions. The defined constant initial conditions are the RBF kernel and one vs one decision function shape for the SVM, alpha equal to 1 and maximum number of iterations of value 2000 for the Multi-Layer Perceptron, random state value of 0 for the AdaBoost, *liblinear* solver and binary class mode (one versus rest) for the Logistic Regression.

The data is split into train and test sets, after which a hyperparameter grid search is performed. The hyperparameter grid values are shown in Table 5.

Classifier	Hyperparameters
Nearest Neighbors	'n_neighbors': np.arange(1, 10), 'weights': ['uniform', 'distance'], 'algorithm': ['auto', 'ball_tree', 'kd_tree', 'brute']
RBF SVM	'C': [1, 10, 100, 1000], 'gamma': [2, 1, 0.1, 0.01, 0.001, 0.0001]
Gaussian Process	'kernel': [1 * RBF(), 1 * RBF(1.0), 1 * DotProduct(), 1 * Matern(), 1 * RationalQuadratic(), 1 * WhiteKernel()]
Decision Tree	criterion': ['gini', 'entropy'], 'max_depth': np.arange(3, 15)
Random Forest	'n_estimators': [200, 700], 'max_features': ['auto', 'sqrt', 'log2'], 'max_depth': np.arange(3,15)
Multi-layer Perceptron	'activation': ['identity', 'logistic', 'tanh', 'relu'], 'solver': ['lbfgs', 'sgd', 'adam'], 'hidden_layer_sizes': [(1,), (2,), (3,), (4,), (5,), (6,), (7,), (8,), (9,), (10,), (11,), (12,), (13,), (14,), (15,), (16,), (17,), (18,), (19,), (20,), (21,)]
AdaBoost	'n_estimators': np.arange(10, 200, 10)
Logistic Regression	'penalty': ['l1', 'l2'], 'C': [0.001, 0.01, 0.1, 1, 10, 100, 1000]

Table 5: Hyperparameters for each classifier. Values are extracted according to *scikit-learn* classifier-specific documentation, mainly parameter range and possible values.

Finally, each model is trained, and fitting to the test data is performed, in order to obtain a vector of predicted labels for each model.

4.7 Evaluation Metrics

A *scikit-learn* classification report is used to evaluate the quality of predictions from each classification algorithm. The predicted labels are compared to the input/real labels. The report includes precision, recall and F1-score for each given class, as well as the accuracy of the classifier. The macro average (averaging the unweighted mean per label) and weighted average (averaging the support-weighted mean per label, i.e., the mean considering the real difference between class sample sizes) of the previous four values is then calculated. The support value was omitted from the report, as it did not add new information between classifiers (this valued is used as a metric of sample size inequality) As an addition, also using *scikit-learn*, the ROC AUC score is computed and added to the corresponding final report. An example of a final report for a single classifier is shown in Table 6.

4.8 Lessons Learned and Limitations

The preliminary system could still be developed in future work, counting that access to sufficient volunteers performing sleep exams can be established. The system also requires more computing

Classifier	ROC AUC		precision	recall	f1-score
Nearest Neighbors	72	Exam	89	87	88
		Exam SD	76	79	78
		accuracy	84	84	84
		macro avg	82	83	83
		weighted avg	84	84	84

Table 6: Example of report table for a single classifier. All numerical values are shown in percentage (all values range from 0 to 100). The classifier name is displayed in the first column and the ROC AUC value in the second column. The remaining columns correspond to a default classification report, with "Exam" and "Exam SD" being the two labels.

power due to the high duration signal acquisition, so a GPU may be needed going forward. Signal processing could follow and approach similar to the one developed during the dissertation work.

A visual, auditory or haptic feedback device will be needed for a final solution. Adaptation of the algorithms to real-time, or at least near it, will be required. Other constraints such as comfort and battery life of the wearable sensors during long periods of usage will need to be addressed. The development of a portable charging adaptor could be an option.

Regarding the signals used, only the raw EEG from the Brainlink Pro is used, due to an incompatibility issue with the *mindwave* library in Python 2.7 resulting low sampling rates for signals other than the raw EEG. These low frequency regions were discarded prior to pre-processing. Future work would require an acquisition protocol made from scratch, or the use of the software *LucidCode*. However, this software has shown in preliminary work to be very processing heavy, and it would be more so in conjunction with the driving simulation, so two computers or a high-spec one would be required.

For the E4 wristband, the high-movement application of driving, in contrast to sleep, poses the issue of movement-related noise, which introduces artifacts in the frequency band of the PPG signal. This is shown in Figure 9 in Section 3.2.2.4, from Empatica support. In the current iteration, this noise is not extracted from the PPG signal. Due to Empatica's policy, the secondary measures IBI and HR, obtained from the PPG signal, are not computed in high-movement situations. Therefore, both IBI.csv and HR.csv cannot be used in this work, but still have interesting applications when it comes to the preliminary methodology proposal.

Regarding the heart rate feature extraction, as mentioned in the previous paragraph, movement related noise in not filtered off the BVP signal, which may induce an inaccurate measure of heart rate for some samples. Future filtering of movement-noise should be performed in the future, using the accelerometer data provided by the E4 wristband.

More varied features can be implemented in future work, mainly in time-domain, to further avoid feature correlation while maintaining a large amount of features. Regarding the PSD extraction, parameter refinement for welch, periodogram, and multitaper functions may be needed in future work.

The current experiment is still limited to drowsiness detection, instead of its prevision. Further

work should focus on long-time acquisitions or several acquisitions at different times of the day, in order to monitor and classify the circadian rhythm.

Chapter 5

Results and Discussion

This chapter includes the Preliminary Results, with the E4 wristband and Brainlink Pro acquisition. Following this, the section Study Participants and Sleep Questionnaires provides insight on the acquired volunteers, demography, and sleep questionnaire results. The following three chapter focus on three different dataset/label classification reports of evaluation metrics, consisting on Practice Run Discrimination (practice vs exam run), Individual Classification (distinguishing an individual from a non-sleep-deprived dataset) and Sleep Deprivation Detection (distinguishing sleep-deprived exams from normal ones).

5.1 Preliminary Results

5.1.1 E4 Wristband Acquisition

With the Python application, data can also be visualized through the developed Python algorithm, as seen in the examples from Figure 21.

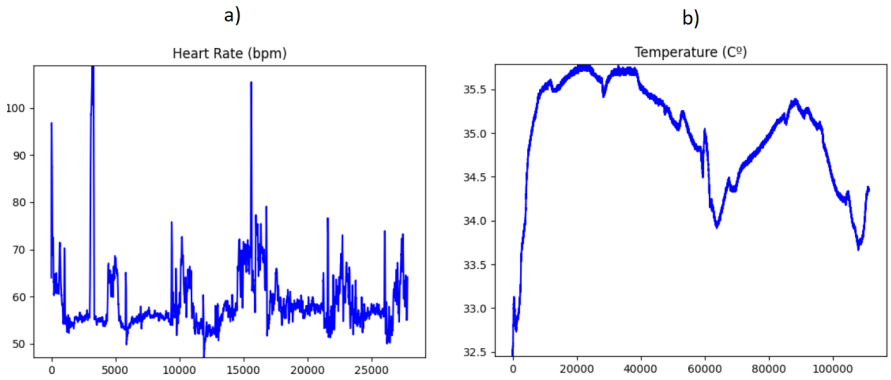


Figure 21: Example of full night's sleep signal acquisition for the E4 in Python. a) Heart Rate in bpm for each data iteration. b) Skin temperature in °C for each data iteration.

From a subjective observation of Figure 22, the medium HR seems to decline and rise in cycles, while the temperature suffers a dip mid-sleep, which is congruent with literature.

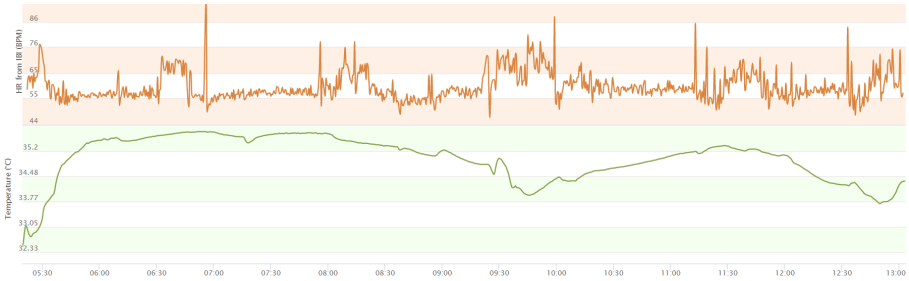


Figure 22: Example of full night's sleep data for the E4 in the Empatica dashboard. Heart Rate (bpm) and Temperature (°C) evolution through time can be visualized, above and below respectively. Values for specific times can be visualized by simply mousing over the corresponding graph region.

5.1.2 Neurosky Brainlink Pro Acquisition

Acquisition works in tandem with the software Lucid Scribe, which allows for real-time data visualization and recording. An example of the Lucid Scribe interface is seen in Figure 23, which presents the raw Neurosky signal (contaminated with blink-related muscle motion), attention and meditation levels (used for entertainment or therapy apps), the EEG signal, Blink Click, and Blink Strength.

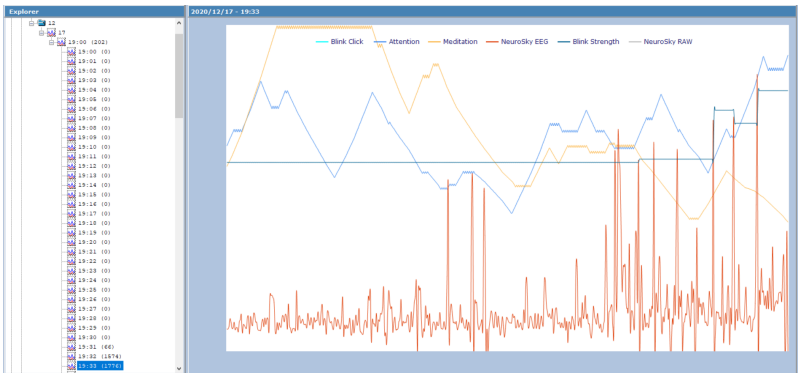


Figure 23: Lucid Scribe recording example from Neurosky Brainlink. Data is recorded in real-time, with w signals or secondary derivations being shown as selected. Each signal or secondary derivation is color labeled. Recorded data can be accessed through the date and time of recording (Left "Explorer" bar).

Other signals of interest not shown in Figure 23 include Alpha, Delta, and Theta waves (interesting for sleep studies, as mentioned in Section 2.4.1.2), as well as the derivation NeuroSky REM.

The recorded data can be read and visualized in a Python application, similar to the example shown in Figure 21 for the E4.

5.2 Study Participants and Sleep Questionnaires

All the participants were young adults with driver's license, aged 21 to 24 years old. For the non-sleep deprived group, there were 11 participants in total, 9 male and 3 female. The education level was mostly graduates, 8 out of 11, with the 3 remaining having completed high school. None of the participants were previously diagnosed with any chronic sleep diseases. Approximately half of the volunteers with good sleep quality claimed to have drunk coffee less than 12 hours before the experiment.

												Mean	STD
Nsleep	7	7	6	5	5	8	7	7.5	10	7	9	7.14	1.45
PSQI	7	7	8	7	8	5	11	4	7	6	4	6.73	1.91
SATED	6	7	3.5	6.5	3.5	6	5.5	6	3	8.5	7	5.68	1.63

Table 7: Normal sleep population questionnaire results. The table includes the number of hours of sleep (Nsleep), PSQI and SATED score for each volunteer, as well as the mean and standard deviation (STD).

SATED and Pittsburgh questionnaire results were calculated in Microsoft Excel. As seen in Table 7, the normal sleep group presented a PSQI mean value of 6.73 and a standard deviation of 2.34, while SATED scores had a mean value of 5.68 and a standard deviation of 1.94. At their extremes, these values are within the range of average sleep quality found in Manzar et al. (2016) and Dalmases et al. (2018), for PSQI and SATED score respectively. The amount of sleep during the night previous to the exams had a mean value of 7 hours, with a standard deviation of 1.5 hours.

Regarding the sleep-deprived dataset, data was obtained from 2 young adults, a female with PSQI of 12 and SATED score of 1.5, and a male with PSQI of 7 and SATED score of 3. As expected from bad sleep quality individuals, PSQI was higher than the normal and SATED was below the normal in the female subject, while the male subject had normal PSQI and bad SATED. Both individuals were awake for more than 24 hours. None of these individuals drunk coffee 12 hours previously to the experiment. Due to the low amount of sleep-deprived subjects, 10 exams were performed by the male individual: 5 under sleep-deprivation and 5 under a normal sleep schedule. The normal sleep sessions were used for the classification performed in a single individual, Section 5.4. This classification served as a comparison to the classification performed in Section 5.5, with the complete non-sleep deprived group and sleep-deprived group.

5.3 Practice Run Discrimination

It is worth noting that the ROC AUC could not be computed for the RBF SVM, due to issues with the probability algorithm used for ROC AUC computation.

Regarding the practice ("Practice") vs exam ("Exam") classification, results are shown in Table A.1 in the Appendix. The discussed results are shown in Table 8.

Classifier	ROC AUC		precision	recall	f1-score	Classifier	ROC AUC		precision	recall	f1-score
AdaBoost	76	Exam	77	59	67	Gaussian Process	76	Exam	68	59	63
		Practice	68	83	75			Practice	65	73	69
		accuracy	71	71	71			accuracy	66	66	66
		macro avg	72	71	71			macro avg	66	66	66
		weighted avg	72	71	71			weighted avg	66	66	66
Multi-layer Perceptron	76	Exam	74	61	67	Decision Tree	59	Exam	71	26	38
		Practice	68	79	73			Practice	56	90	69
		accuracy	70	70	70			accuracy	59	59	59
		macro avg	71	70	70			macro avg	63	58	53
		weighted avg	71	70	70			weighted avg	63	59	54
Random Forest	76	Exam	67	63	65						
		Practice	67	71	69						
		accuracy	67	67	67						
		macro avg	67	67	67						
		weighted avg	67	67	67						

Table 8: Discussed practice run discrimination results.

Nearly all classifiers reached 65% averages and 70% ROC AUC, excluding Decision Tree and QDA. The best classifier was AdaBoost, with averages and accuracy above 70%, and 76% ROC AUC value. Closely behind were the MLP, Random Forest and Gaussian Process, all with the same ROC AUC of 76%, but worse means and accuracy, slightly below 70%.

The Decision Tree classifier was the worst performing in this classification, with the most discrepancy between classes. The values of precision, recall and F1-score were, respectively, 71%, 26% and 38% for the exam data, and 56%, 90% and 69% for the practice data. The macro averages were 63%, 58% and 53% with the weighted averages being nearly identical. The accuracy and ROC AUC for this classifier had both the value of 59%.

5.4 Individual Classification

Regarding the individual exams ("Individual") vs other exams ("Other") classification, results are shown in Table A.2 in the Appendix. The discussed results are shown in Table 9.

For all classifiers, the "Individual" class had more precision than recall, with the opposite being found in the "Other" class. F1-scores were better for the "Individual" class (except in Naive Bayes).

The best performing classifier was Random Forest, with 84% AUC and 78% accuracy and averages. MLP, Gaussian Process, Logistic Regression and Nearest Neighbors all attained ROC AUC of 80%, accuracy of 70-76% and averages in the range 70-78%.

Classifier	ROC AUC		precision	recall	f1-score	Classifier	ROC AUC		precision	recall	f1-score
Random Forest	84	Other	71	75	73	Logistic Regression	81	Other	64	88	74
		Individual	83	80	82			Individual	89	68	77
		accuracy	78	78	78			accuracy	76	76	76
		macro avg	77	78	77			macro avg	77	78	75
		weighted avg	78	78	78			weighted avg	79	76	76
Multi-layer Perceptron	80	Other	58	88	70	QDA	68	Other	62	50	55
		Individual	88	60	71			Individual	71	80	75
		accuracy	71	71	71			accuracy	68	68	68
		macro avg	73	74	71			macro avg	66	65	65
		weighted avg	77	71	71			weighted avg	68	68	68
Gaussian Process	81	Other	64	88	74						
		Individual	89	68	77						
		accuracy	76	76	76						
		macro avg	77	78	75						
		weighted avg	79	76	76						

Table 9: Discussed individual classification results.

The worst performing classifier was QDA, with 68 ROC AUC, accuracy and averages between 65-70%.

5.5 Sleep Deprivation Detection

Finally, for the sleep-deprived exam ("Exam SD") vs non-sleep deprived exam (Exam) classification, results are shown in Table A.3 in the Appendix. The discussed results are shown in Table 10.

Classifier	ROC AUC		precision	recall	f1-score	Classifier	ROC AUC		precision	recall	f1-score
Random Forest	95	Exam	95	87	91	AdaBoost	93	Exam	89	87	88
		Exam SD	79	92	85			Exam SD	76	79	78
		accuracy	88	88	88			accuracy	84	84	84
		macro avg	87	89	88			macro avg	82	83	83
		weighted avg	89	88	89			weighted avg	84	84	84
Naive Bayes	94	Exam	97	82	89	Logistic Regression	90	Exam	86	82	84
		Exam SD	74	96	84			Exam SD	69	75	72
		accuracy	87	87	87			accuracy	80	80	80
		macro avg	86	89	86			macro avg	78	79	78
		weighted avg	89	87	87			weighted avg	80	80	80
Multi-layer Perceptron	92	Exam	88	80	84	Decision Tree	78	Exam	86	80	83
		Exam SD	68	79	73			Exam SD	67	75	71
		accuracy	80	80	80			accuracy	78	78	78
		macro avg	78	80	78			macro avg	76	78	77
		weighted avg	81	80	80			weighted avg	79	78	79

Table 10: Discussed sleep deprivation detection results.

Overall, the "Exam SD" class achieved better recall than precision, with the opposite happening to the "Exam" class. F1-scores were overall better in the "Exam" class.

The obtained results for this classification were the best of all three datasets and labels. The best classifier was Random Forest, with 95% ROC AUC, and accuracy and averages in the range

87-89%. Closely behind where Naive Bayes, AdaBoost, MLP and Logistic regression, with ROC AUC 90-94%, accuracies and averages between 78% and 89%.

The worst classifier was the Decision Tree, with 78% ROC AUC and 76-79% accuracy and averages.

5.6 Lessons Learned and Limitations

Upon analysis of subject feedback, regarding the proposed solution, volunteers found that the experience was mostly close to reality regarding control and perception, and the sensors were comfortable to use. The main complaints presented were about the sensibility of the wheel being too high and the pedals being too difficult to step on, compared to real vehicles. Better calibration of the wheel can be achieved in future work through the Logitech software, with it not being considered for this study due to it changing the experiment conditions between subjects. The problem posed by the tough-to-press pedals can only be corrected through alternative equipment.

The *Oculus Rift* VR headset could be an option for more realistic visual and auditory stimuli, since it is also compatible with the simulation software. However, it would not be practical to use together with Brainlink Pro, and it would require an additional monetary cost for these experiments. A more realistic approach would be to enhance the auditory experience by implementing earphones, and visual perception by switching the display screen to other means more suited to depth perception and larger field of view.

An auditory notification of termination of the Brainlink Pro's acquisition script could be useful in the future. However, real-time visualization is still needed. In order to guarantee sensor connection and correct electrode placement. Automatic checks for these issues could also be of interest.

The low amount of volunteers poses a threat to the statistical validity of further results, mainly when it comes to gender inequality and age range, as well as the very low amount of sleep-deprived individuals. Further work should invest in obtaining more volunteers (at least 30). The sleep questionnaire results support that nearly study participants have sleep habits corresponding to a young adult population, with some scoring worse than the norm.

Regarding classification results, practice run discrimination and individual classification had comparable results to each other, both slightly above average (70 to 80%) regarding their evaluation metric values. When it comes to the practice discrimination, results proved that while the simulation environment had some impact on the performance of at least some subjects, this was somewhat reduced by the 10-minute practice sequence. Regarding the individual classification, results don't allow for completely discarding the effect of the low sample size in the sleep-deprived population.

However, the results found in the discrimination of sleep-deprived exams were significantly better than other dataset-label experiments, which leads us to believe that sleep deprivation classification is possible with the proposed methodology with very good evaluation metrics to back them up (above 90%).

The Decision Tree Classifier was the worst performing overall, which is to be expected due to its simplicity. The best performing classifiers were Random Forest, Naive Bayes, AdaBoost, MLP and Gaussian Process. Further hyperparameter tuning can be useful for future improvement of these classifiers, as well as for attaining usability of the not so successful ones.

Chapter 6

Conclusions

This chapter is subdivided in Current Achievements and Limitations, and Future Work.

6.1 Current Achievements and Limitations

Sleep questionnaires have the disadvantage of being too reliant on the veracity of subject testimony. Polysomnography and marketed solutions for drowsy driving assessment have validated objective results, some of which are derived from physiological signals. However, these solutions lack the portability that a solution integrating wearable devices offers.

The proposed system integrates commonly used algorithms in PPG and EEG-based Machine Learning, obtaining promising results when it comes to the detection of last-night sleep deprivation.

The preliminary solution (integrated with signal processing implemented in the final solution) could also be relevant for sleep exams in general, since the current polysomnography montage is uncomfortable for the subject, and analysis can be quickened and automated, with no significant loss of efficacy. However, PSG is still a clinical gold standard that may be difficult to substitute in the short term.

The proposed wearable devices have few leads/electrodes, which may decrease the quality of the acquired signals and difficult the extrapolation of relevant information from ECG and EEG, but research has shown that interesting results can be obtained from wearable devices. The sensors are costly, particularly the E4 wristband, but it is expected that they become cheaper as research in the field progresses.

The driving simulation revealed itself to be safe, repeatable, fast, and effective in its implementation, with future work addressing the issues found from the volunteers' feedback: wheel sensitivity too high, pedal sensitivity too low, unsatisfactory auditory perception, reduced field of view.

The current iteration has low sample size, due to the low amount of volunteers acquired, which may impair the statistical relevance of the results.

The classification results were promising, with ROC AUC, precision, recall, f1-score, and accuracy reaching values in the range 80-95%, in sleep deprivation detection. The best performing classifiers were Random Forest, Naive Bayes, AdaBoost, MLP and Gaussian Process.

Acquisition issues were found while using the mindwave library with Brainlink Pro, with significant sampling rate drops in some subjects. Heart rate measurements with the E4 wristband were limited due to high movement noise during the experiment.

Overall, the proposed solution far surpasses the current solutions in portability and day-to-day applicability, with most of the previously discussed issues being solvable with further research and work.

6.2 Future Work

Future work should apply to other kinds of sleep deprivation, such as chronic sleep deprivation. Also, it should implement the prediction of the sleep-deprived state through monitoring of the circadian rhythm, in order to differentiate itself further from in-market solutions. Integration of the sleep questionnaires with the objective sleep evaluation methods may be of interest for a future circadian rythm monitoring system.

Real-time feedback should be a worry in the future: for the preliminary solution, because most clinicians are used to work with immediate values or signal waves to perform their diagnosis; and for the proposed solution, because of the urgency to warn the driver of the possible danger. This would require real-time acquisition and visualization methods, with fast classification and feedback. The final solution must also have some sort of auditory, visual, or vibratory feedback to warn the driver of its sleep-deprived state.

Most of the problems found within the simulation can be solved through setting tuning, or ultimately different hardware.

Expanding the dataset, particularly in the sleep-deprived population, is a big must in future work, in order to guarantee the statistical relevance of the results. Further improvement of the results can be obtained through hyperparameter tuning and larger sample size.

Further work with Brainlink Pro must find and alternative software for acquisition. Suggestions include Lucid Code, developed during the preliminary work, or a costum acquisition script.

In a future iteration, movement noise removal must be applied in the PPG signal, in order to provide good basis for interbeat interval, heart rate, and heart rate variability computation.

References

Ganongs review of medical physiology. Mcgraw-hill education, 2019.

Khald Ali I. Aboalayon, Miad Faezipour, Wafaa S. Almuhammadi, and Saeid Moslehpour. Sleep stage classification using eeg signal analysis: A comprehensive survey and new investigation. *Entropy*, 18, 8 2016. ISSN 10994300. doi: 10.3390/e18090272.

U. Rajendra Acharya, H. Fujita, Vidya K. Sudarshan, Shreya Bhat, and Joel E.W. Koh. Application of entropies for automated diagnosis of epilepsy using eeg signals: A review. *Knowledge-Based Systems*, 88:85–96, 2015. ISSN 0950-7051. doi: https://doi.org/10.1016/j.knosys.2015.08.004. URL https://www.sciencedirect.com/science/article/pii/S0950705115003081.

P. Alhola and P. Polo-Kantola. Sleep deprivation: Impact on cognitive performance. *Neuropsychiatr Dis Treat*, 3(5):553–567, 2007.

Sangeeta Bagha and Laxmi Shaw. Article: A real time analysis of ppg signal for measurement of spo2 and pulse rate. *International Journal of Computer Applications*, 36(11):45–50, December 2011. Full text available.

Asa Ben-Hur and Jason Weston. A user's guide to support vector machines. In *Methods in Molecular Biology*, pages 223–239. Humana Press, October 2009. doi: 10.1007/978-1-60327-241-4_13. URL https://doi.org/10.1007/978-1-60327-241-4_13.

Ivan Benítez, Núria Roure, Lucía Pinilla, Esther Sapiña-Beltran, Daniel J. Buysse, Ferran Barbé, and Jordi de Batlle. Validation of the satisfaction, alertness, timing, efficiency and duration (SATED) questionnaire for sleep health measurement. *Annals of the American Thoracic Society*, 17(3):338–343, March 2020. doi: 10.1513/annalsats.201908-628oc. URL https://doi.org/10.1513/annalsats.201908-628oc.

Richard B. Berry. Chapter 1 - sleep stages and basic sleep monitoring. In Richard B. Berry, editor, *Fundamentals of Sleep Medicine*, pages 1 – 11. W.B. Saunders, Saint Louis, 2012. ISBN 978-1-4377-0326-9. doi: https://doi.org/10.1016/B978-1-4377-0326-9.00001-4. URL http://www.sciencedirect.com/science/article/pii/B9781437703269000014.

Daniel Bone, Chi-Chun Lee, Theodora Chaspari, James Gibson, and Shrikanth Narayanan. Signal processing and machine learning for mental health research and clinical applications [perspectives]. *IEEE Signal Processing Magazine*, 34:196–195, 09 2017. doi: 10.1109/MSP.2017.2718581.

Naomi Breslau, Thomas Roth, Leon Rosenthal, and Patricia Andreski. Sleep disturbance and psychiatric disorders: A longitudinal epidemiological study of young adults. *Biological Psychiatry*, 39(6):411 – 418, 1996. ISSN 0006-3223. doi: https://doi.org/10.1016/

0006-3223(95)00188-3. URL http://www.sciencedirect.com/science/article/pii/0006322395001883.

Daniel J. Buysse, Charles F. Reynolds, Timothy H. Monk, Susan R. Berman, and David J. Kupfer. The pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, 28(2):193–213, May 1989. doi: 10.1016/0165-1781(89)90047-4. URL https://doi.org/10.1016/0165-1781(89)90047-4.

Sean M. Caples, Carol L. Rosen, Win K. Shen, Apoor S. Gami, William Cotts, Michael AdamsSUP4/SUP, Parvin Dorostkar, Kalyanan Shivkumar, Virend K. Somers, Timothy I. Morgenthaler, Edward J. Stepanski, and Conrad Iber. The scoring of cardiac events during sleep. *Journal of Clinical Sleep Medicine*, 03(02):147–154, March 2007. doi: 10.5664/jcsm.26816. URL https://doi.org/10.5664/jcsm.26816.

Carlos Carreiras. BioSPPy: Biosignal processing in Python, 2015–. URL https://github.com/PIA-Group/BioSPPy.

Sudhansu Chokroverty and Sushanth Bhat. Chapter 1 - polysomnographic recording technique. In Sudhansu Chokroverty and Robert J. Thomas, editors, *Atlas of Sleep Medicine*, pages 1 – 25. W.B. Saunders, St. Louis, second edition, 2014. ISBN 978-1-4557-1267-0. doi: https://doi.org/10.1016/B978-1-4557-1267-0.00001-1. URL http://www.sciencedirect.com/science/article/pii/B9781455712670000011.

Mireia Dalmases, Ivan D. Benítez, Anna Mas, Oriol Garcia-Codina, Antonia Medina-Bustos, Joan Escarrabill, Esteve Saltó, Daniel J. Buysse, Núria Roure, Manuel Sánchez de-la Torre, Montse Rué, Ferran Barbé, and Jordi de Batlle. Assessing sleep health in a european population: Results of the catalan health survey 2015. *PLOS ONE*, 13(4):e0194495, April 2018. doi: 10.1371/journal.pone.0194495. URL https://doi.org/10.1371/journal.pone.0194495.

Brett Duce, Conchita Rego, Jasmina Milosavljevic, and Craig Hukins. The AASM recommended and acceptable EEG montages are comparable for the staging of sleep and scoring of EEG arousals. *Journal of Clinical Sleep Medicine*, 10(07):803–809, July 2014. doi: 10.5664/jcsm.3880. URL https://doi.org/10.5664/jcsm.3880.

Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, August 1997. doi: 10.1006/jcss.1997.1504. URL https://doi.org/10.1006/jcss.1997.1504.

Dragoljub Gajic, Zeljko Djurovic, Jovan Gligorijevic, Stefano Di Gennaro, and Ivana Savic-Gajic. Detection of epileptiform activity in EEG signals based on time-frequency and non-linear analysis. *Frontiers in Computational Neuroscience*, 9, March 2015. doi: 10.3389/fncom.2015.00038. URL https://doi.org/10.3389/fncom.2015.00038.

A. A. Graham. Plethysmography: safety, effectiveness, and clinical utility in diagnosing vascular disease. *Health Technol Assess (Rockv)*, (7):1–46, 1996.

Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A. Engemann, Daniel Strohmeier, Christian Brodbeck, Roman Goj, Mainak Jas, Teon Brooks, Lauri Parkkonen, and Matti S. Hämäläinen. MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7(267):1–13, 2013. doi: 10.3389/fnins.2013.00267.

Elise Guillodo, Christophe Lemey, Mathieu Simmonet, Juliette Ropars, and Sofian Berrouiguet. Sleep monitoring and wearables : a systematic review of clinical trials and future applications (preprint). 12 2018. doi: 10.2196/preprints.13044.

Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, September 2020. doi: 10.1038/s41586-020-2649-2. URL https://doi.org/10.1038/s41586-020-2649-2.

Jibo He, William Choi, Yan Yang, Junshi Lu, Xiaohui Wu, and Kaiping Peng. Detection of driver drowsiness using wearable devices: A feasibility study of the proximity sensor. *Applied Ergonomics*, 65:473–480, November 2017. doi: 10.1016/j.apergo.2017.02.016. URL https://doi.org/10.1016/j.apergo.2017.02.016.

E. Helfenbein, R. Firoozabadi, S. Chien, E. Carlson, and S. Babaeizadeh. Development of three methods for extracting respiration from the surface ECG: a review. *J Electrocardiol*, 47(6): 819–825, 2014.

Luis G. Hernández, Oscar Martinez Mozos, José M. Ferrández, and Javier M. Antelis. EEG-based detection of braking intention under different car driving conditions. *Frontiers in Neuroinformatics*, 12, May 2018. doi: 10.3389/fninf.2018.00029. URL https://doi.org/10.3389/fninf.2018.00029.

Jianfeng Hu. Automated detection of driver fatigue based on AdaBoost classifier with EEG signals. *Frontiers in Computational Neuroscience*, 11, August 2017. doi: 10.3389/fncom.2017.00072. URL https://doi.org/10.3389/fncom.2017.00072.

Hugh Hunkin, Daniel L. King, and Ian T. Zajac. EEG neurofeedback during focused attention meditation: Effects on state mindfulness and meditation experiences. *Mindfulness*, 12(4):841–851, November 2020. doi: 10.1007/s12671-020-01541-0. URL https://doi.org/10.1007/s12671-020-01541-0.

John D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science Engineering*, 9 (3):90–95, 2007. doi: 10.1109/MCSE.2007.55.

S. H. Hwang, S. Seo, H. N. Yoon, D. W. Jung, H. J. Baek, J. Cho, J. W. Choi, Y. J. Lee, D. Jeong, and K. S. Park. Sleep period time estimation based on electrodermal activity. *IEEE Journal of Biomedical and Health Informatics*, 21(1):115–122, 2017. doi: 10.1109/JBHI.2015.2490480.

S. Imhoff, M. Lavallière, N. Teasdale, and P. Fait. Driving assessment and rehabilitation using a driving simulator in individuals with traumatic brain injury: A scoping review. *NeuroRehabilitation*, 39(2):239–251, Jun 2016.

Mohammad Tariqul Islam, Ishmam Zabir, Sk. Tanvir Ahamed, Md. Tahmid Yasar, Celia Shahnaz, and Shaikh Anowarul Fattah. A time-frequency domain approach of heart rate estimation from photoplethysmographic (ppg) signal. *Biomedical Signal Processing and Control*, 36:146–154, 2017. ISSN 1746-8094. doi: https://doi.org/10.1016/j.bspc.2017.03.020. URL https://www.sciencedirect.com/science/article/pii/S1746809417300721.

Manuella Kadar and Grigore-Dan Iordachescu. Smart testing environment for the evaluation of students' attention. *Interaction Design and Architecture(s) Journal - IxDA*, 2:205–217, 07 2017.

C. Karmakar, A. Khandoker, T. Penzel, C. Sch?bel, and M. Palaniswami. Detection of respiratory arousals using photoplethysmography (PPG) signal in sleep apnea patients. *IEEE J Biomed Health Inform*, 18(3):1065–1073, May 2014.

Onur Koçak, Tuncay Bayrak, Aykut Erdamar, Levent Ozparlak, Ziya Telatar, and O. Erogul. *Automated Detection and Classification of Sleep Apnea Types Using Electrocardiogram (ECG) and Electroencephalogram (EEG) Features*. 01 2012. ISBN 978-953-307-902-8. doi: 10.5772/22782.

Balázs Kégl. The return of adaboost.mh: multi-class hamming trees, 2013.

John LaRocco, Minh Dong Le, and Dong-Guk Paeng. A systemic review of available low-cost EEG headsets used for drowsiness detection. *Frontiers in Neuroinformatics*, 14, October 2020. doi: 10.3389/fninf.2020.553352. URL https://doi.org/10.3389/fninf.2020.553352.

Gregory R. Lee, Ralf Gommers, Filip Wasilewski, Kai Wohlfahrt, and Aaron O'Leary. Pywavelets/pywt: Pywavelets v1.0.3, 2019. URL https://zenodo.org/record/2634243.

Jong-Min Lee, Dae-Jin Kim, In-Young Kim, Kwang-Suk Park, and Sun I. Kim. Detrended fluctuation analysis of eeg in sleep apnea using mit/bih polysomnography data. *Computers in Biology and Medicine*, 32(1):37 – 47, 2002. ISSN 0010-4825. doi: https://doi.org/10.1016/S0010-4825(01)00031-2. URL http://www.sciencedirect.com/science/article/pii/S0010482501000312.

Suk-Min Lee, Jeong-Woo Kim, and Seong-Whan Lee. Detecting driver's braking intention using recurrent convolutional neural networks based eeg analysis. In *2017 4th IAPR Asian Conference on Pattern Recognition (ACPR)*, pages 840–845, 2017. doi: 10.1109/ACPR.2017.86.

Jacqueline M. Leung and Micaela E. Martinez. Circadian rhythms in environmental health sciences. *Current Environmental Health Reports*, 7(3):272–281, July 2020. doi: 10.1007/s40572-020-00285-2. URL https://doi.org/10.1007/s40572-020-00285-2.

C. Lin, C. Chuang, C. Huang, S. Tsai, S. Lu, Y. Chen, and L. Ko. Wireless and wearable eeg system for evaluating driver vigilance. *IEEE Transactions on Biomedical Circuits and Systems*, 8(2):165–176, 2014. doi: 10.1109/TBCAS.2014.2316224.

Joseph Loscalzo. *Harrisons cardiovascular medicine*. McGraw-Hill Education, 2017.

Brendan P. Lucey, Jennifer S. Mcleland, Cristina D. Toedebusch, Jill Boyd, John C. Morris, Eric C. Landsness, Kelvin Yamada, and David M. Holtzman. Comparison of a single-channel EEG sleep study to polysomnography. *Journal of Sleep Research*, 25(6):625–635, June 2016. doi: 10.1111/jsr.12417. URL https://doi.org/10.1111/jsr.12417.

Raman K. Malhotra and Alon Y. Avidan. Chapter 3 - sleep stages and scoring technique. In Sudhansu Chokroverty and Robert J. Thomas, editors, *Atlas of Sleep Medicine*, pages 77 – 99. W.B. Saunders, St. Louis, second edition edition, 2014. ISBN 978-1-4557-1267-0. doi: https://doi.org/10.1016/B978-1-4557-1267-0.00003-5. URL http://www.sciencedirect.com/science/article/pii/B9781455712670000035.

Giuseppe Mancia. Autonomic modulation of the cardiovascular system during sleep. *New England Journal of Medicine*, 328(5):347–349, February 1993. doi: 10.1056/nejm199302043280511. URL https://doi.org/10.1056/nejm199302043280511.

Md. Dilshad Manzar, Wassilatul Zannat, M. Ejaz Hussain, Seithikurippu R. Pandi-Perumal, Ahmed S. Bahammam, Doaa Barakat, Nwakile Izuchukwu Ojike, Awad Olaish, and D. Warren Spence. Dimensionality of the pittsburgh sleep quality index in the young collegiate adults. *SpringerPlus*, 5(1), September 2016. doi: 10.1186/s40064-016-3234-x. URL https://doi.org/10.1186/s40064-016-3234-x.

C. McCarthy, N. Pradhan, C. Redpath, and A. Adler. Validation of the empatica e4 wristband. pages 1–4, 2016. doi: 10.1109/EMBSISC.2016.7508621.

Mary L. McHugh. The chi-square test of independence. *Biochemia Medica*, pages 143–149, 2013. doi: 10.11613/bm.2013.018. URL https://doi.org/10.11613/bm.2013.018.

N. Milstein and I. Gordon. Validating Measures of Electrodermal Activity and Heart Rate Variability Derived From the Empatica E4 Utilized in Research Settings That Involve Interactive Dyadic States. *Front Behav Neurosci*, 14:148, 2020.

U. K. Misra. *Clinical electroencephalography*. Elsevier India, New Delhi, 2018. ISBN 9788131252956.

T. Mollayeva, P. Thurairajah, K. Burton, S. Mollayeva, C. M. Shapiro, and A. Colantonio. The Pittsburgh sleep quality index as a screening tool for sleep dysfunction in clinical and non-clinical samples: A systematic review and meta-analysis. *Sleep Med Rev*, 25:52–73, Feb 2016.

Doris Moser, Peter Anderer, Georg Gruber, Silvia Parapatics, Erna Loretz, Marion Boeck, Gerhard Kloesch, Esther Heller, Andrea Schmidt, Heidi Danker-Hopfe, Bernd Saletu, Josef Zeitlhofer, and Georg Dorffner. Sleep Classification According to AASM and Rechtschaffen
amp; Kales: Effects on Sleep Scoring Parameters. *Sleep*, 32(2):139–149, 02 2009. ISSN 0161-8105. doi: 10.1093/sleep/32.2.139. URL https://doi.org/10.1093/sleep/32.2.139.

S. Mousavi, F. Afghah, and U. R. Acharya. SleepEEGNet: Automated sleep stage scoring with sequence to sequence deep learning approach. *PLoS One*, 14(5):e0216456, 2019.

Fionn Murtagh. Multilayer perceptrons for classification and regression. *Neurocomputing*, 2(5): 183–197, 1991. ISSN 0925-2312. doi: https://doi.org/10.1016/0925-2312(91)90023-5. URL https://www.sciencedirect.com/science/article/pii/0925231291900235.

A. Muzet, J. Ehrhart, V. Candas, J. P. Libert, and J. J. Vogt. Rem sleep and ambient temperature in man. *International Journal of Neuroscience*, 18(1-2):117–125, January 1983. doi: 10.3109/00207458308985885. URL https://doi.org/10.3109/00207458308985885.

Pedro Nogueira, Joana Urbano, Luís Paulo Reis, Henrique Lopes Cardoso, Daniel Castro Silva, Ana Paula Rocha, Joaquim Gonçalves, and Brígida Mónica Faria. A review of commercial and medical-grade physiological monitoring devices for biofeedback-assisted quality of life improvement studies. *Journal of Medical Systems*, 42(6), April 2018. doi: 10.1007/s10916-018-0946-1. URL https://doi.org/10.1007/s10916-018-0946-1.

K OKAMOTOMIZUNO, K TSUZUKI, K MIZUNO, and T IWAKI. Effects of partial humid heat exposure during different segments of sleep on human sleep stages and body temperature. *Physiology & Behavior*, 83(5):759–765, January 2005. doi: 10.1016/j.physbeh.2004.09.009. URL https://doi.org/10.1016/j.physbeh.2004.09.009.

Alexandra Oliveira, Joyce Aguiar, Eliana Silva, Brígida Mónica Faria, Helena Gonçalves, Luís Teófilo, Joaquim Gonçalves, Victor Carvalho, Henrique Lopes Cardoso, and Luís Paulo Reis. Assessing daily activities using a PPG sensor embedded in a wristband-type activity tracker. In *Trends and Innovations in Information Systems and Technologies*, pages 108–119. Springer International Publishing, 2020. doi: 10.1007/978-3-030-45697-9_11. URL https://doi.org/10.1007/978-3-030-45697-9_11.

S. Ollander, C. Godin, A. Campagne, and S. Charbonnier. A comparison of wearable and stationary sensors for stress detection. pages 004362–004366, 2016. doi: 10.1109/SMC.2016.7844917.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Donald B. Percival and Debashis Mondal. A wavelet variance primer. In *Time Series Analysis: Methods and Applications*, pages 623–657. Elsevier, 2012. doi: 10.1016/b978-0-444-53858-1.00022-3. URL https://doi.org/10.1016/b978-0-444-53858-1.00022-3.

M. Raghu Ram, K. Venu Madhav, E. Hari Krishna, Nagarjuna Reddy Komalla, and K. Ashoka Reddy. A novel approach for motion artifact reduction in ppg signals based on as-lms adaptive filter. *IEEE Transactions on Instrumentation and Measurement*, 61(5):1445–1457, 2012. doi: 10.1109/TIM.2011.2175832.

Irina Rish. An empirical study of the naïve bayes classifier. *IJCAI 2001 Work Empir Methods Artif Intell*, 3, 01 2001.

Warren R. Ruehland, Fergal J. O'Donoghue, Robert J. Pierce, Andrew T. Thornton, Parmjit Singh, Janet M. Copland, Bronwyn Stevens, and Peter D. Rochford. The 2007 AASM recommendations for EEG electrode placement in polysomnography: Impact on sleep and cortical arousal scoring. *Sleep*, 34(1):73–81, January 2011. doi: 10.1093/sleep/34.1.73. URL https://doi.org/10.1093/sleep/34.1.73.

J. C. Sagot, C. Amoros, V. Candas, and J. P. Libert. Sweating responses and body temperatures during nocturnal sleep in humans. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 252(3):R462–R470, March 1987. doi: 10.1152/ajpregu.1987.252.3.r462. URL https://doi.org/10.1152/ajpregu.1987.252.3.r462.

Vincent J. Samar, Ajit Bopardikar, Raghuveer Rao, and Kenneth Swartz. Wavelet analysis of neuroelectric waveforms: A conceptual tutorial. *Brain and Language*, 66(1):7–60, January 1999. doi: 10.1006/brln.1998.2024. URL https://doi.org/10.1006/brln.1998.2024.

Susana Aguiar Santos, Boudewijn Venema, and Steffen Leonhardt. Accelerometer-assisted PPG measurement during physical exercise using the LAVIMO sensor system. *Acta Polytechnica*, 52(5), January 2012. doi: 10.14311/1646. URL https://doi.org/10.14311/1646.

Ben Sawyer, Grace Teo, and Mustapha Mouloua. Driveid: Safety innovation through individuation. *Work*, 41:4273–4278, 2012. ISSN 10519815. doi: 10.3233/WOR-2012-0722-4273.

Thomas E. Scammell, Elda Arrigoni, and Jonathan O. Lipton. Neural circuitry of wakefulness and sleep. *Neuron*, 93(4):747 – 765, 2017. ISSN 0896-6273. doi: https://doi.org/10.1016/j.neuron.2017.01.014. URL http://www.sciencedirect.com/science/article/pii/S0896627317300387.

Patrick Schober, Christa Boer, and Lothar A. Schwarte. Correlation coefficients. *Anesthesia & Analgesia*, 126(5):1763–1768, May 2018. doi: 10.1213/ane.0000000000002864. URL https://doi.org/10.1213/ane.0000000000002864.

Arthur Schuster. The periodogram and its optical analogy. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 77(515):136–140, February 1906. doi: 10.1098/rspa.1906.0011. URL https://doi.org/10.1098/rspa.1906.0011.

Gulbadan Sikander and Shahzad Anwar. Driver fatigue detection systems: A review. *IEEE Transactions on Intelligent Transportation Systems*, 20(6):2339–2352, 2019. doi: 10.1109/TITS.2018.2868499.

James E Sinex. Pulse oximetry: Principles and limitations. *The American Journal of Emergency Medicine*, 17(1):59–66, January 1999. doi: 10.1016/s0735-6757(99)90019-0. URL https://doi.org/10.1016/s0735-6757(99)90019-0.

Sónia Soares, Sara Ferreira, and António Couto. Driving simulator experiments to study drowsiness: A systematic review. *Traffic Injury Prevention*, 21(1):29–37, January 2020. doi: 10.1080/15389588.2019.1706088. URL https://doi.org/10.1080/15389588.2019.1706088.

Y. Y. Song and Y. Lu. Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry*, 27(2):130–135, Apr 2015.

Kai Spiegelhalder, Bernd Feige, Dominik Paul, Dieter Riemann, Ludger Tebartz van Elst, Erich Seifritz, Jürgen Hennig, and Magdolna Hornyak. Cerebral correlates of muscle tone fluctuations in restless legs syndrome: A pilot study with combined functional magnetic resonance imaging and anterior tibial muscle electromyography. *Sleep Medicine*, 9(2):177 – 183, 2008. ISSN 1389-9457. doi: https://doi.org/10.1016/j.sleep.2007.03.021. URL http://www.sciencedirect.com/science/article/pii/S1389945707001281.

R. Subramanian, J. Wache, M. K. Abadi, R. L. Vieriu, S. Winkler, and N. Sebe. Ascertain: Emotion and personality recognition using commercial sensors. *IEEE Transactions on Affective Computing*, 9(2):147–160, 2018. doi: 10.1109/TAFFC.2016.2625250.

Akara Supratak, Hao Dong, Chao Wu, and Yike Guo. DeepSleepNet: A model for automatic sleep stage scoring based on raw single-channel EEG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(11):1998–2008, November 2017. doi: 10.1109/tnsre.2017.2721116. URL https://doi.org/10.1109/tnsre.2017.2721116.

Kevin Sweeney. *Motion Artifact Processing Techniques for Physiological Signals*. PhD thesis, 04 2013.

Mitra Taghizadeh-Sarabi, Kavous Salehzadeh Niksirat, Sohrab Khanmohammadi, and Mohammadali Nazari. EEG-based analysis of human driving performance in turning left and right using hopfield neural network. *SpringerPlus*, 2(1), December 2013. doi: 10.1186/2193-1801-2-662. URL https://doi.org/10.1186/2193-1801-2-662.

M. G. Terzano, L. Parrino, A. Smerieri, R. Chervin, S. Chokroverty, C. Guilleminault, M. Hirshkowitz, M. Mahowald, H. Moldofsky, A. Rosa, R. Thomas, and A. Walters. Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep. *Sleep Med*, 3(2):187–199, Mar 2002.

Alaa Tharwat. Linear vs. quadratic discriminant analysis classifier: a tutorial. *International Journal of Applied Pattern Recognition*, 3(2):145, 2016. doi: 10.1504/ijapr.2016.079050. URL https://doi.org/10.1504/ijapr.2016.079050.

David J Thomson et al. Multitaper analysis of nonstationary and nonlinear time series data. *Nonlinear and nonstationary signal processing*, pages 317–394, 2000.

Kevin K. Tremper. Pulse oximetry. *Chest*, 95(4):713–715, April 1989. doi: 10.1378/chest.95.4.713. URL https://doi.org/10.1378/chest.95.4.713.

Vladas Valiulis. *The Effect of Transcranial Magnetic Stimulation on Brain Bioelectrical Activity*. PhD thesis, 09 2014.

Nelleke van Wouwe, Pierre Valk, and Bertil Veenstra. Sleep monitoring: A comparison between three wearable instruments. *Military medicine*, 176:811–6, 07 2011. doi: 10.7205/MILMED-D-10-00389.

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

Olivia Walch, Yitong Huang, Daniel Forger, and Cathy Goldstein. Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device. *Sleep*, 42(12), 08 2019. ISSN 0161-8105. doi: 10.1093/sleep/zsz180. URL https://doi.org/10.1093/sleep/zsz180. zsz180.

P. Welch. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2):70–73, 1967. doi: 10.1109/TAU.1967.1161901.

Wes McKinney. Data Structures for Statistical Computing in Python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 56 – 61, 2010. doi: 10.25080/Majora-92bf1922-00a.

Zhongheng Zhang. Introduction to machine learning: k-nearest neighbors. *Annals of Translational Medicine*, 4(11):218–218, June 2016. doi: 10.21037/atm.2016.03.37. URL https://doi.org/10.21037/atm.2016.03.37.

Appendix A

Tables

Classifier	ROC AUC		precision	recall	f1-score
Nearest Neighbors	72	Exam	67	57	61
		Practice	64	73	68
		accuracy	65	65	65
		macro avg	65	65	65
		weighted avg	65	65	65
RBF SVM		Exam	67	57	61
		Practice	64	73	68
		accuracy	65	65	65
		macro avg	65	65	65
		weighted avg	65	65	65
Gaussian Process	76	Exam	68	59	63
		Practice	65	73	69
		accuracy	66	66	66
		macro avg	66	66	66
		weighted avg	66	66	66
Decision Tree	59	Exam	71	26	38
		Practice	56	90	69
		accuracy	59	59	59
		macro avg	63	58	53
		weighted avg	63	59	54
Random Forest	76	Exam	67	63	65
		Practice	67	71	69
		accuracy	67	67	67
		macro avg	67	67	67
		weighted avg	67	67	67
Multi-layer Perceptron	76	Exam	74	61	67
		Practice	68	79	73
		accuracy	70	70	70
		macro avg	71	70	70
		weighted avg	71	70	70
AdaBoost	76	Exam	77	59	67
		Practice	68	83	75
		accuracy	71	71	71
		macro avg	72	71	71
		weighted avg	72	71	71
Naive Bayes	71	Exam	56	96	71
		Practice	88	29	44
		accuracy	62	62	62
		macro avg	72	62	57
		weighted avg	72	62	57
QDA	65	Exam	52	98	68
		Practice	88	15	25
		accuracy	55	55	55
		macro avg	70	56	47
		weighted avg	70	55	46
Logistic Regression	70	Exam	64	61	62
		Practice	64	67	65
		accuracy	64	64	64
		macro avg	64	64	64
		weighted avg	64	64	64

Table A.1: Practice vs Exam run classification results (in percent format).

Classifier	ROC AUC		precision	recall	f1-score
Nearest Neighbors	79	Other	59	81	68
		Individual	84	64	73
		accuracy	71	71	71
		macro avg	72	73	71
		weighted avg	74	71	71
RBF SVM		Other	63	75	69
		Individual	82	72	77
		accuracy	73	73	73
		macro avg	72	74	73
		weighted avg	75	73	73
Gaussian Process	81	Other	64	88	74
		Individual	89	68	77
		accuracy	76	76	76
		macro avg	77	78	75
		weighted avg	79	76	76
Decision Tree	74	Other	57	75	65
		Individual	80	64	71
		accuracy	68	68	68
		macro avg	69	70	68
		weighted avg	71	68	69
Random Forest	84	Other	71	75	73
		Individual	83	80	82
		accuracy	78	78	78
		macro avg	77	78	77
		weighted avg	78	78	78
Multi-layer Perceptron	80	Other	58	88	70
		Individual	88	60	71
		accuracy	71	71	71
		macro avg	73	74	71
		weighted avg	77	71	71
AdaBoost	75	Other	62	81	70
		Individual	85	68	76
		accuracy	73	73	73
		macro avg	73	75	73
		weighted avg	76	73	73
Naive Bayes	79	Other	54	94	68
		Individual	92	48	63
		accuracy	66	66	66
		macro avg	73	71	66
		weighted avg	77	66	65
QDA	68	Other	62	50	55
		Individual	71	80	75
		accuracy	68	68	68
		macro avg	66	65	65
		weighted avg	68	68	68
Logistic Regression	81	Other	64	88	74
		Individual	89	68	77
		accuracy	76	76	76
		macro avg	77	78	75
		weighted avg	79	76	76

Table A.2: Single Individual's vs Other's run classification results (in percent format).

Classifier	ROC AUC		precision	recall	f1-score
Nearest Neighbors	72	Exam	89	87	88
		Exam SD	76	79	78
		accuracy	84	84	84
		macro avg	82	83	83
		weighted avg	84	84	84
RBF SVM		Exam	90	82	86
		Exam SD	71	83	77
		accuracy	83	83	83
		macro avg	81	83	81
		weighted avg	84	83	83
Gaussian Process	90	Exam	88	82	85
		Exam SD	70	79	75
		accuracy	81	81	81
		macro avg	79	81	80
		weighted avg	82	81	81
Decision Tree	78	Exam	86	80	83
		Exam SD	67	75	71
		accuracy	78	78	78
		macro avg	76	78	77
		weighted avg	79	78	79
Random Forest	95	Exam	95	87	91
		Exam SD	79	92	85
		accuracy	88	88	88
		macro avg	87	89	88
		weighted avg	89	88	89
Multi-layer Perceptron	92	Exam	88	80	84
		Exam SD	68	79	73
		accuracy	80	80	80
		macro avg	78	80	78
		weighted avg	81	80	80
AdaBoost	93	Exam	89	87	88
		Exam SD	76	79	78
		accuracy	84	84	84
		macro avg	82	83	83
		weighted avg	84	84	84
Naive Bayes	94	Exam	97	82	89
		Exam SD	74	96	84
		accuracy	87	87	87
		macro avg	86	89	86
		weighted avg	89	87	87
QDA	92	Exam	79	98	87
		Exam SD	92	50	65
		accuracy	81	81	81
		macro avg	85	74	76
		weighted avg	83	81	79
Logistic Regression	90	Exam	86	82	84
		Exam SD	69	75	72
		accuracy	80	80	80
		macro avg	78	79	78
		weighted avg	80	80	80

Table A.3: Last Night Sleep Deprivation vs Non-Last Night Sleep Deprivation run classification results (in percent format).

Appendix B

Figures

INSTRUCTIONS:

The following questions relate to your usual sleep habits during the past month only. Your answers should indicate the most accurate reply for the majority of days and nights in the past month. Please answer all questions.

During the past month,

1. When have you usually gone to bed? _____
2. How long (in minutes) has it taken you to fall asleep each night? _____
3. What time have you usually gotten up in the morning? _____
4. A. How many hours of actual sleep did you get at night? _____
   B. How many hours were you in bed? _____

5. During the past month, how often have you had trouble sleeping because you	Not during the past month (0)	Less than once a week (1)	Once or twice a week (2)	Three or more times a week (3)
A. Cannot get to sleep within 30 minutes				
B. Wake up in the middle of the night or early morning				
C. Have to get up to use the bathroom				
D. Cannot breathe comfortably				
E. Cough or snore loudly				
F. Feel too cold				
G. Feel too hot				
H. Have bad dreams				
I. Have pain				
J. Other reason (s), please describe, including how often you have had trouble sleeping because of this reason (s):				
6. During the past month, how often have you taken medicine (prescribed or "over the counter") to help you sleep?				
7. During the past month, how often have you had trouble staying awake while driving, eating meals, or engaging in social activity?				
8. During the past month, how much of a problem has it been for you to keep up enthusiasm to get things done?				
9. During the past month, how would you rate your sleep quality overall?	Very good (0)	Fairly good (1)	Fairly bad (2)	Very bad (3)

Scoring

Component 1	#9 Score	C1 _____
Component 2	#2 Score (<15min (0), 16-30min (1), 31-60 min (2), >60min (3)) + #5a Score (if sum is equal 0=0; 1-2=1; 3-4=2; 5-6=3)	C2 _____
Component 3	#4 Score (>7(0), 6-7 (1), 5-6 (2), <5 (3)	C3 _____
Component 4	(total # of hours asleep) / (total # of hours in bed) x 100 >85%=0, 75%-84%=!, 65%-74%=2, <65%=3	C4 _____
Component 5	# sum of scores 5b to 5j (0=0; 1-9=1; 10-18=2; 19-27=3)	C5 _____
Component 6	#6 Score	C6 _____
Component 7	#7 Score + #8 score (0=0; 1-2=1; 3-4=2; 5-6=3)	C7 _____

Add the seven component scores together _____ ***Global PSQI*** _____

Figure B.1: PSQI questionnaire instructions, questions, components, and scoring (modified from Aurora University document).

CONSENTIMENTO INFORMADO, LIVRE E ESCLARECIDO PARA PARTICIPAÇÃO EM INVESTIGAÇÃO

Conforme a lei 67/98 de 26 de Outubro e a "Declaração de Helsínquia" da Associação Médica Mundial (Helsínquia 1964; Tóquio 1975; Veneza 1983; Hong Kong 1989; Somerset West 1996, Edimburgo 2000; Washington 2002, Tóquio 2004, Seul 2008)

Designação do Estudo: Sono ao Volante 2.0 - Sistema de Informação para a previsão do sono ao volante e a deteção de distúrbio ou privação crónica do sono.

Consórcio: Laboratório de Inteligência Artificial e Ciência de Computadores (LIACC) da Faculdade de Engenharia da Universidade do Porto, Optimizer - Serviços e Consultoria Informática Lda, Instituto do Sono (ISCCI) - Centro Clínico e Investigação, e Instituto Politécnico do Cávado e Ave (IPCA).

Fui informado(a) de que o Estudo de Investigação acima mencionado se destina a obter dados para previsão e deteção de distúrbios ou privação de sono, com o objetivo de avaliar a capacidade e agilidade na tomada de decisão e medir o impacto nas atividades diárias, pessoais e profissionais.

Para o efeito irá conduzir um veículo em simulação de computador, recorrendo a uma interface de volante, caixa de velocidades e joystick, no decorrer de 2 seções: uma primeira em que dormiu o tempo de sono normal e uma segunda em que terá entre 18h a 24h horas de privação de sono. Enquanto realiza esta ação, serão recolhidos dados biométricos da atividade cerebral e cardíaca com os sensores *Brainlink Pro* e *E4 Wristband*, respetivamente. Responderá ainda a dois questionários que têm como objetivo avaliar a qualidade de sono e identificar o ciclo circadiano.

Foi-me garantido que a captura de vídeo efetuada durante a experiência não será divulgada a terceiros.

Fui ainda informado(a) que neste estudo está prevista a realização de testes num simulador e preenchimento de questionários tendo-me sido explicado em que consistem e quais os seus possíveis efeitos.

Foi-me garantido que todos os dados relativos à identificação dos Participantes neste estudo são confidenciais e que será mantido o anonimato.

Sei que posso recusar participar ou interromper a qualquer momento a participação no estudo, sem nenhum tipo de penalização.

Fui informado(a) de que não está contemplado qualquer ressarcimento ou remuneração para participação no estudo assim como não apresenta qualquer custo para os participantes.

Fui informado que não está contemplada qualquer compensação na eventualidade de acidentes provocados pela privação de sono requerida na 2ª sessão da experiência.

Compreendi a informação que me foi dada, tive oportunidade de fazer perguntas e as minhas

Figure B.2: Agreement of consent, in Portuguese (page 1 of 2).

dúvidas foram esclarecidas.

Em nome da equipa agradecemos a sua participação.

Investigador responsável: Professor Doutor Luís Paulo Reis

Assinatura: … … … … … … … … … … … … … … ….. … … … … … … … … … … … … …

-o-o-o-o-o-o-o-o-o-o-o-o-o-o-o-o-o-o-o-o-

Eu, abaixo-assinado aceito participar de livre vontade no estudo acima mencionado. Também autorizo a divulgação dos resultados obtidos no meio científico, com as garantias de anonimato e confidencialidade dos dados fornecidos.

Data Assinatura

… … /… … /… … …………………………………………………………………………………………………

Se não for o participante do estudo a assinar:

Eu, abaixo-assinado ____(nome completo do representante legal do participante do estudo), na qualidade de representante legal de ____ (nome completo do participante do estudo) autorizo de livre vontade a participação daquele que legalmente represento no estudo acima mencionado.

Data Assinatura

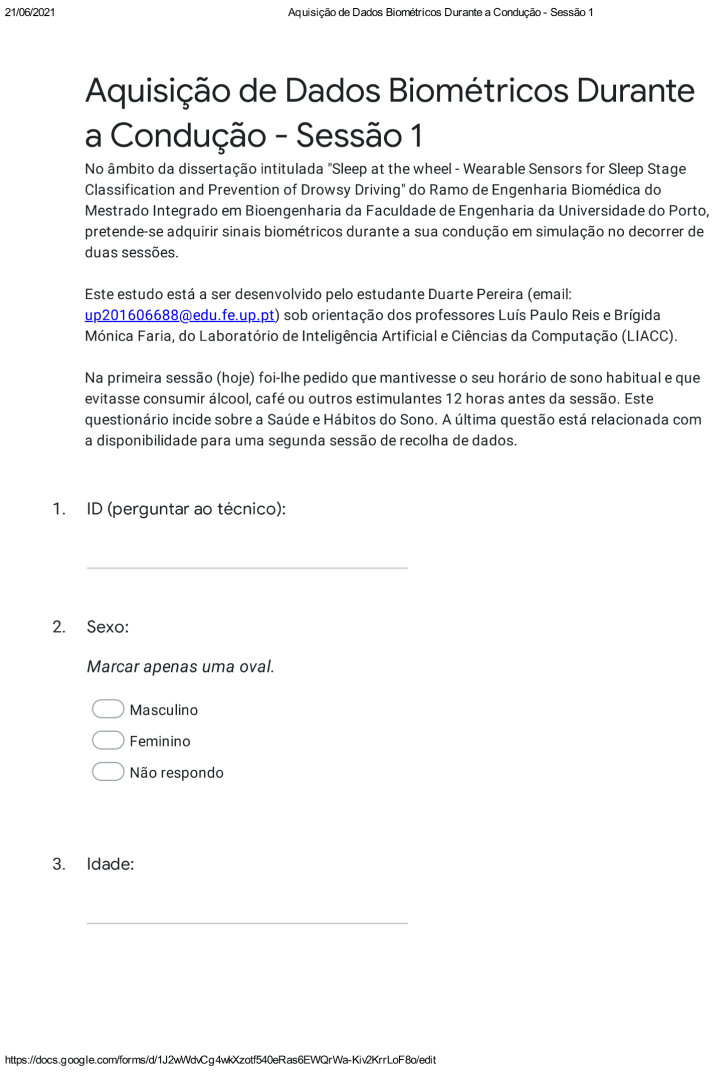
… … /… … /… … …………………………………………………………………………………………………

ESTE DOCUMENTO É COMPOSTO POR 2 PÁGINAS E FEITO EM DUPLICADO: UMA VIA VAI PARA O/A INVESTIGADOR/A, OUTRA PARA A PESSOA QUE CONSENTE.

Figure B.3: Agreement of consent, in Portuguese (page 2 of 2).

Appendix C

Questionnaire



21/06/2021 Aquisição de Dados Biométricos Durante a Condução - Sessão 1

Aquisição de Dados Biométricos Durante a Condução - Sessão 1

No âmbito da dissertação intitulada "Sleep at the wheel - Wearable Sensors for Sleep Stage Classification and Prevention of Drowsy Driving" do Ramo de Engenharia Biomédica do Mestrado Integrado em Bioengenharia da Faculdade de Engenharia da Universidade do Porto, pretende-se adquirir sinais biométricos durante a sua condução em simulação no decorrer de duas sessões.

Este estudo está a ser desenvolvido pelo estudante Duarte Pereira (email: up201606688@edu.fe.up.pt) sob orientação dos professores Luís Paulo Reis e Brígida Mónica Faria, do Laboratório de Inteligência Artificial e Ciências da Computação (LIACC).

Na primeira sessão (hoje) foi-lhe pedido que mantivesse o seu horário de sono habitual e que evitasse consumir álcool, café ou outros estimulantes 12 horas antes da sessão. Este questionário incide sobre a Saúde e Hábitos do Sono. A última questão está relacionada com a disponibilidade para uma segunda sessão de recolha de dados.

1. ID (perguntar ao técnico):

2. Sexo:

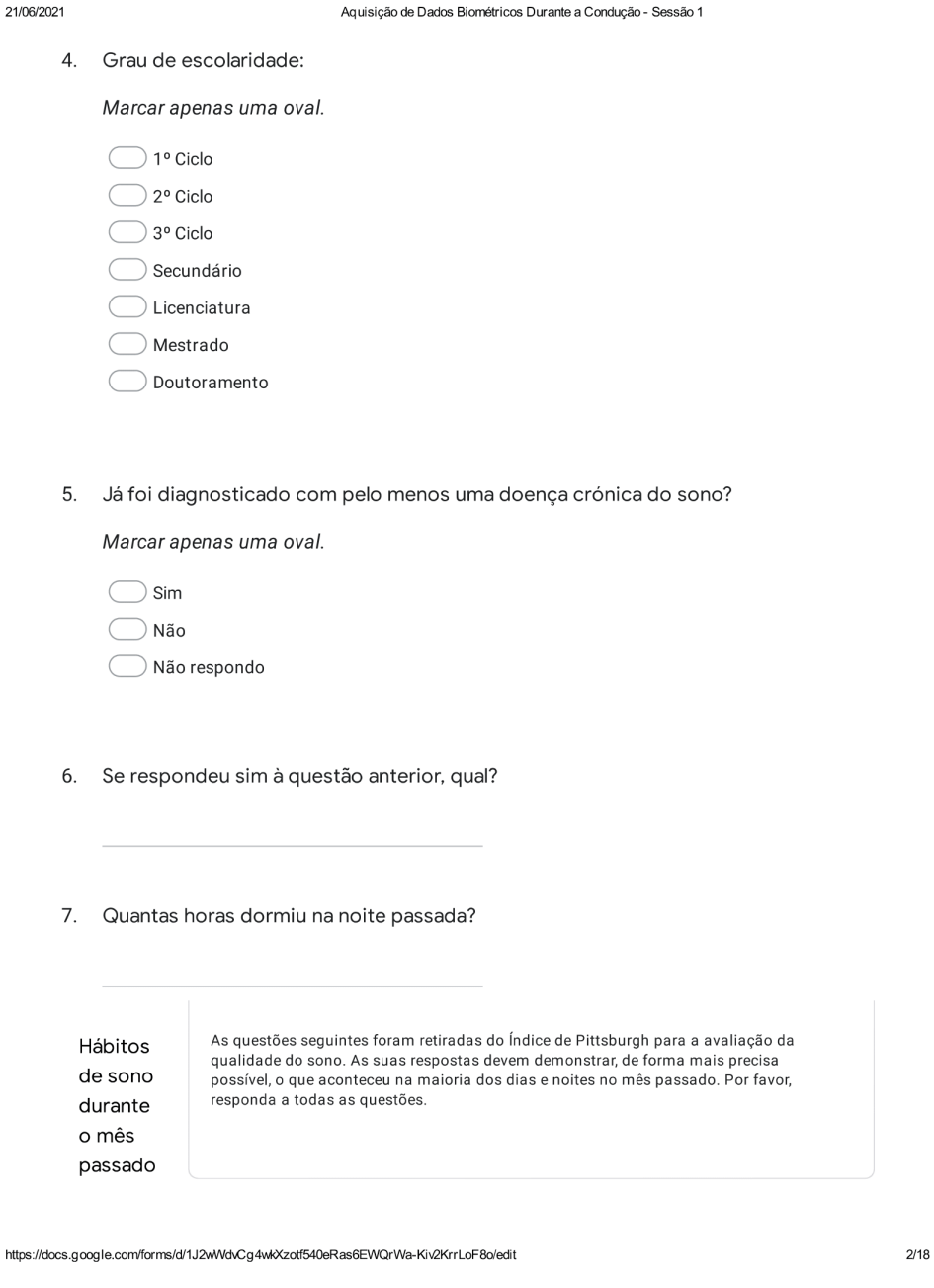
Marcar apenas uma oval.

- Masculino
- Feminino
- Não respondo

3. Idade:

https://docs.google.com/forms/d/1J2wWdvCg4wkXzotf540eRas6EWQrWa-Kiv2KrrLoF8o/edit 1/18

Figure C.1: Page 1 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)



21/06/2021 Aquisição de Dados Biométricos Durante a Condução - Sessão 1

4. Grau de escolaridade:

Marcar apenas uma oval.

- 1º Ciclo
- 2º Ciclo
- 3º Ciclo
- Secundário
- Licenciatura
- Mestrado
- Doutoramento

5. Já foi diagnosticado com pelo menos uma doença crónica do sono?

Marcar apenas uma oval.

- Sim
- Não
- Não respondo

6. Se respondeu sim à questão anterior, qual?

7. Quantas horas dormiu na noite passada?

Hábitos de sono durante o mês passado

As questões seguintes foram retiradas do Índice de Pittsburgh para a avaliação da qualidade do sono. As suas respostas devem demonstrar, de forma mais precisa possível, o que aconteceu na maioria dos dias e noites no mês passado. Por favor, responda a todas as questões.

https://docs.google.com/forms/d/1J2wWdvCg4wkXzotf540eRas6EWQrWa-Kiv2KrrLoF8o/edit 2/18

Figure C.2: Page 2 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

8. 1 - Durante o mês passado, a que horas foi habitualmente dormir?

Exemplo: 08:30

9. 2 - Durante o mês passado, quanto tempo (em minutos) habitualmente levou para adormecer à cada noite?

10. 3 - Durante o mês passado, a que horas habitualmente acordou?

Exemplo: 08:30

11. 4 - Durante o mês passado, quantas horas de sono realmente teve à noite? (isto pode ser diferente do número de horas que permaneceu na cama)

5 - Durante o mês passado, com que frequência teve problemas de sono porque:

12. 5.1 - não conseguia dormir em 30 minutos.

Marcar apenas uma oval.

- ◯ nunca no mês passado
- ◯ menos de uma vez por semana
- ◯ uma ou duas vezes por semana
- ◯ três ou mais vezes por semana

Figure C.3: Page 3 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

13. 5.2 - despertou a meio da noite ou de madrugada.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

14. 5.3 - teve que se levantar à noite para ir à casa de banho.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

15. 5.4 - não conseguia respirar de forma satisfatória.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

Figure C.4: Page 4 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

16. 5.5 - tossia ou ressonava alto.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

17. 5.6 - sentia muito frio.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

18. 5.7 - sentia muito calor.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

Figure C.5: Page 5 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

19. 5.8 - tinha pesadelos.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

20. 5.9 - tinha dor.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

21. 5.10 - outra razão (por favor descreva sucintamente)

22. 5.11 - Caso tenha respondido à questão a cima, com que frequência teve problemas de sono por essa causa?

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

Figure C.6: Page 6 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

23. 6 - Durante o mês passado, como avaliaria a qualidade geral do seu sono?

Marcar apenas uma oval.

	1	2	3	4	5	
Muito Mau	◯	◯	◯	◯	◯	Muito Bom

24. 7 - Durante o mês passado, com que frequência tomou medicação (por conta própria) para ajudar a adormecer?

Marcar apenas uma oval.

- ◯ nunca no mês passado
- ◯ menos de uma vez por semana
- ◯ uma ou duas vezes por semana
- ◯ três ou mais vezes por semana

25. 8 - Durante o mês passado, com que frequência teve dificuldades em permanecer acordado enquanto estava a conduzir?

Marcar apenas uma oval.

- ◯ nunca no mês passado
- ◯ menos de uma vez por semana
- ◯ uma ou duas vezes por semana
- ◯ três vezes por semana

Figure C.7: Page 7 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

26. 9 - Durante o mês passado, com que frequência teve problemas em sentir-se suficientemente ativo ao realizar suas atividades?

Marcar apenas uma oval.

- ◯ nunca no mês passado
- ◯ menos de uma vez por semana
- ◯ uma ou duas vezes por semana
- ◯ três vezes por semana

Hábitos de sono durante o mês passado

As questões seguintes foram retiradas do Índice de Pittsburgh para a avaliação da qualidade do sono. As suas respostas devem demonstrar, de forma mais precisa possível, o que aconteceu na maioria dos dias e noites no mês passado. Por favor, responda a todas as questões.

27. 1 - Durante o mês passado, a que horas foi habitualmente dormir?

Exemplo: 08:30

28. 2 - Durante o mês passado, quanto tempo (em minutos) habitualmente levou para adormecer à cada noite?

29. 3 - Durante o mês passado, a que horas habitualmente acordou?

Exemplo: 08:30

30. 4 - Durante o mês passado, quantas horas de sono realmente teve à noite? (isto pode ser diferente do número de horas que permaneceu na cama)

Figure C.8: Page 8 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

5 - Durante o mês passado, com que frequência teve problemas de sono porque:

31. 5.1 - não conseguia dormir em 30 minutos.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

32. 5.2 - despertou a meio da noite ou de madrugada.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

33. 5.3 - teve que se levantar à noite para ir à casa de banho.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

Figure C.9: Page 9 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

34. 5.4 - não conseguia respirar de forma satisfatória.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

35. 5.5 - tossia ou ressonava alto.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

36. 5.6 - sentia muito frio.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

Figure C.10: Page 10 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

37. 5.7 - sentia muito calor.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

38. 5.8 - tinha pesadelos.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

39. 5.9 - tinha dor.

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três ou mais vezes por semana

40. 5.10 - outra razão (por favor descreva sucintamente)

Figure C.11: Page 11 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

41. 5.11 - Caso tenha respondido à questão a cima, com que frequência teve problemas de sono por essa causa?

Marcar apenas uma oval.

- ◯ nunca no mês passado
- ◯ menos de uma vez por semana
- ◯ uma ou duas vezes por semana
- ◯ três ou mais vezes por semana

42. 6 - Durante o mês passado, como avaliaria a qualidade geral do seu sono?

Marcar apenas uma oval.

	1	2	3	4	5	
Muito Mau	◯	◯	◯	◯	◯	Muito Bom

43. 7 - Durante o mês passado, com que frequência tomou medicação (por conta própria) para ajudar a adormecer?

Marcar apenas uma oval.

- ◯ nunca no mês passado
- ◯ menos de uma vez por semana
- ◯ uma ou duas vezes por semana
- ◯ três ou mais vezes por semana

Figure C.12: Page 12 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

44. 8 - Durante o mês passado, com que frequência teve dificuldades em permanecer acordado enquanto estava a conduzir?

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três vezes por semana

45. 9 - Durante o mês passado, com que frequência teve problemas em sentir-se suficientemente ativo ao realizar suas atividades?

Marcar apenas uma oval.

- nunca no mês passado
- menos de uma vez por semana
- uma ou duas vezes por semana
- três vezes por semana

Escala de Saúde do Sono

As seguintes questões foram retiradas do questionário "Satisfaction Alertness Timing Efficiency and Duration" (SATED) e referem-se a vários aspetos do seu sono. Para cada uma delas assinale a opção que melhor se adequa a si, numa escala de 0 (nunca) a 5 (sempre).

46. 10 - Deita-se e levanta-se mais ou menos à mesma hora todos os dias?

Marcar apenas uma oval.

	1	2	3	4	5	
Nunca						Sempre

Figure C.13: Page 13 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

47. 11 - Está satisfeito(a) com o seu sono?

Marcar apenas uma oval.

	1	2	3	4	5	
Nunca	◯	◯	◯	◯	◯	Sempre

48. 12 - Encontra-se já a dormir (ou na cama) entre as 2 e 4 da manhã?

Marcar apenas uma oval.

	1	2	3	4	5	
Nunca	◯	◯	◯	◯	◯	Sempre

49. 13 - Passa menos de 30 minutos acordado à noite? (inclui o tempo que leva a adormecer e os despertares durante o sono)

Marcar apenas uma oval.

	1	2	3	4	5	
Nunca	◯	◯	◯	◯	◯	Sempre

50. 14 - Dorme entre 7 e 9 horas por dia?

Marcar apenas uma oval.

	1	2	3	4	5	
Nunca	◯	◯	◯	◯	◯	Sempre

Obrigado! Por favor, espere nesta página.

Por favor, aguarde pelo inicio do teste antes de seguir para a próxima secção.

Figure C.14: Page 14 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

Avançar para a pergunta 51

Dificuldades sentidas durante o teste de condução.

Por favor, indique as dificuldades que sentiu no decorrer do teste.

51. 1 - Sentiu dificuldades no controlo do veículo?

Marcar apenas uma oval.

	1	2	3	4	5	
Pouco	◯	◯	◯	◯	◯	Muito

52. 1.1 - Indique quais:

53. 2 - Sentiu dificuldades na perceção em regime de simulação (por exemplo, enjoo ou olhar em redor em cruzamentos)?

Marcar apenas uma oval.

	1	2	3	4	5	
Pouco	◯	◯	◯	◯	◯	Muito

Figure C.15: Page 15 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

54. 2.1 - Indique quais:

55. 3 - Sentiu desconforto ao usar algum do equipamento (incluindo os sensores)?

Marcar apenas uma oval.

	1	2	3	4	5	
Pouco	◯	◯	◯	◯	◯	Muito

56. 3.1 - Indique as causas de desconforto:

57. 4 - Sentiu dificuldades em termos de manter a distância de segurança com outros veículos?

Marcar apenas uma oval.

	1	2	3	4	5	
Pouco	◯	◯	◯	◯	◯	Muito

Figure C.16: Page 16 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

58. 5 - Sentiu dificuldades em reconhecer potenciais obstáculos e a responder atempadamente aos mesmos?

Marcar apenas uma oval.

	1	2	3	4	5	
Pouco	◯	◯	◯	◯	◯	Muito

59. 6 - Sentiu dificuldades em ter em conta outros veículos e em reagir aos erros dos mesmos de forma responsável?

Marcar apenas uma oval.

	1	2	3	4	5	
Pouco	◯	◯	◯	◯	◯	Muito

60. 7 - Sentiu dificuldades em conduzir de forma sustentável (sem forçar o veículo desnecessariamente)?

Marcar apenas uma oval.

	1	2	3	4	5	
Pouco	◯	◯	◯	◯	◯	Muito

Informação útil à 2ª sessão

Na 2ª sessão, será requerido que repita o exame de hoje com entre 16 a 24h de privação de sono. Como tal, por favor preencha esta informação de modo a combinarmos uma segunda sessão. Irá ser contactado(a) após a sessão para marcação definitiva.

Figure C.17: Page 17 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)

61. Datas e horas de disponibilidade (sugerem-se manhãs de dias em que não tenha atividades que sofram devido à privação de sono):

Figure C.18: Page 18 of 18 of the PSQI, SATED and experimental feedback questionnaire, in Portuguese. (exported from Google Forms)