



Framework para la detección de somnolencia en escenarios de conducción usando Interfaces Cerebro-Máquina

Grado en Ingeniería Informática

Trabajo Fin de Grado

Autor:

José Manuel Hidalgo Rogel

Tutor/es:

Gregorio Martínez Pérez

Sergio López Bernal



**Facultad
Informática
Universidad
Murcia**

28 de Junio de 2021

Framework para la detección de somnolencia en escenarios de conducción usando Interfaces Cerebro-Máquina

Trabajo Fin de Grado

Autor

José Manuel Hidalgo Rogel

Tutor/es

Gregorio Martínez Pérez

Departamento de Ingeniería de la Información y las Comunicaciones

Sergio López Bernal

Departamento de Ingeniería de la Información y las Comunicaciones



Grado en Ingeniería Informática



Murcia, 28 de Junio de 2021

A mi familia por todas las facilidades y el apoyo que me han brindado para que pueda estudiar lo que realmente me apasiona. También Sergio, Enrique, Alberto y Gregorio por los incontables consejos y la supervisión que me han dedicado.

*El genio se hace con
un 1% de talento y
un 99% de trabajo*

Albert Einstein.

Declaración firmada sobre originalidad del trabajo

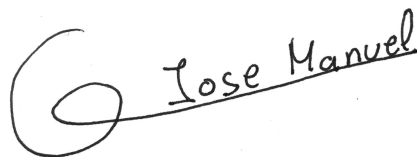
D./Dña. **José Manuel Hidalgo Rogel**, con DNI **48848317C**, estudiante de la titulación de **Grado en Ingeniería Informática** de la Universidad de Murcia y autor del TF titulado **“Framework para la detección de somnolencia en escenarios de conducción usando Interfaces Cerebro-Máquina”**.

De acuerdo con el Reglamento por el que se regulan los Trabajos Fin de Grado y de Fin de Máster en la Universidad de Murcia (aprobado C. de Gob. 30-04-2015, modificado 22-04-2016 y 28-09-2018), así como la normativa interna para la oferta, asignación, elaboración y defensa de los Trabajos Fin de Grado y Fin de Máster de las titulaciones impartidas en la Facultad de Informática de la Universidad de Murcia (aprobada en Junta de Facultad 27-11-2015)

DECLARO:

Que el Trabajo Fin de Grado presentado para su evaluación es original y de elaboración personal. Todas las fuentes utilizadas han sido debidamente citadas. Así mismo, declara que no incumple ningún contrato de confidencialidad, ni viola ningún derecho de propiedad intelectual e industrial

Murcia, a 28 de Junio de 2021

A handwritten signature in black ink, consisting of a large, stylized 'G' followed by the words 'Jose Manuel' written in a cursive script.

Fdo.: José Manuel Hidalgo Rogel
Autor del TFG

Resumen

La somnolencia al volante supone un gran peligro para los conductores, siendo uno de los principales factores presentes en los accidentes de tráfico. Además, incrementa los daños sufridos respecto a otros accidentes debido a la incapacidad de reacción por parte del conductor. La criticidad de esta problemática requiere el uso de sistemas capaces de medir en tiempo real el estado cognitivo del individuo para así tomar decisiones que eviten el accidente.

Debido a esto, la neurociencia cognitiva, disciplina encargada de estudiar las funciones mentales que se desarrollan en el sistema nervioso, ha permitido establecer una relación entre las señales cerebrales (Electroencefalografía (EEG)) y la presencia o ausencia de somnolencia. Estas señales pueden ser recopiladas y posteriormente analizadas mediante el uso de Interfaces Cerebro-Máquina (BCIs). Por lo tanto, estos sistemas abren una puerta de cara a la identificación de la somnolencia.

Las motivaciones que contribuyen a la realización de este Trabajo Final de Grado (TFG) comprenden, desde una labor social gracias a la prevención de accidentes de tráfico, junto con el acercamiento al sector de la BCI, en crecimiento y con grandes perspectivas de futuro. Además de la posibilidad de explorar algoritmos basados en Machine Learning y características que no han sido ampliamente analizadas en la literatura, junto con la escasez de propuestas de un modelo general capaz de detectar la somnolencia en múltiples usuarios.

Con el objetivo de abordar los retos anteriores, este trabajo presenta el diseño de un framework inteligente capaz de detectar la somnolencia del conductor en base a la señal EEG, las cuales se obtienen con una BCI. Tras la adquisición de las ondas cerebrales, estas señales deben ser procesadas para eliminar ruidos, así como ser clasificadas para identificar eventos relevantes asociados a la somnolencia. Entre las capacidades del framework destacan el empleo de diferentes conjuntos de características de interés de la señal, además de la variedad de técnicas de aprendizaje, tanto supervisado como no supervisado y, de algoritmos específicos empleados en cada caso.

Dentro de la fase de clasificación, el framework cuenta tanto con modelos de Machine Learning individuales, capaces de realizar una predicción para un individuo en particular, como con varios modelos generales, capaces de detectar la somnolencia de todos los sujetos. Cada modelo general tiene la particularidad de ser entrenado con un porcentaje del total de sujetos, de esta forma, se han definido tres modelos.

Para la implementación y posterior evaluación del framework, se hace uso de un *dataset* público, debido a la imposibilidad durante el trabajo de usar una BCI para adquirir datos de sujetos. Como conjuntos de características de interés, se han explorado

tanto la dimensión temporal de la señal EEG, como de la frecuencial (Densidad Espectral de Potencia (PSD)). De forma adicional, se estudia la Electrooculografía (EOG), correspondiente a los movimientos oculares del sujeto.

Se ha definido el uso de técnicas de regresión, clasificación y *clustering* como tipos de modelos de Machine Learning (ML). Dentro de regresión y clasificación, se hace uso de Support Vector Machine (SVM), k-Nearest Neighbors (kNN), Decision Tree (DT), Gaussian Process (GP) y Random Forest (RF), mientras que, para *clustering* se opta por *k-Means*, Gaussian Mixture (GM), Bayesian Gaussian Mixture (BGM) y Affinity Propagation (AP). Para la posterior evaluación de estos algoritmos se hace uso de unas métricas particulares específicas para cada familia de algoritmos, siendo Error Cuadrático Medio (RMSE) y Coeficiente de Determinación (R^2) las dedicadas a regresión; *accuracy*, *precision*, *recall* y *f1-score* correspondientes a clasificación y, para finalizar, Coeficiente Silhouette (SC) y Índice Davies-Bouldin (DBI) para *clustering*.

Pese a que SVM es el algoritmo de ML más usado en la literatura, en este framework se han evaluado otros algoritmos, tales como kNN y RF, capaces de ofrecer un rendimiento superior haciendo uso de los mismos datos. Esto ha sucedido en todos los modelos implementados, tanto los individuales como los generales. Sumado a esto, los modelos individuales han tenido el mejor rendimiento en cuanto a predicciones, seguidos de los modelos generales. Por su parte, en *clustering*, pese a los resultados aceptables de las métricas, no se discernían los niveles de somnolencia en la agrupación que se realizaba. Adicionalmente, se comprobó que todos los conjuntos de características empleados para analizar la señal ofrecen buenos resultados, incluyendo, aunque en menor medida que las anteriores, los datos del análisis temporal de la señal.

Extended Abstract

Drowsiness is defined as a person's tendency to fall asleep. However, in certain situations, it is a great danger. For instance the driving scenario, which will be the study area of this work. Because of this, several systems monitoring the subject's behaviour have been proposed aiming to detect drowsiness in drivers. These techniques represent an advance in the safety of drivers, but they have a significant drawback since they produce false positives.

In this sense, Cognitive Neuroscience, the area of knowledge responsible for studying the nervous system that supports mental functions, including drowsiness, has proposed different techniques for its quantification. The first ones were based on analysing subjects' behaviour used by the "traditional" systems mentioned above. After that, solutions based on self-assessment with scales emerged. This technique is based on asking the subject how drowsy he/she has felt in the last minutes. However, the negative part of these methods is the inclusion of subjectivity when the subject comes to his/her self-evaluation. Hence, the need to objectively quantify the sleepiness of an individual arises. For this reason, neurophysiological tests have been developed, based on the monitoring of the patient's brain signals (Electroencephalography), to quantify drowsiness.

Electroencephalography measures the electrical signal produced in the brain through sensors called electrodes. This signal has the peculiarity of its non-stationary behaviour since its amplitude and frequencies are not fixed in time. In addition, it is non-linear, referring to the fact that two EEG waves that overlap in time do not produce a product wave of the previous two multiplied by certain factors. Finally, it is non-Gaussian since the aggregation of these waves do not produce a Gaussian distribution of the data. The literature has shown that the different levels of brain activity are related to the different cognitive states of the subject. Due to this, the need arises to study the EEG signal in different frequency bands, being the lower frequency rhythms (delta, theta and alpha) directly related to the states of relaxation and drowsiness.

The proposed technological solution to study the EEG is the Brain-Computer Interface. BCIs allow the monitoring of signals passively to determine patterns of cognitive activity (drowsiness). Furthermore, they are also used to carry out medical evaluations and have the ability to interact with devices based on brain activity and in real-time.

Related to the invasiveness with which the electrodes are placed on the subject, two types of BCIs are distinguished. On the one hand, the invasive ones. In this type of BCIs, the electrodes are located in the different layers of the skull. Therefore surgical intervention is necessary. On the other hand, there are non-invasive BCIs. The way they

work is by capturing brain signals through electrodes placed directly on the subject's scalp. However, non-invasive BCIs lead to other difficulties since the quality of the signals collected is affected by the movements of the subject or the hair itself, which adds noise to the signal read. Furthermore, the specificity of the studied area is also reduced since an outer electrode collects the activity of larger groups of neurons. The most widespread systems available in order to place the electrodes is the 10-20 system, in which the arrangement of the electrodes is at intervals of 10 or 20% of separation from the measurements of the subject's skull.

A BCI system consists of the following steps. First of all, the EEG signal is acquired from the subject's electrodes, then sent to the processing device. Once they are in the device, it is necessary to process them in order to eliminate the noise from the signals, which was previously mentioned as being added in non-invasive BCIs. For this procedure, different techniques are applied, such as the Notch filter and the band-pass, resampling and Independent Component Analysis (ICA). After performing this filtering phase, it is common to reduce the number of samples of the signal, thus reducing the size of the data and speeding up its subsequent classification. An additional step that is usually used in most studies that are not in real-time is the Independent Component Analysis (ICA). After eliminating the artefacts noise from the signal, the features are extracted from it, with which the signal will be classified later. Specifically, this study focuses on classification techniques using Machine Learning. Finally, based on the result of the classification, the system produces a response, for example, notifying the user or sending an order to a device.

Regarding the state of the art, given that sleepiness studies use BCIs, their experiments consist of the same steps of a BCI. First, the brain signals must be acquired. For this, in a large number of articles, it is opted for the generation of its own *dataset*, on the other hand, some studies choose to use a public dataset. The signal processing is a common factor in all the articles. The use of the Notch filter, band-pass and sample reduction is the order of the day. In addition, as long as the study conditions allow it, the ICA algorithm is also applied.

The features extracted from the signal are very varied. In its temporal dimension, statistical values are extracted (typical deviation, percentiles, etc.) together with more autoregressive models (AR). Next, in the frequency dimension, Power Spectral Density (PSD) is highly used. Finally, a study is also made of both previous domains together (time-frequency domain), with techniques such as Discrete Wave Transform (DWT). In addition to the features of the brain signal, it is also common to extract features of the subject's behaviour (blinking, yawning, reaction time,...).

When it comes to classifying signals, supervised learning models are the most widespread and the ones with the best results in the literature. In particular, Support Vector Machine (SMV) stands out as an algorithm used in both classification and regression, followed quite far by others such as k-Nearest Neighbours or Decision Trees, which are much more popular in other branches of the BCI study. The results obtained with these algorithms take values between 0.5 and 0.15 of RMSE for regression, while the

accuracy is around 80% and 99%. Definitely, the unsupervised learning branch is much less developed in the literature. There are very few studies, and the most widely used algorithms are usually k-Means or Affinity Propagation. Its accuracy range from 67 to 77%. In supervised learning, data is usually hand-labelled by experts or through some self-assessment technique with scales. The use of the subject's reaction time or PERcentage of eye CLOSure (PERCLOS) value is also common.

The aspects that motivated the development of this research include: 1) the lack of studies that analyse the state of the art in terms of sleepiness detection using existing BCI and datasets, 2) the shortage of studies that evaluate the importance of different features from the EEG to accurately detect drowsiness, 3) the lack of studies that compare the performance of different techniques based on ML and, finally, 4) the lack of studies that analyse the suitability of general models to detect sleepiness with BCIs.

For the implementation and subsequent evaluation of the framework, a public dataset is used due to the impossibility during work of using a BCI to acquire data from subjects. The EEG raw data was processed using a 60 Hz Notch filter. Band pass in the 1-30Hz range, reduction of samples to 60 Hz and ICA was applied in order to eliminate remaining noise from blinks, beats, etc. Eight features were extracted from each EEG channel, forming a total of 136 EEG features per epoch. The PSD data was already provided by the processed dataset, so it was not necessary to generate it from the EEG, the average of each frequency band was taken among the 17 channels, giving rise to five PSD features per epoch. Regarding the EOG, the vertical channel data were used and, from these, the number of blinks performed by the subject in each epoch.

In this way, three sources of features were finally obtained, the first one, containing the 136 features of the EEG, the second, using the five features from the PSD, lighter than the previous one and widely used in the literature, is known to give good results. The last one uses the features from the PSD plus the one from the EOG.

Regarding the evaluated algorithms, supervised ML techniques (Support Vector Machine (SVM), k-Nearest Neighbors (kNN), Decision Trees (DT), Gaussian Process (GP) and Random Forest (RF)) and other unsupervised techniques (k-Means, Gaussian Mixture (GM), Bayesian Gaussian Mixture (BGM) and Affinity Propagation (AP)) were used. The selection of algorithms from both families has been based on having alternatives that encompass the best results obtained in the literature, other algorithms that are used but less frequently and, expectedly, with worse results, and finally, algorithms that they rely on the Gaussian properties of the data to make predictions, to check if they perform well with non-Gaussian data (EEG signal). Furthermore, Hyperparameter tuning and cross-validation was used in order to obtain optimal results with each algorithm.

The framework considers the use of individual models for each subject, in addition to a general model trained with the data of all the subjects of the dataset. Two alternatives of general models are also explored for which a certain part of the subjects is reserved for the subsequent evaluation of the model.

Starting with the individual models and regression, the best combination of algorithm

and features was Random Forest making use of both PSD and EOG. By the way, in classification, it also was RF, but this time, using just the PSD. On the other hand, general models generally produced the same results. RF with PSD and EOG was the best option for regression, while RF or kNN, both with PSD, was for classification.

Although SVM is the most used ML algorithm in the literature, this study has evaluated other algorithms, such as kNN and RF, are capable of offering a superior performance using the same data. This has happened in all the models implemented, both individual and general. Added to this, the individual models had the best performance in terms of predictions, followed by the general models. On the other hand, in clustering, despite the acceptable results of the metrics, the levels of drowsiness were not discernible in the grouping that was carried out. Additionally, it was found that all the sets of features used to analyse the signal offer good results, including, although to a lesser extent than the previous ones, the data from the temporal analysis of the signal.

As future work, this study considers the generation of its own dataset using a BCI in order to compare the results. Next, apply Deep Learning techniques to estimate drowsiness, due to its growing popularity and the results, in many cases better, compared to classic ML algorithms. Finally, the integration of this framework for the detection of drowsiness to the framework implemented in previous works by the research team is also proposed.

Keywords: Brain-Computer Interfaces, Electroencephalography, Framework, Machine Learning, Deep Learning.

Índice general

| | | |
|----------|---------------------------------------------------------------------------------|-----------|
| 1 | Introducción | 1 |
| 2 | Background | 4 |
| 2.1 | Neurociencia | 4 |
| 2.1.1 | Neurociencia cognitiva | 4 |
| 2.1.2 | Técnicas para la cuantificación de la somnolencia | 5 |
| 2.1.3 | Aspectos fisiológicos para la medición de la vigilia y la somnolencia | 7 |
| 2.2 | La Interfaz Cerebro-Máquina | 8 |
| 2.2.1 | Electroencefalografía | 10 |
| 2.2.2 | Procesamiento de la señal EEG | 10 |
| 2.2.3 | Clasificación de la señal EEG | 12 |
| 2.2.4 | Electrooculografía | 14 |
| 3 | Estado del arte | 18 |
| 3.1 | <i>Datasets</i> en entornos de somnolencia | 18 |
| 3.2 | Procesamiento de datos cognitivos | 19 |
| 3.3 | Extracción de características | 20 |
| 3.4 | Clasificación de señales cerebrales | 21 |
| 3.4.1 | Eficiencia de sistemas para la detección de la somnolencia | 21 |
| 4 | Análisis de objetivos y metodología | 24 |
| 4.1 | Objetivos del trabajo | 24 |
| 4.2 | Metodología | 25 |
| 5 | Diseño y resolución del trabajo realizado | 27 |
| 5.1 | Escenario de partida | 27 |
| 5.2 | Procesamiento y extracción de características | 28 |
| 5.2.1 | Sobre el EEG | 29 |
| 5.2.2 | Sobre el PSD | 31 |
| 5.2.3 | Sobre el EOG | 32 |
| 5.3 | Clasificación de la señal en base a modelos de Machine Learning | 33 |
| 5.3.1 | Discretización de los valores PERCLOS | 34 |
| 6 | Análisis de resultados | 37 |
| 6.1 | Modelos individuales por usuario | 37 |
| 6.1.1 | Algoritmos de regresión | 37 |

| | | |
|----------|--------------------------------------------------|-----------|
| 6.1.2 | Algoritmos de clasificación | 39 |
| 6.1.3 | Algoritmos de agrupamiento | 41 |
| 6.2 | Modelo general para todos los usuarios | 42 |
| 6.3 | Otros modelos estudiados | 43 |
| 6.3.1 | Modelo 90-10 | 44 |
| 6.3.2 | Modelo 70-30 | 45 |
| 6.4 | Comparativa entre métodos | 45 |
| 6.4.1 | Sobre regresión | 46 |
| 6.4.2 | Sobre clasificación | 47 |
| 7 | Conclusiones y vías futuras | 48 |
| | Bibliografía | 50 |
| | Lista de Acrónimos y Abreviaturas | 59 |

Índice de figuras

| | | |
|-----|------------------------------------------------------------------------------------------------------------------|----|
| 2.1 | Lóbulos cerebrales [1]. | 5 |
| 2.2 | Comparativa entre las señales EEG en estado alerta frente a somnoliento. | 15 |
| 2.3 | Sistema Internacional 10-20 para la colocación de los electrodos [2]. | 16 |
| 2.4 | Visualización de las bandas de frecuencia junto al estado mental al que están asociadas. | 16 |
| 2.5 | Ejemplo de separación en dos clases con SVM [3]. | 17 |
| 2.6 | Ejemplo de electrooculograma en el que se muestran los valores del canal vertical. | 17 |
| | | |
| 5.1 | Disposición de los electrodos en el experimento. | 28 |
| 5.2 | Procesamiento y división en épocas de la señal EEG original. | 30 |
| 5.3 | Procesamiento completo de la señal de un sujeto. | 31 |
| 5.4 | Cálculo de las <i>features</i> asociadas al PSD. | 32 |
| 5.5 | Detección de los parpadeos en el canal vertical del EOG una vez establecido el parámetro <i>thresh</i> | 33 |
| 5.6 | Discretización del PERCLOS en tres clases. | 35 |
| | | |
| 6.1 | Predicción del PERCLOS con Random Forest haciendo uso del PSD y el EOG. | 39 |
| 6.2 | Predicción de las clases PERCLOS. | 41 |
| 6.3 | Matriz de confusión relativa a las predicciones del ejemplo con RF y PSD. | 42 |
| 6.4 | Agrupamiento de los diferentes vectores de características realizado por AP con PSD. | 44 |

Índice de tablas

| | | |
|-----|--------------------------------------------------------------------------------------------|----|
| 2.1 | Escala de Somnolencia de Karolinska en inglés. | 7 |
| 2.2 | Distinción de las bandas de frecuencia cerebrales [4]. | 11 |
| 3.1 | Tabla comparativa entre distintos <i>datasets</i> disponibles. | 19 |
| 3.2 | Comparativa entre trabajos de la literatura analizados. | 23 |
| 5.1 | Distinción de las bandas frecuenciales dentro del <i>dataset</i> SEED-VIG. . | 29 |
| 5.2 | Umrales de PERCLOS establecidos en este estudio. | 35 |
| 6.1 | Comparativa de los algoritmos regresivos probados. | 38 |
| 6.2 | Comparativa de los algoritmos de clasificación probados. | 40 |
| 6.3 | Comparativa de los algoritmos de agrupamiento probados. | 43 |
| 6.4 | Comparativa de los algoritmos regresivos probados para el modelo general. | 43 |
| 6.5 | Comparativa de los algoritmos de clasificación probados para el modelo general. | 44 |
| 6.6 | Comparativa de los algoritmos regresivos probados para el modelo 90-10. | 45 |
| 6.7 | Comparativa de los algoritmos de clasificación probados para el modelo 90-10. | 45 |
| 6.8 | Comparativa de los algoritmos regresivos probados para el modelo 70-30. | 46 |
| 6.9 | Comparativa de los algoritmos de clasificación probados para el modelo 70-30. | 46 |

1 Introducción

La somnolencia es la tendencia de una persona a quedarse dormida, la cual representa una necesidad fisiológica básica al igual que la sed o el hambre [5]. No obstante, en ciertas situaciones puede suponer un gran peligro, siendo un claro ejemplo de ello durante la conducción. La somnolencia interviene entre el 15 y el 30% de los accidentes de tráfico en España, aumentando significativamente el daño sobre el conductor, pues este no es capaz de reaccionar y, por tanto, frenar u oponer una resistencia física al impacto [6, 7, 8]. Además, la somnolencia provoca un incremento del tiempo de reacción, una toma de decisiones más lenta y la predisposición a cometer más errores.

A día de hoy se han propuesto gran cantidad de sistemas con el objetivo de detectar la somnolencia, desde cámaras que monitorizan las expresiones faciales y la apertura de los ojos y la boca, hasta sensores que analizan de forma continua la frecuencia cardíaca y la conductividad eléctrica de nuestra piel avisándonos mediante una fuerte vibración y una luz amarilla cuando nota riesgo de somnolencia [9]. Sin embargo, es probable que estos sistemas produzcan falsas alarmas, ya que estos atributos no siempre van acompañados de somnolencia.

En los últimos años, el campo de la neurociencia cognitiva ha avanzado mucho y se ha demostrado que la EEG, es decir, la técnica capaz de medir los campos eléctricos producidos por la actividad cerebral, es un indicador fisiológico altamente eficaz y objetivo para evaluar los estados de vigilia. Numerosos estudios de EEG sugieren que las actividades delta (0.1-3.5 Hz), theta (4-7 Hz) y alfa (8-13 Hz) están altamente correlacionadas con la fatiga, la somnolencia y el desempeño deficiente de tareas [10, 11, 12]. La solución tecnológica propuesta para la captura y procesamiento de la señal EEG con el fin de estimar el estado de somnolencia de un sujeto al volante es la BCI.

Las BCIs son canales bidireccionales entre el cerebro y dispositivos externos que permiten la monitorización de ondas eléctricas cerebrales. Pueden ser clasificadas en dos categorías en base a su grado de invasividad en el paciente. Por una parte están las invasivas, para las que se requiere de una operación quirúrgica para su implantación y, por otro lado, las no invasivas, que se colocan sobre la cabeza del paciente. Con respecto a las interfaces no invasivas, estas consisten en electrodos colocados en la superficie del cuero cabelludo y se encargan de captar la actividad de grandes grupos de neuronas durante la emisión y transmisión de impulsos eléctricos (EEG). De esta forma es posible monitorizar la actividad del encéfalo y la activación de regiones concretas del mismo.

La señal obtenida con el uso de estas interfaces no invasivas es la superposición de la actividad de todas las neuronas. Por lo tanto se obtiene una medición menos precisa que la proporcionada por las invasivas, ya que el cráneo y piel debilitan la

señal. De la misma manera, estas señales vienen acompañadas de ruido, el cuál debe ser eliminado antes de realizar el estudio de la señal. Sin embargo, la ventaja de evitar la implantación quirúrgica de la BCI reduce de manera drástica los riesgos a los que el sujeto se expone. Además, el coste de implantación es mucho menor. Esto ocasiona que el uso de interfaces no invasivas esté mucho más extendido que el de las invasivas, siendo el tipo de BCI que se emplea para el estudio de la somnolencia.

A pesar de los avances obtenidos por los trabajos existentes en la literatura, algunos retos siguen necesitando mas esfuerzos desde el punto de vista de investigación. Concretamente, hay una falta de trabajos que analicen el estado del arte en cuanto a detección de somnolencia haciendo uso de BCI y *datasets* existentes. Además hay una escasez de trabajos que evalúen la importancia de distintas *features* (o fuentes de datos) provenientes de la BCI para detectar de forma precisa el estado de somnolencia. Otra de las limitaciones es la falta de trabajos que comparen el rendimiento de distintas técnicas basadas en ML desde distintos enfoque (supervisado, no supervisado) sobre *datasets* que traten con Somnolencia y BCI. Por último, es necesario más trabajo enfocado en el análisis de la idoneidad de modelos de ML globales frente a individuales para detectar somnolencia con BCI.

Para mejorar los retos anteriores, este TFG tiene como objetivo general crear un framework inteligente capaz de detectar somnolencia en base al EEG adquirido por una BCI mediante la predicción del PERcentage of eye CLOSure (PERCLOS) del conductor. El PERCLOS es una medida que determina el tiempo que el sujeto permanece con los ojos al menos un 80% cerrados en un intervalo de tiempo determinado. Para lograr esto, es necesario estudiar previamente el estado del arte, definir los conjuntos de características y algoritmos a estudiar, buscar un conjunto de datos, procesar los datos presentes en dicho conjunto, implementar diversos modelos que hacen uso de los diferentes conjuntos previamente definidos y, por último, cotejar los resultados obtenidos con la literatura.

A continuación, con el objetivo de validar el framework, se ha recurrido a un conjunto de datos públicos ante la imposibilidad de usar una BCI real. Se realizó una comparativa entre las opciones disponibles y se seleccionó aquella que mejor se adaptaba a los requisitos impuestos para esta investigación. En base a la revisión previa del estado del arte, se definen los conjuntos de características de la señal y de algoritmos de aprendizaje a implementar en el framework. Dentro de los conjuntos de *features*, se tomaron datos directamente del dominio temporal de la señal EEG, otros tantos, del dominio frecuencial conocido como PSD (la variación de energía que hay dentro del EEG, en función de la frecuencia) y por último del EOG, el cual mide los movimientos de los músculos oculares. Respecto a los algoritmos evaluados, se emplearon técnicas de ML supervisadas (SVM, kNN, DT, GP, RF) y otras no supervisadas (k-Means, GM, BGM y AP). Tanto para los conjuntos de *features* como para los algoritmos, se opta por los que, a priori, en la literatura ofrecen buenos resultados y son muy nombrados. Contrariamente a esto, también se seleccionan *features* y algoritmos que no son tan populares pero que se sospecha que pueden ofrecer buenos resultados. Una vez quedan

definidos, se procesa el *dataset* a fin de obtener las características deseadas.

En el framework se considera el uso de modelos individuales específicos para cada sujeto, además de un modelo general entrenado con los datos de todos los sujetos del *dataset*. También se exploran dos alternativas de modelos generales para las cuales se reserva cierta parte de los sujetos para realizar la posterior evaluación del modelo.

Una vez implementados los diferentes algoritmos, se ha comparado el rendimiento en la parte de aprendizaje supervisado usando las métricas de RMSE y R^2 , mientras que para clasificación se emplea *accuracy*, *precision*, *recall* y *f1-score*. Por su parte, en la rama de aprendizaje no supervisado se usa SC y DBI. A partir de estas métricas, se concluye que el aprendizaje supervisado, tanto regresión como clasificación, ofrecen unos resultados muy buenos, siendo los modelos individuales los que han ofrecido las predicciones más acertadas. En particular, la mejor combinación de regresión, RF con PSD y EOG, presenta un RMSE de 0.084 ± 0.024 . Por otra parte, en clasificación, se vuelve a destacar RF, pero esta vez solo con PSD, ofreciendo un *accuracy* (tasa de acierto) de $85.7\% \pm 6\%$. Finalmente, los algoritmos no supervisados, pese a presentar en algunos casos unas métricas buenas, no han conseguido agrupar los datos de manera que se obtenga una diferenciación entre estados de somnolencia.

Por último se compararon los resultados de la literatura con los del framework desplegado. En estos se identifica que el algoritmo más popular de la literatura, (SVM), se ha visto superado en este framework por otros no tan explorados como son RF y kNN tanto en modelos individuales como generales. Dentro de los conjuntos de características estudiados, PSD ha sido el ganador, aunque en general los tres se han comportado bien, siendo EEG el más novedoso que se ha empleado, aunque también el que peores resultados ha ofrecido.

2 Background

Para comprender el trabajo presentado posteriormente es necesario que el lector conozca qué es la neurociencia cognitiva, cómo se describen los estados de vigilia y somnolencia y qué aspectos clave diferencian uno de otro. Además, es necesario conocer qué técnicas existen para su cuantificación, qué son las BCIs y cómo han ayudado estas últimas al progreso de la neurociencia cognitiva.

2.1 Neurociencia

La neurociencia o neurobiología, como indica Shepherd et al. [13] es un campo de la ciencia que estudia las células del sistema nervioso y su organización dentro de circuitos funcionales que procesan la información y median en el comportamiento. Es fundamental comprender la fisiología interna del cerebro para interpretar posteriormente los datos y las zonas de activación cerebral. En primer lugar, hay que distinguir entre el encéfalo y el telencéfalo, siendo el segundo una parte del encéfalo. Esta estructura cerebral ocupa el 85% total del encéfalo [1] y es a la que comúnmente se hace referencia cuando se habla del término *cerebro*.

El cerebro [14, 15, 16] está a su vez dividido en dos hemisferios, izquierdo y derecho. Cada hemisferio cerebral está dividido, a su vez, en cuatro lóbulos (frontal, temporal, parietal y occipital) nombrados análogamente a su respectivo hueso craneal que lo cubre (ver Figura 2.1). Cada lóbulo se encarga de coordinar diferentes funciones del organismo. El lóbulo **frontal** está involucrado en funciones motoras, en la atención y toma de decisiones, así como en el razonamiento y la conducta, entre otras. El lóbulo **parietal** es el encargado de integrar información de los sentidos así como la manipulación de conceptos u objetos. Por su parte, el lóbulo **occipital** es el centro principal para procesar la visión y darle una interpretación. Por último, el lóbulo **temporal** está relacionado con la información auditiva, en reconocer el lenguaje y formar recuerdos.

2.1.1 Neurociencia cognitiva

La neurociencia cognitiva se describe en [10, 11] como un área de estudio interdisciplinar emergente de dos disciplinas: la psicología cognitiva, que estudia las funciones mentales superiores; y la neurociencia, encargada de estudiar el sistema nervioso que sustenta dichas funciones mentales. Se encarga de comprender la relación entre la existencia de un estado mental a partir de varios conjuntos ordenados de células

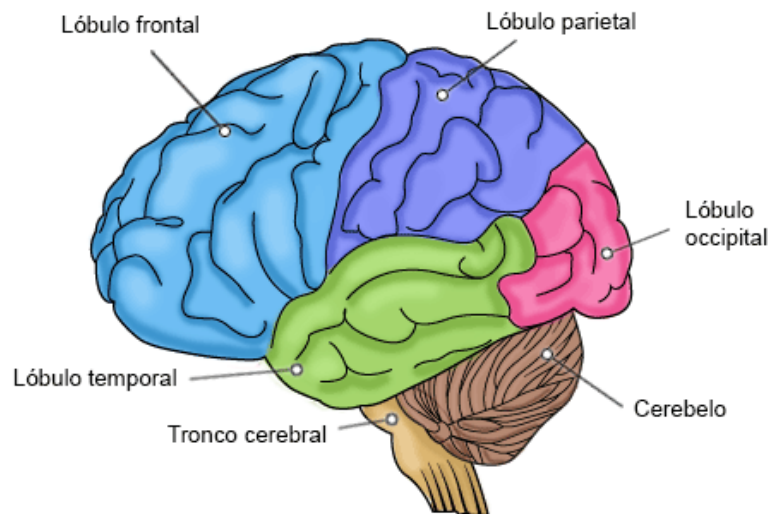


Figura 2.1: Lóbulos cerebrales [1].

con propiedades electrofisiológicas que se comunican gracias a reacciones químicas. Las neuronas, junto con los lóbulos de la corteza cerebral, son el principal objeto de estudio de esta ciencia.

Dentro del ámbito cognitivo, una de las ramas de estudio se centra en identificar qué sucesos dan lugar a un aumento o decremento del estado de alerta y qué aspectos clave se dan a nivel fisiológico en cada uno de los diferentes niveles del mismo. El nivel establecido como *usual* es el estado de vigilia. Se define como un estado consciente, y se caracteriza por un nivel de actividad alto en relación al intercambio de información entre el sujeto y su ambiente. Por otra parte, la somnolencia se define como la habilidad de transición de la vigilia al sueño o como la disminución del estado de alerta dando lugar a las primeras fases del sueño. En este trabajo se estudia qué sucesos dan lugar a un aumento o decremento de este estado y qué aspectos clave se dan a nivel fisiológico en cada uno de estos estados.

Respecto a la somnolencia, se definen en la literatura dos niveles de la misma. Por un lado, la somnolencia **global**, que hace referencia al nivel de adormecimiento característico que presenta el sujeto en su vida cotidiana. Por otra parte, la somnolencia **aguda** determina la tendencia de un sujeto a quedarse dormido en un momento puntual, no persistiendo en el tiempo. De ahora en adelante, para este estudio, cuando se hable de somnolencia se hará referencia al nivel de somnolencia **aguda** de los sujetos.

2.1.2 Técnicas para la cuantificación de la somnolencia

En la literatura se expone que la cuantificación de la somnolencia es un proceso complejo, distinguiendo tres clases de métodos para realizar dicho estudio [5, 17, 18]. La primera es a partir de **mediciones del comportamiento del sujeto** basadas en el bostezo, expresiones faciales, movimientos de la cabeza o frecuencia del cierre

ocular. Dichas mediciones se suelen llevar a la práctica mediante Tests de Vigilancia Psicomotora (PVT), tests de tiempo de reacción o mediciones del PERCLOS, entre otras.

La segunda clase de método consiste en una **autoevaluación con escalas** y, dependiendo del nivel de somnolencia que se desea estudiar, se utilizan unas escalas u otras. Para la somnolencia global se usan la Escala de Somnolencia de Epworth (ESS) y el Inventario de Actividad vigilia-sueño (SWAI), mientras que para la somnolencia aguda, se suelen emplear la Escala de Somnolencia de Karolinska (KSS), la Escala de Somnolencia de Stanford (SSS) y NASA Task Load Index (NASA-TLX). En general, todas las autoevaluaciones mencionadas anteriormente preguntan al sujeto acerca de cómo de despierto se encontraría en diversas situaciones o qué tan somnoliento se ha sentido en los últimos minutos. Por su parte, el sujeto deberá responder a dichas preguntas en base a un rango de valores establecido por el test. En la Tabla 2.1 se puede ver un ejemplo de estas evaluaciones, concretamente, de la Escala de Somnolencia de Karolinska. La autoevaluación con escalas es el procedimiento más sencillo y barato a la hora de realizar la experimentación sobre los sujetos. La parte negativa de estos métodos es la inclusión de un grado de subjetividad por parte del propio sujeto a la hora de autoevaluar su estado.

Por último se tienen los **tests neurofisiológicos**, los cuales van un paso más allá de las auto-evaluaciones, ya que persiguen el objetivo de cuantificar la somnolencia de manera objetiva. Ejemplos de estos son el Test de Latencias Múltiples del Sueño (MSLT) y el Test de Mantenimiento de la Vigilia (MWT). Ambas pruebas requieren la instalación de electrodos en el sujeto, una habitación adaptada y un técnico experto. Los factores previamente mencionados limitan la realización de estas pruebas de forma masiva. A pesar de esto, con el paso de los años se han diseñado soluciones que permiten realizar este tipo de experimentos junto a otros que anteriormente no eran realizables debido a la falta de presupuesto o infraestructura. La solución en cuestión es el uso de sistemas BCI. Estos sistemas son un compromiso entre los test neurofisiológicos (MWT y MSLT) de ámbito sanitario y los estudios realizados por equipos de investigación en los que no se requiere de una metodología tan profesional para su realización.

De los métodos anteriormente mencionados, es necesario conocer el funcionamiento de la Escala de Somnolencia de Karolinska [19] y del valor PERCLOS, ya que serán objeto de estudio a lo largo del trabajo. KSS sirve para medir de forma subjetiva el nivel de somnolencia en un momento concreto. Para ello, el sujeto deberá indicar qué nivel de la escala (ver Tabla 2.1) refleja mejor el estado psicofísico que ha experimentado durante los minutos previos a la realización de la evaluación.

El PERCLOS es una medida psicofisiológica del sujeto que cuantifica el porcentaje de tiempo que ha estado con los ojos al menos un 80% cerrados [20] durante el intervalo de tiempo de la medición (ver Ecuación 2.1). A partir de dicha medida, se establecen diferentes límites en la literatura respecto a un estado de alerta y un estado de somnolencia leve o severa en base al contexto del individuo y de lo crucial que sea la situación en la que se pueda producir la somnolencia.

| Karolinska Sleepiness Scale (KSS) | |
|-----------------------------------|---------------------------------------------------------|
| Rating | Verbal descriptions |
| 1 | Extremely alert |
| 2 | Very alert |
| 3 | Alert |
| 4 | Fairly alert |
| 5 | Neither alert nor sleepy |
| 6 | Some signs of sleepiness |
| 7 | Sleepy, but no effort to keep alert |
| 8 | Sleepy, some effort to keep alert |
| 9 | Very sleepy, great effort to keep alert, fighting sleep |

Tabla 2.1: Escala de Somnolencia de Karolinska en inglés.

Respecto al uso de PERCLOS como medida del nivel de alerta de un sujeto, hay controversia en el mundo científico acerca de si es o no una medida fiable para dicha labor. Por una parte, se tienen artículos defendiéndola [21, 22], mientras que por otra, también alegan que hay otras técnicas que ofrecen mejores resultados [23]. Lo que sí es innegable, es que se han diseñado gran cantidad de sistemas basados en esta métrica que, posteriormente, han sido analizados y validados por otras entidades [24, 25, 26].

$$PERCLOS = \frac{\text{tiempo_ojos_cerrados}(\geq 80\%)}{\text{tiempo_total}} \quad (2.1)$$

2.1.3 Aspectos fisiológicos para la medición de la vigilia y la somnolencia

Los diferentes métodos para la cuantificación de la somnolencia (ver Sección 2.1.2) y la incorporación de las BCIs (ver Sección 2.2) han permitido la realización de un mayor número de experimentos. Esto ha permitido relacionar la actividad fisiológica recogida a partir de los electrodos y de otras mediciones del estado del sujeto (movimiento de la cabeza, conteo de parpadeos, ritmo cardiaco, etc.) con los diferentes niveles de actividad reflejados en estos datos y, por lo tanto, con el nivel de vigilia o somnolencia del sujeto.

La técnica más extendida a día de hoy para trabajar con señales cerebrales es la Electroencefalografía (EEG), pudiendo analizar las señales adquiridas mediante su división en bandas de frecuencia, también denominados ritmos (para saber más detalles ver Sección 2.2.1). Gracias a todos estos experimentos, se ha podido relacionar el estado de vigilia quiescente, situación en la que el sujeto está quieto pero tiene la capacidad de moverse, con una actividad cerebral de fondo de bajo voltaje. En dicha actividad predominan los ritmos alfa y beta. De la misma manera, se han relacionado los primeros síntomas de la somnolencia y las primeras fases del sueño con la desaparición de

actividades gamma y beta; junto con un incremento en las alfa, theta y delta. Esto se aprecia claramente en la Figura 2.2, en la que se presentan dos fragmentos de las ondas cerebrales de un paciente, la primera en estado de alerta, mientras que la segunda se corresponde a un estado somnoliento. Comparándolas, se aprecia claramente que las ondas son más variables cuando se está alerta, es decir, se dan más frecuencias altas, además de que su amplitud es mayor que en estado de somnolencia. De forma adicional, se detecta la aparición de *spindles* del sueño y complejos K. Ambos elementos se definen como pequeñas y breves fluctuaciones que se dan en la señal de manera repentina y con una forma característica en cada caso. Por último, queda añadir que en estado de somnolencia, se aumenta el número de parpadeos y el tiempo que se tienen los ojos cerrados, además de la aparición de bostezos [10, 11, 12, 27, 28].

2.2 La Interfaz Cerebro-Máquina

Las Interfaces Cerebro-Máquina, en inglés *Brain-Computer Interfaces*, son tecnologías que han logrado avanzar en el estudio de la neurociencia cognitiva a pasos de gigante. Estos sistemas brindan la capacidad de analizar neuronas o grupos de estas de forma específica a través de electrodos, permitiendo así ampliar el conocimiento sobre el funcionamiento de las mismas y de cómo las diferentes actividades complejas asociadas a sus respectivos lóbulos de la corteza cerebral interactúan entre sí para que estas se den.

Las BCIs [16, 29, 30, 31] son dispositivos cuya finalidad radica en la lectura de la actividad cerebral a través de electrodos siendo uno de sus objetivos el **monitorizar** la actividad de forma pasiva y poder así analizar patrones de actividad cognitiva, motora o emocional. Del mismo modo, permite **evaluar o diagnosticar** las capacidades cognitivas o emocionales en base a patrones cerebrales detectados permitiendo identificar anormalidades. Finalmente, habilitan **interaccionar** con dispositivos, procesando la actividad cerebral en tiempo real. Si el objetivo de la interacción es el ámbito médico de la rehabilitación, entonces se denomina **intervención** [32].

Existen dos tipos de BCIs en base a la forma en la que se adquieren las señales y, en concreto, al grado de invasividad de los electrodos en el sujeto. El primer tipo son las BCIs **invasivas**, donde los electrodos son implantados en las distintas capas del cráneo dependiendo de su propósito. Es por esto que se requiere de una intervención quirúrgica para introducirlos. Son técnicas mucho más delicadas que sus contrapuestas pero a cambio ofrecen una calidad máxima de resolución y precisión de las muestras. Este tipo de técnicas se emplean en el tratamiento de enfermedades neurodegenerativas que afectan varias actividades que el cuerpo realiza, como el equilibrio o el movimiento. Por otra parte, el segundo tipo de interfaces son las **no invasivas**, siendo la electroencefalografía la técnica más común, no requiriendo una intervención quirúrgica. Sus aplicaciones son mucho más extensas ya que ofrecen una mayor facilidad a la hora de realizar experimentos con diferentes sujetos.

En las BCIs no invasivas, los electrodos se sitúan directamente en el cuero cabelludo, dando lugar a señales más imprecisas en cuanto a especificidad, ya que a día de hoy un electrodo no invasivo es incapaz de monitorizar neuronas individualmente, sino que recoge la actividad de grupos neuronales. Por otra parte, la calidad de las señales también se ve afectada debido a dificultades añadidas, como el movimientos del sujeto o la interposición del cabello con los electrodos. Respecto a la disposición de los electrodos en los sistemas no invasivos, lo más común es emplear el **Sistema Internacional 10-20** (ver Figura 2.3), el cuál define la colocación de los electrodos extracraneales según indica Malmivuo et al. [2]. Se toman como puntos de referencia el nasión, que es el punto de intersección del hueso frontal y de dos huesos nasales del cráneo humano, que se manifiesta en forma de hendidura; y el inión, definido como la proyección más prominente del hueso occipital en la parte posterioinferior del cráneo humano. A partir de estos puntos, se miden los perímetros del cráneo que forman los planos sagital y transversal. La colocación de los electrodos estará determinada por la división de esos perímetros en intervalos del 10% y del 20%, de ahí el nombre del sistema. Además, también señala que es común la colocación de otros electrodos en zonas intermedias a las especificadas por este sistema.

Un sistema BCI sigue las siguientes fases de funcionamiento independientemente del hardware, software o tipo de adquisición de la señal que se emplee:

1. Generación y posterior adquisición de la señal EEG a partir de los electrodos dispuestos en el sujeto.
 2. Envío de datos cerebrales a un dispositivo de control. Los protocolos de comunicación que se suelen emplear son Bluetooth Low Energy (BLE), en la comunicación del casco con el controlador; y User Datagram Protocol (UDP) o Lab Streaming Layer (LSL) para transmitir los datos entre el controlador y otros dispositivos de la red para su posterior procesamiento.
 3. Procesamiento de la señal para aumentar su calidad y eliminar artefactos, es decir, distorsiones, adiciones o errores que no tiene correlación con el sujeto o región anatómica estudiada. Esta fase se explica con detalle en la Sección 2.2.2.
 4. Extracción de características, en inglés *features*. Una vez se tiene la señal procesada, esta se trocea en épocas (en inglés *epochs*) y, a partir de estas secciones, se genera para cada una un vector con la información relevante de esa porción de la señal, que será el objeto de estudio en los diferentes experimentos que se realicen.
 5. Clasificación de la señal en base al vector de características. En este punto es importante todo el trabajo realizado previamente ya que la calidad de los resultados obtenidos dependerá directamente del procesamiento previo de la señal y de la selección de características a estudiar. En la Sección 2.2.3 se detallan las diferentes técnicas que se emplean en esta fase. Es importante no confundir el proceso
-

de clasificación de la señal con el término *clasificación* empleado en aprendizaje computacional, el cual describe una rama del aprendizaje supervisado.

6. Interfaz de control que traduce las señales clasificadas en alertas acerca del estado del sujeto o en comandos significativos para un escenario de aplicación particular.

2.2.1 Electroencefalografía

La electroencefalografía [33], conocida por sus siglas EEG del inglés, es una técnica capaz de detectar y analizar la actividad cerebral, basada en el registro de actividad bioeléctrica. El procedimiento se efectúa mediante la colocación de sensores llamados electrodos en el cuero cabelludo. Cada electrodo envía una señal a la BCI, siendo capaz de mostrar la fluctuación rítmica de la actividad eléctrica del cerebro (ondas cerebrales) visualmente como una línea ondulante. De esta manera, es posible el estudio de las ondas cerebrales de los sujetos en toda clase de experimentos.

Es importante destacar que las señales EEG de los sujetos poseen un comportamiento **no estacionario**, ya que sus propiedades (amplitud y frecuencias) varían con el tiempo [34]. Además, estas señales son **no lineales**, puesto que no obedecen a las propiedades de superposición y escalado, refiriéndose a que dos ondas EEG que se superponen en el tiempo no producen una onda producto de las dos anteriores multiplicadas por ciertos factores [35, 36]. Finalmente, se les atribuye un comportamiento **no gaussiano**, es decir, que la agregación de estas no tiende a una distribución Gaussiana de los datos. Estos aspectos influirán posteriormente en partes de su procesamiento y de la extracción de características.

Estas señales pueden ser analizadas en diferentes bandas de frecuencia (ver Tabla 2.2). El motivo de esta división es el de relacionar distintos estados cognitivos del sujeto con diferentes niveles de actividad cerebral, siendo los ritmos de más baja frecuencia (theta y delta) los característicos en etapas de sueño. Posteriormente, le sigue el ritmo alfa, el cuál se asocia a un estado de relajación del sujeto y a la meditación, siendo mucho más apreciable cuando el paciente tiene los ojos cerrados. Los últimos dos ritmos, beta y gamma, implican estados mentales de concentración y de carga mental grande e, incluso, situaciones de estrés en el caso de las gamma. Este desglose, junto con una representación de cada ritmo, queda detallado en la Figura 2.4.

Hay alternativas al EEG, como es la **Magnetoencefalografía** (MEG), que detecta la actividad cerebral, pero en este caso, mediante los campos magnéticos producidos por las corrientes eléctricas en el cerebro. Con esta técnica se consigue mayor precisión de la actividad cerebral pero tiene un mayor coste económico el realizarla.

2.2.2 Procesamiento de la señal EEG

Una parte fundamental a la hora de trabajar con señales fisiológicas y, en este caso, con EEG, es que estas suelen venir acompañadas de señales no deseadas que dificultan

| Nombre | Frecuencia (Hz) | Relación con la vigilia |
|--------------------------|-----------------|--------------------------------------------------------|
| Ritmo Delta (δ) | 0.1 - 3.5 | Sueño profundo. |
| Ritmo Theta (θ) | 4 - 7.5 | Sueño ligero y sueño REM. |
| Ritmo Alfa (α) | 8 - 13 | Estado de relajación y poco nivel de actividad mental. |
| Ritmo Beta (β) | 14 - 30 | Estado mental de concentración. |
| Ritmo Gamma (γ) | > 30 | Actividades cerebrales intensas y estrés. |

Tabla 2.2: Distinción de las bandas de frecuencia cerebrales [4].

la legibilidad de la actividad cerebral y cuyo origen suele ser ajeno al cerebro. Es por esto que se precisa de una etapa previa de procesamiento de la señal antes de proceder a la extracción de características. De no ser así, los resultados obtenidos se verían afectados en cierta medida por este ruido añadido dando lugar a posibles resultados incorrectos.

En la literatura se describen los subsecuentes pasos a la hora de procesar la señal. Se comienza con una etapa de filtrado de la señal mediante los filtros **Notch** y **paso banda**. El filtro Notch o de *muesca*, como bien indica su traducción al castellano, elimina una frecuencia específica que crea interferencias en el circuito eléctrico. En particular, es de gran utilidad para eliminar los artefactos producidos por la frecuencia de la red eléctrica, la cuál es de 50 Hz en Europa y de 60 Hz en América. Por otra parte, el filtro paso banda tiene como función permitir el paso de un rango concreto de frecuencias mientras que el resto son obviadas.

Posteriormente se procede con una etapa de reducción del número de muestras, en inglés (***downsampling***), con el objetivo de reducir el tamaño de los datos. Si a la vez que se aplica este proceso se desea evitar la pérdida de calidad de la señal, es necesario tener en cuenta el Teorema de muestreo de Nyquist-Shannon. Este teorema enuncia que para reconstruir una señal de manera exacta, es necesario que la frecuencia de muestreo sea al menos el doble que la mayor frecuencia que existe en esa señal.

Tras reducir la frecuencia de muestro de la señal, se aplica el **Análisis de Componentes Independientes (ICA)**. Este algoritmo tiene como objetivo encontrar una representación lineal de datos de modo que los componentes sean estadísticamente independientes, o tan independientes como sea posible. Para poder aplicar ICA, es necesario que las subcomponentes que la formen sean estadísticamente independientes y no gaussianas, cosa que cumple el EEG. [37].

Para comprender de forma sencilla qué se logra con este algoritmo, se plantea una conjetura denominada *Cocktail Party Problem* [38]. En la versión más simplificada, se expone un escenario en el que dos personas tienen una conversación. Cada una de ellas posee un micrófono que graba la conversación, dando como resultado dos grabaciones en las que está presente la voz de ambos sujetos. El problema que se plantea a continuación es, ¿Cómo podemos separar las dos voces en cada archivo para obtener grabaciones aisladas de los hablantes? Para resolver esta situación, se hace uso de ICA.

ICA es de gran utilidad en procesamiento del EEG, ya que permite eliminar de la señal los artefactos que no se ha podido eliminar con el filtrado previo como podrían ser latidos del corazón o parpadeos. Aunque hay ocasiones, por ejemplo, la estimación en tiempo real del estado del sujeto, en las que el coste computacional de ICA no es admisible y por lo tanto no se aplica.

2.2.3 Clasificación de la señal EEG

Tras obtener el vector de características a estudiar, se procede a clasificar la señal en base a dicho vector en búsqueda de resultados. En la literatura se distinguen dos tipos de estudio. En primer lugar, se tienen los **estudios analíticos**, caracterizados por recoger ciertas *features* con el fin de estudiar el comportamiento que presentan ante las distintas condiciones a las que se someten los sujetos. Un ejemplo de este tipo de estudio sería la periódica realización de tests de autoevaluación con escalas. Estos tests se realizan de forma simultánea a alguna prueba mental, como puede ser el uso de un simulador de conducción. Posteriormente se contrastan los distintos estados reportados por los sujetos con sus características medidas, estableciendo así relaciones entre ellas. De este modo, una posible relación sería: *“Cuando el sujeto presenta un valor superior a 7/9 en KSS, el valor de la banda alfa siempre aumenta respecto a su valor normal”*.

Por otra parte los **estudios predictivos** tienen como objetivo la correlación entre *features* y diferentes estados del sujeto a través de modelos de ML. De esta forma, gracias a modelos entrenados, sería posible hacer estimaciones acerca del estado de ese mismo sujeto e, incluso, sobre otros. Dichos modelos tienen finalidades de lo más diversas, desde la estimación del estado emocional hasta el nivel de somnolencia o de carga de trabajo mental que tiene el sujeto.

Como se indica en [39], si la técnica de ML empleada es aprendizaje supervisado, es necesario adjuntar una **etiqueta** (y) a cada vector de características (X). De esta forma, el algoritmo debe establecer una relación entre la forma de los datos de entrada y la salida que se espera en base a la experiencia que va adquiriendo durante su fase de *entrenamiento* o *supervisión*. Dentro del aprendizaje supervisado se tienen dos tipos de modelos. Los primeros son los modelos de **regresión**, los cuales hacen predicciones sobre valores continuos como el precio o la edad. Al contrario que los modelos regresivos, los modelos de **clasificación** realizan estimaciones sobre valores discretos, por ejemplo, tipos de animales, el idioma de un texto o el tiempo atmosférico. La técnica contrapuesta al aprendizaje supervisado, es el aprendizaje no supervisado. Para este último, no se requiere del etiquetado de los datos, ya que es el propio algoritmo el que los agrupa en base a patrones comunes que detecta dentro de los distintos datos que se le suministran. Esta técnica se conoce en inglés como *clustering*.

Es necesario añadir que en el estudio cognitivo, hay un algoritmo que destaca sobre el resto respecto a su uso. Dicho algoritmo es **SVM** y se emplea tanto para clasificación como para regresión. Su funcionamiento se basa en buscar un hiperplano, es decir, un subespacio con una dimensión más baja que su espacio ambiental. En el caso de un

plano, sería una recta. El objetivo del hiperplano es separar de forma óptima a los puntos de varias clases [3, 40]. Debido a que opera a través de *kernels*, es excelente para resolver problemas no lineales, como es el caso de las señales EEG. En particular, el *kernel* más usado en el ámbito cognitivo es Radial Basis Function (RBF). En la Figura 2.5 se ilustra el funcionamiento de SVM con datos de dos clases y el hiperplano que realiza la división.

En los estudios predictivos, cada vez es más popular el uso de Deep Learning (DL). El aprendizaje profundo [39, 41], es una rama del Machine Learning en la que se emplean algoritmos inspirados en la estructura y el funcionamiento del cerebro para generar modelos de aprendizaje, denominados Artificial Neural Networks (ANNs), conformando el núcleo del DL. Estos modelos se basan en una colección de nodos, llamados *neuronas*, interconectadas entre sí y organizadas por capas, unas visibles y otras ocultas. Cada neurona, es una pequeña unidad de cálculo, formada por una serie de entradas y que produce una única salida. Esta salida, a su vez, puede ser enviada a múltiples neuronas. Por su parte, las entradas pueden ser las *features* de entrada del problema o bien, las salidas de otras neuronas. Por último, las neuronas de la última capa de la red neuronal, son las encargadas de devolver el resultado de la estimación, como por ejemplo, reconocer un objeto en una imagen [42].

Una vez se tiene un modelo entrenado, es necesario analizar qué tan bueno es. Para ello, se proporciona un conjunto de métricas específicas para cada tipo de aprendizaje [43]. Gracias a estas métricas, no solo es posible comparar los resultados que se obtienen al emplear diferentes algoritmos, sino que también permite compararlos con los resultados de otros trabajos. En los modelos regresivos, las métricas más populares son el RMSE, encargado de medir la cantidad de error que existe entre los datos originales (etiquetados) y las predicciones que se han realizado; cuanto más cercano a cero, mejor serán las predicciones realizadas. De forma complementaria al RMSE, se emplea Coeficiente de correlación (CC) o R^2 para determinar el grado de relación que se da entre las *features* empleadas y los resultados obtenidos; la mejor puntuación posible es 1.0 y puede ser negativa. Gracias a CC y R^2 , se tiene una medida de qué tan buenas serán las predicciones que realice el modelo ante muestras que no haya visto previamente.

Por otra parte, en los modelos de clasificación, las métricas cambian. De forma general se emplean las siguientes:

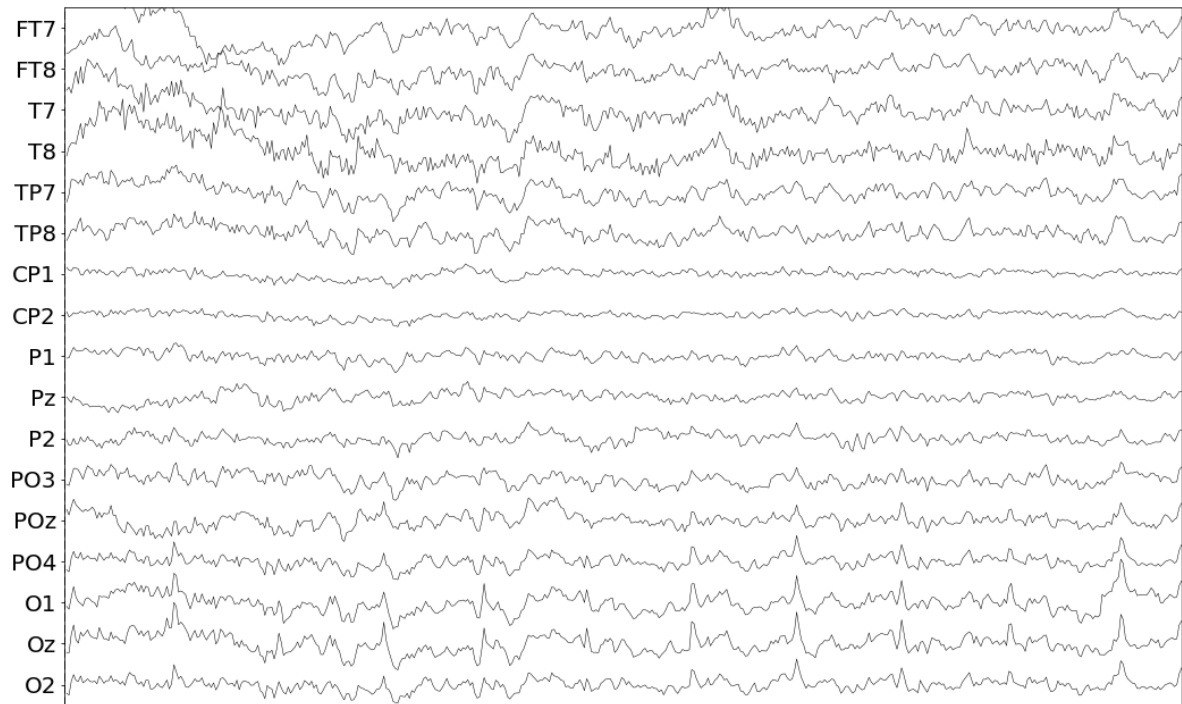
- **Accuracy:** Es el ratio de predicciones que se han realizado de forma correcta. Un valor de uno indica que todas las predicciones son correctas.
 - **Precision:** Mide el ratio entre el número de aciertos y el total de aciertos más falsos positivos. Cuanto más próxima a uno sea, menos falsos positivos producirá el modelo.
 - **Recall:** Esta medida es similar a precision, pero en este caso, el ratio medido es entre el número de aciertos y el total de aciertos más falsos negativos. Por lo tanto, cuanto más próxima a uno sea, menos falsos negativos producirá el modelo.
-

- **F1-score:** Es una métrica derivada de precision y recall, calculada a partir de una media ponderada entre ambas y, de manera análoga, el mejor valor posible es uno.

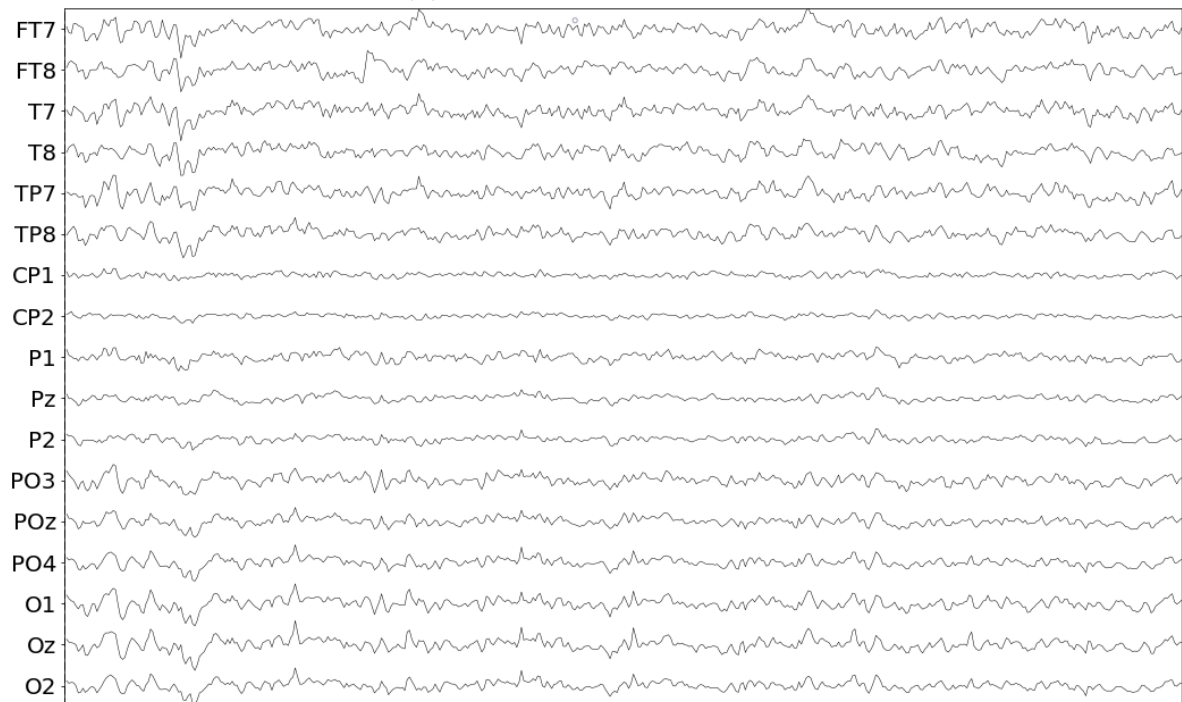
Por último, es importante destacar que evaluar el rendimiento de un algoritmo de agrupación no es tan trivial como contar el número de errores o la *precision* y el *recall* de un algoritmo de clasificación supervisado. Se busca determinar si esta agrupación define separaciones de los datos similares de forma significativa [44]. Algunas métricas que miden estos aspectos son SC, con valores entre $[-1, 1]$ e indica qué tan bien están definidos los clústers, donde un mayor valor refleja una mejor definición. Mientras que DBI mide la separación entre los clústers que se definen. En este caso, cero es el mejor valor.

2.2.4 Electrooculografía

De forma adicional al EEG, se dispone de otras técnicas con el objetivo de recopilar señales fisiológicas de los sujetos. Una de ellas es la **Electrooculografía** [45], conocido como EOG por su acrónimo en inglés. Esta técnica permite medir los movimientos de los músculos oculares mediante el registro de la diferencia de potencial existente entre la córnea y la retina haciendo uso de electrodos (de forma similar al EEG) tal y como se muestra en la Figura 2.6. Cada pico de potencial se corresponde con un parpadeo del sujeto. Gracias a esta técnica es posible monitorizar en los sujetos aspectos relacionados con el parpadeo y movimiento ocular sacádico (movimientos rápidos del ojo). Es interesante añadir que esta técnica no es incompatible con la anterior, se puede realizar de forma simultánea un EEG y un EOG de un sujeto.



(a) Sujeto en estado de alerta.



(b) Sujeto en estado de somnolencia.

Figura 2.2: Comparativa entre las señales EEG en estado alerta frente a somnoliento.

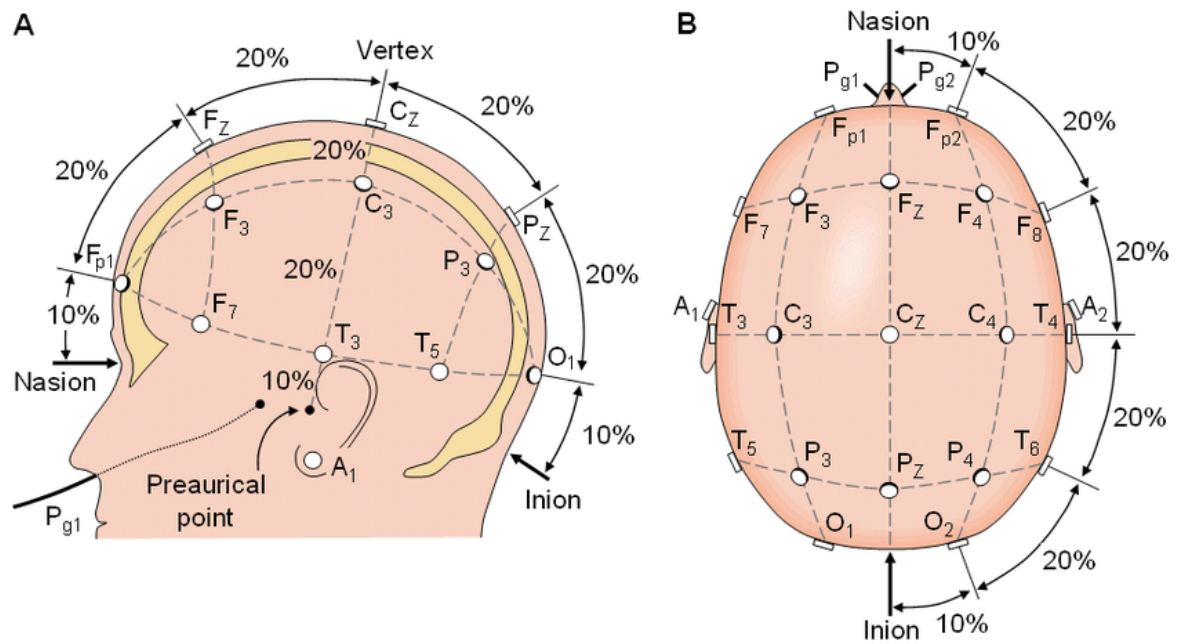


Figura 2.3: Sistema Internacional 10-20 para la colocación de los electrodos [2].

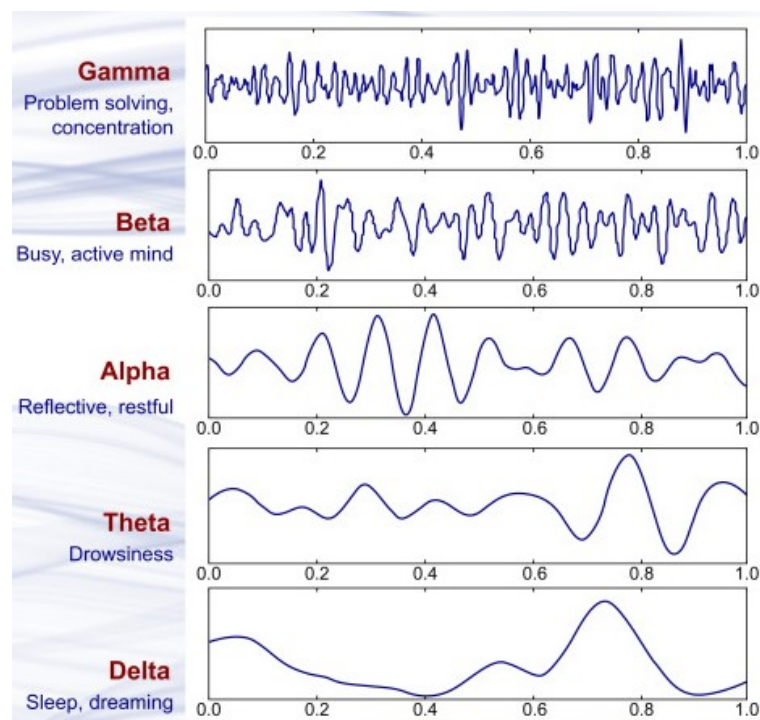


Figura 2.4: Visualización de las bandas de frecuencia junto al estado mental al que están asociadas.

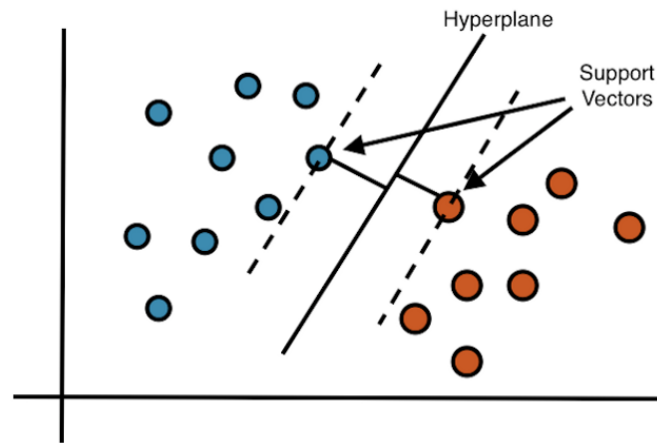


Figura 2.5: Ejemplo de separación en dos clases con SVM [3].

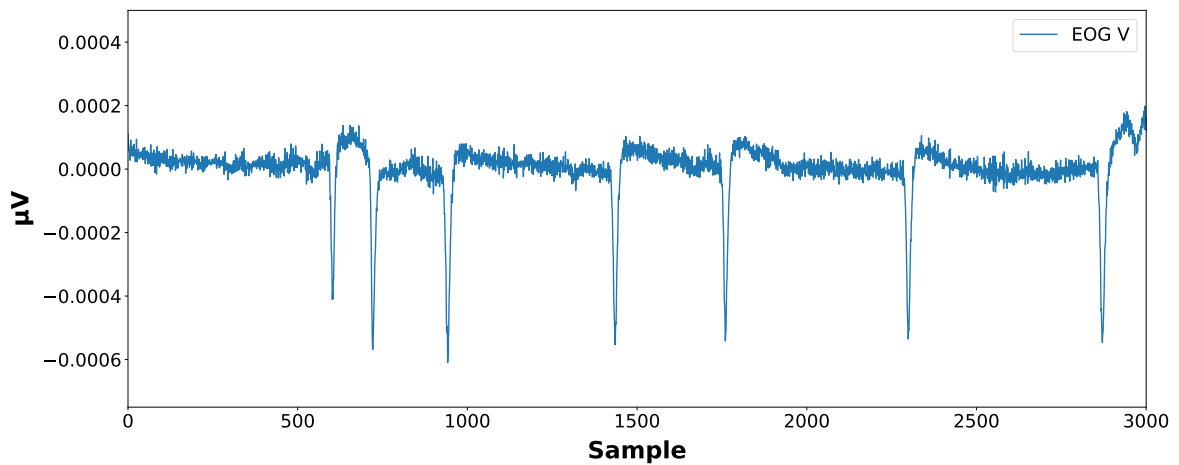


Figura 2.6: Ejemplo de electrooculograma en el que se muestran los valores del canal vertical.

3 Estado del arte

En esta sección se analiza cómo en la informática se han adoptado los conocimientos teóricos de la neurociencia cognitiva y qué metodología de trabajo siguen los experimentos realizados, tanto a nivel analítico como predictivo. Además, se ha estudiado qué características se extraen de los sujetos en dichos experimentos, qué procedimientos se han utilizado para clasificar la señal y qué tan buenos han sido los resultados obtenidos. En el ámbito informático, no es extraño ver que ambos términos, somnolencia y fatiga, se engloban dentro del término *cansancio* a la hora de realizar estudios y establecer soluciones.

3.1 *Datasets* en entornos de somnolencia

El punto de partida de toda experimentación que hace uso de BCIs es la generación y adquisición de los datos. Estos datos se suelen organizar de manera que conforman un conjunto de datos, denominado *dataset* en inglés, del que se quieren obtener ciertos resultados de cara a probar una hipótesis o de realizar un estudio. En gran cantidad de los artículos de este ámbito, se opta por la generación de un *dataset* propio, mediante la construcción de un escenario de estudio y la participación de varios sujetos en el mismo. Esto permite establecer las condiciones propicias para el experimento y recopilar todos los datos de interés deseados. Por otra parte, hay otros estudios que optan por el uso de un conjunto de datos público, ya sea para contrastar resultados, para agilizar el proceso, o por falta de infraestructura.

En la Tabla 3.1, se muestran varios *datasets* públicos relacionados con el estudio del estado cognitivo y la somnolencia. Concretamente, en [46], se sometió a 27 sujetos a un estudio de atención sostenida, en el que debían realizar dos sesiones de conducción. Cada sesión constaba con una duración de, al menos, 90 minutos. El objetivo, en ambas sesiones, era corregir la trayectoria del vehículo siempre que este se desviase del carril. La primera sesión se realizó sin proporcionar al sujeto un feedback quinestésico, es decir, el sujeto no percibía una sensación de movimiento del vehículo. De este modo, tenía que guiarse únicamente de su sentido de la vista para corregir la posición del coche. Sin embargo, en la segunda sesión sí que se le proporcionaba este feedback. Por otro lado, para los *datasets* [47, 48] también se hizo uso de un simulador de conducción. En el primero de los dos mencionados, se proporciona para cada uno de los 12 sujetos dos grabaciones de su EEG relativas a su estado cognitivo, una clasificada como alerta y otra como cansado. En cambio, para el segundo *dataset* mencionado se realizó uno o dos

experimentos por persona. Todos los experimentos se realizaban bajo las condiciones de una conducción monótona y con el objetivo de recopilar las señales fisiológicas de los sujetos. Por último, en [49], se realizaron dos mediciones para los diez participantes, cada una con 150 segundos de duración. La primera, con los ojos abiertos y, la segunda, cerrados. Los sujetos se encontraban en estado de reposo en el momento de la grabación, habiendo realizado previamente unos 100 minutos de tareas de atención sostenida.

| Dataset | Sujetos | Canales EEG | Canales EOG | Etiquetado | Descripción |
|---------|---------|-------------|-------------|-----------------------|-------------------------------------------------------------------------------------------------------------------|
| [46] | 27 | 32 | × | Posición del vehículo | Dos sesiones de 90 minutos en un simulador. En ambas se mide el tiempo de reacción al desviarse. |
| [47] | 12 | 40 | × | × | Ofrece dos ficheros de EEG para cada sujeto, referentes a su estado cognitivo: uno alerta y otro cansado. |
| [48] | 21 | 17 | 6 | PERCLOS | Una o dos sesiones de 118 minutos en un simulador. Cada ocho segundos se toma un registro del PERCLOS del sujeto. |
| [49] | 10 | 64 | 3 | × | Cinco minutos de señal EEG tras realizar 105 minutos de tareas que suponen un esfuerzo mental. |

Tabla 3.1: Tabla comparativa entre distintos *datasets* disponibles.

3.2 Procesamiento de datos cognitivos

Una vez se tiene el conjunto de datos a estudiar, es momento de procesar dichos datos, con el objetivo de eliminar el ruido, distorsiones y otras imperfecciones que puedan darse. Esto permite obtener unos datos de mayor calidad que, posteriormente, podrán ser analizados de manera más concluyente. A la hora procesar las señales EEG se observa que, en la práctica, se sigue una metodología clara, tal como queda descrito en la Sección 2.2.2. Por consiguiente, el filtrado paso banda, la reducción del número de muestras y, en determinadas ocasiones, el ICA se convierten en un factor en común a lo largo de la literatura. Ejemplos de trabajos siguiendo esta metodología son [50, 51, 52, 53, 54].

3.3 Extracción de características

Con los datos previamente procesados, se procede a extraer las características a estudiar. Se observan tres fuentes de extracción de características a partir de las señales EEG. Cada fuente se corresponde con un dominio en el que se puede estudiar dicha señal, como queda señalado en [4] y [55], junto con la información recopilada en la Tabla 3.2. El primero y más común de estos tres dominios es el temporal y, a partir de este, se obtienen *features* en base a modelos matemáticos y otras operaciones algebraicas. El más común y extendido es el modelo *autoregresivo*, consistiendo en una representación de un proceso aleatorio, en este caso, la señal EEG, en el que la variable de interés depende de sus observaciones pasadas. A partir de este modelo nacen otras alternativas como los *adaptativos autoregresivos* [56, 57]. Otra alternativa novedosa en el dominio temporal es la del estudio de la señal a partir de Cuaterniones, una extensión de los números reales que añaden las unidades imaginarias i , j y k , que permiten representar objetos en un espacio tridimensional en base a su orientación y rotación lo cual se ha demostrado que es útil para el estudio del EEG. También es común extraer de la señal valores estadísticos como la varianza, la desviación típica y los cuantiles, entre otros [54].

Gracias a la teoría de Fourier, es posible transformar una señal a su representación en el dominio frecuencial, haciendo uso de la Transformada Rápida de Fourier (FFT). De esta manera, es posible analizar qué frecuencias predominan en la señal original y con qué amplitud. A esto se le conoce como PSD y suele medirse por bandas de frecuencia (ver Tabla 2.2). El PSD representa la variación de energía que hay dentro de una señal, en función de la frecuencia por unidad de masa [58]. En otras palabras, el PSD muestra para cada frecuencia si la energía presente es mayor o menor. En gran cantidad de experimentos se extraen características del PSD, ya que tienen relación directa con las bandas frecuenciales y, por lo tanto, dan buenos resultados a la hora de estimar el estado cognitivo [51, 54, 57, 59, 60, 61].

Por último, y debido al comportamiento no estacionario, no lineal y no gaussiano que presentan las señales EEG (ver Sección 2.2.1), es de utilidad disponer de una representación y descomposición de la información frecuencial de la señal que esté vinculada al dominio temporal (dominio tiempo-frecuencia). Es por esto que se emplean métodos como la Transformación de Onda Discreta (DWT), encargada de transformar la señal de partida en un vector de coeficientes, usando un conjunto de funciones base ortogonales llamadas *wavelets* [57, 62, 63].

De forma adicional a las características de la señal cerebral, también es común extraer otras características de interés del sujeto. Entre ellas se encuentran el ritmo cardiaco, la frecuencia del parpadeo y el número de parpadeos [64, 65]. También se opta por la detección de *alpha spindles* [52, 53], debido a la relación que tienen con la somnolencia (ver Sección 2.1.3); y potenciales evocados, que se definen como picos positivos de voltaje, causados por estímulos auditivos, visuales o somatosensoriales inesperados [16, 66].

3.4 Clasificación de señales cerebrales

A la hora de clasificar las señales se dan varios aspectos en común. El primero, que la mayoría de las veces se emplea aprendizaje supervisado y, el segundo, que dentro del aprendizaje supervisado, es común que se haga uso de varios algoritmos en concreto [55]. El más popular con diferencia es **SVM** [12, 51, 54, 56, 61, 63, 65], detallado en la Sección 2.2.3. Seguido a este, se hace también uso de modelos lineales (*Ridge*, *Logistic Regression* o *Lasso*), *Naive Bayes* y *Nearest Neighbors* [55, 67, 68, 69]. En menor medida, pero con más popularidad en otras ramas del estudio del EEG, se opta también por k-NN, LDA, *regression trees* y *random forest* [55, 70, 71].

Con mucha diferencia, el *clustering*, que forma parte del aprendizaje no supervisado, es la rama de ML menos explorada en cuanto a la estimación de la somnolencia, siendo sus resultados tampoco los más destacables. Pero, es de interés conocer que se usan algoritmos como k-Means, AP y GM, con tasas de acierto entre el 67 y el 77% [72, 73, 74].

En cuanto al enfoque del Deep Learning, las ANNs más empleadas son *Recurrent Self-Evolving Fuzzy Neural Networks (RSEFNNs)*, *Convolutional Neural Networks (CNNs)* y *Long Short-Term Memory (LSTM)* [55, 57, 59, 63, 64]. Cada vez están cobrando más importancia ya que producen unos resultados, en muchos casos, superiores a los métodos tradicionales de este ámbito [63, 71].

Para el estudio de la somnolencia, las etiquetas que se emplean para los problemas de regresión son los propios valores medidos mediante KSS, NASA-TLX o PERCLOS (ver Sección 2.1.2). Mientras que, si el problema es resuelto con un modelo de clasificación, las etiquetas generadas son discretizaciones, en distintos niveles de somnolencia, de los valores de las etiquetas que se usan en los métodos regresivos.

3.4.1 Eficiencia de sistemas para la detección de la somnolencia

En el marco de los estudios predictivos, se puede observar una gran variedad de resultados que se muestran con detalle en la Tabla 3.2. Concretamente, se especifica el conjunto de características, el etiquetado de los datos y los algoritmos de ML empleados. Tras estos campos, se indican los resultados obtenidos. Es interesante destacar que el porcentaje de acierto de estos toma valores entre el 80% y el 99%, dependiendo del experimento. Los mejores valores alcanzados se logran empleando SVM, ELM y RSEFNN. En los artículos para los que se empleó un método regresivo, los resultados se ofrecen en base al RMSE, con resultados entre 0.15 y 0.5, siendo el mínimo valor posible alcanzable el 0, equivalente a que no se ha cometido ningún fallo. De forma complementaria al RMSE, también se indica el CC o el R^2 dependiendo del experimento. Ambas métricas se emplean para medir el grado de relación que se da entre las *features* empleadas y los resultados obtenidos.

De manera particular, cabe destacar los trabajos llevados a cabo en [51, 63], ya que en ambos estudios se etiquetaron las señales EEG a mano, gracias a expertos

neurólogos. Además, se empleó SVM como método de aprendizaje para la clasificación de la señal. Las tasas de acierto (accuracy) obtenidas por los modelos son de 94.7% y 99.30% respectivamente, presentando diferencias en el vector de características usado. En el primer caso, se optó por la extracción de cuatro *features* del dominio tiempo-frecuencia con la DWT juntos con otras medidas oculares, mientras que en el segundo se seleccionaron cuatro *features* para cada una de las bandas de frecuencia δ , θ , α y β del PSD.

Continuando con los métodos regresivos, [54] consiguió unos resultados de 0.15 en RMSE y un R^2 de 0.83, haciendo uso de un SVM recurrente. Ambos valores son los mejores de su categoría a lo largo de todos los recopilados. En dicho experimento, se tomaron características de la dimensión temporal (parámetros de Hjorth) y otras tantas de las cinco bandas del PSD (dimensión frecuencial). Para el etiquetado, se utilizó autoevaluación con escalas, concretamente la Escala de Somnolencia de Karolinska (KSS). También es de interés destacar [70, 71], ya que en ambos estudios se realizó una comparativa entre distintos algoritmos que no son tan nombrados en este ámbito, como pueden ser kNN, DT y RF. En ambos casos, los datos eran etiquetados por el equipo de investigación. La tasa de acierto en el primer estudio, haciendo uso de DT es de 74.7%, mientras que con *ensemble boosted trees* era de 77.3%. También se probó con SVM, logrando una tasa del 76.7% en el mejor de los casos. Para el segundo estudio, con RF se obtuvo una tasa del 81.4% mientras que para SVM era del 78.6%. En ambos experimentos, se llegó a la conclusión de que hay alternativas mejores a SVM.

Por su parte, los algoritmos de Deep Learning se están usando cada vez con más frecuencia en estos estudios, generando unos resultados muy buenos, llegando a superar a los algoritmos tradicionales en cada vez más ocasiones. Un ejemplo de esto es [63], ya que con el mismo vector de categorías y el mismo etiquetado que para su experimento con SVM se obtuvo una tasa de acierto del 97.30% frente al 96.90% de SVM. El algoritmo empleado para obtener estos resultados fue Extreme Learning Machine (ELM) y, además del mejorar la tasa de acierto, se reporta en dicho estudio que el tiempo de entrenamiento del SVM fue de 42.010 segundos mientras que para la ELM fue de tan solo 0.166 segundos. Otro estudio a destacar es [59], ya que se tenía como objetivo hacer una estimación en tiempo real, por lo que no se aplicó ICA en el procesamiento de los datos. Para realizar la estimación se extrajeron las *features* del PSD y el etiquetado era el tiempo de reacción del conductor antes los desvíos. Como modelo de aprendizaje se optó por una *RSEFNN* y se obtuvo una tasa de acierto del 96.18%.

| Referencia | Características | Etiquetado | Algoritmo | Resultados |
|----------------------------|-------------------------------------------------------------------------------------------------------|------------------------------------|------------------------------|----------------------------|
| Shen et al. (2008) [60] | PSD $\delta, \theta, \alpha, \beta$ | AVT | Multi-class SVM-c | 85.40% |
| Hu et al. (2009) [65] | 11 del EOG | KSS | SVM-c con RBF | 80.74% |
| Yeo et al. (2009) [51] | 16 del PSD $\delta, \theta, \alpha, \beta$ | A mano por expertos | SVM-c | 99.30% |
| Lin et al. (2014) [61] | PSD $\delta, \theta, \alpha, \beta$ | Tiempo de reacción | SVM-r con kernel función RBF | R^2 :0.207 RMSE:0.816 |
| Garcés et al. (2014) [57] | Tres del EEG, diez del PSD y seis de la WT | A mano por expertos | Linear classification | 72.90% |
| | | | ANN | 85.50% |
| Igasaki et al. (2015) [67] | Ritmo cardiaco y de respiración | KSS | Logistic regression | 84% |
| Chen et al. (2015) [63] | Cuatro de la WT | A mano por expertos | SVM-c | 96.90% |
| | | | ELM con RBF | 97.30% |
| Liu et al. (2016) [59] | PSD $\delta, \theta, \alpha, \beta$ | Tiempo de reacción | RSEFNN | 96.18% |
| Gwak et al. (2018) [70] | PSD $\delta, \theta, \alpha, \beta$, ritmo cardiaco, EOG y datos respecto al simulador de conducción | Expresiones faciales y rendimiento | SVM | 78.60% |
| | | | kNN | 75.30% |
| | | | Random Forest | 81.40% |
| Cuui et al. (2019) [68] | PSD θ, α | Tiempo de reacción | kNN | CC:0.439 RMSE:0.268 |
| | | | Ridge regression | CC:0.504 RMSE:0.362 |
| | | | MLPNN | CC:0.2756 RMSE:0.5422 |
| Akbar et al. (2019) [54] | Tres del EEG y PSD $\delta, \theta, \alpha, \beta, \gamma$ | KSS | SVM-r | R^2 :0.64 RMSE:0.56 |
| | | | Recurrent SVM-r | R^2 :0.83 RMSE:0.15 |

Tabla 3.2: Comparativa entre trabajos de la literatura analizados.

4 Análisis de objetivos y metodología

La presente sección presenta los objetivos del TFG, así como la metodología seguida a lo largo del trabajo.

4.1 Objetivos del trabajo

La premisa con la que parte este TFG es la recopilación, análisis y procesamiento de los datos fisiológicos (EEG y EOG) del sujeto conductor del vehículo que, posteriormente, serán empleados para crear un framework inteligente capaz de detectar la somnolencia en escenarios de conducción y evitar así incidentes automovilísticos. Para poder llevar a cabo dicho objetivo principal, se ha desglosado este mismo en una serie de subobjetivos más específicos:

- Realizar una revisión del estado del arte acerca del estado cognitivo dentro del ámbito de la neurociencia y de las Interfaces Cerebro-Máquina (BCIs), aclarando el por qué estas son de interés en el estudio del estado cognitivo e identificando las limitaciones de los trabajos existentes.
- Distinguir y extraer del estado del arte qué características de interés, qué tipos de modelos y algoritmos de ML, y qué resultados se han obtenido a la hora de realizar estudios y estimaciones en el ámbito de la somnolencia al volante.
- Examinar y seleccionar de entre los distintos *datasets* públicos, relacionados con el estudio del estado cognitivo y la vigilia en el ámbito de la conducción, el que mejor se adapte a los requerimientos de este TFG.
- Analizar y procesar los datos de interés del *dataset* seleccionado para generar varios conjuntos de características de interés.
- Haciendo uso de las características extraídas, entrenar varios modelos de ML con el propósito de identificar el estado de somnolencia del sujeto, ofreciendo dos comparativas diferenciadas. En la primera, el objetivo es identificar qué algoritmo de los empleados dentro de un mismo tipo de modelo ofrece mejores prestaciones. Por otra parte, en la segunda, se busca determinar qué tipo de modelo de aprendizaje funciona mejor en el ámbito aplicado.
- Comparar los resultados obtenidos en este TFG con los que ya se tienen en el estado del arte.

4.2 Metodología

En esta sección se detalla la metodología seguida para cumplir los objetivos marcados en la sección anterior. La estructura seguida está definida en base a los objetivos propuestos y, principalmente, por el orden cronológico en el que las tareas se han desarrollado.

En primer lugar, se ha realizado un estudio previo acerca del funcionamiento del cerebro y de la neurociencia cognitiva. Concretamente, se ha indagado en los aspectos fisiológicos que definen los estados de vigilia y somnolencia, junto con las diferentes técnicas para su detección. Seguido a esto, ha sido necesario conocer el funcionamiento de las BCIs y las distintas fases que conforman su ciclo de funcionamiento. Posteriormente, se investigó sobre la electroencefalografía, ya que es la técnica de adquisición de datos neurológicos más común en BCI. En particular, se analizó qué técnicas de procesamiento de la señal y qué características se pueden extraer de esta señal que sean de interés para el estudio de la somnolencia.

Tras adquirir todos estos conocimientos previos, se llevó a cabo un estudio del estado del arte mediante el que se identificaron una serie de retos que fueron la base del trabajo realizado.

Este estudio se divide en varias etapas, correspondientes a las fases de funcionamiento de la BCI. Para la fase de adquisición de datos se optó por la búsqueda de un conjunto de datos público debido a la imposibilidad de usar una BCI real para adquirir los datos, por lo que fue necesario realizar una comparativa entre las diferentes opciones disponibles. Tras esto, se analizaron los métodos de procesamiento de los datos y la posterior extracción de características que se efectuaban en la literatura. En consecuencia a estos datos, se optó por un procesamiento con filtrado Notch y paso banda, reducción de muestras y, por último, ICA. Para la extracción de características se decidió obtener varios conjuntos de estas, procedentes del EEG, PSD y EOG.

Una vez se tenían los datos procesados y las características extraídas, se estudió qué algoritmos eran los más empleados en el estado del arte y qué resultados se habían obtenido en la etapa de clasificación de la señal. Con esta información, se definió cómo sería dicha etapa en este trabajo. Se especificó que se entrenarían modelos de regresión, clasificación y agrupamiento (*clustering*). Dentro del aprendizaje supervisado (regresión y clasificación) se seleccionó SVM, ya que es el algoritmo más popular en este ámbito. Adicionalmente, se seleccionaron otros algoritmos que, a priori, no son tan extendidos en su uso u ofrecieron peores resultados en la literatura. Concretamente, se optó por kNN, DT, GP y RF. Por otra parte, en el aprendizaje no supervisado, se escogió k-Means, GM, BGM y AP.

Adicionalmente, se concretó que el estudio se centraría principalmente en modelos de aprendizaje individualizados para cada sujeto y, posteriormente, se plantearían diferentes alternativas para conformar un modelo general capaz de estimar el estado de todos los sujetos.

Tras definir cómo sería la etapa de clasificación de la señal, se procede a la implemen-

tación de los diferentes algoritmos y a evaluación de sus resultados en base a diferentes métricas particulares a cada tipo de modelo. Concretamente, se hizo uso del RMSE y de R^2 para los modelos regresivos, mientras que, en los modelos de clasificación, se midió *accuracy*, *precision*, *recall* y *f1-score*. Por último, se empleó SC y DBI para evaluar los modelos de agrupamiento. En todos los casos, se buscó que la eficiencia del algoritmo fuese máxima. Para ello, fue necesario un ajuste de hiperparámetros junto con validación cruzada.

Se realizó una comparativa entre los algoritmos de un mismo modelo en pos de encontrar las combinaciones de algoritmos y características que ofrecían los mejores resultados. A continuación, se hizo un análisis más general en el que se compararon los modelos (regresión, clasificación o *clustering*). Esto permitió concluir qué tipos de modelo y qué algoritmos en particular ofrecen unas mejores prestaciones.

En base a esta información se implementaron una serie de modelos generales agrupando la información de todos los sujetos. Adicionalmente, se exploró la posibilidad de crear modelos entrenados con un subconjunto de los usuarios, reservando los usuarios restantes para evaluación. Para concluir, los resultados obtenidos fueron comparados de manera análoga a los modelos individuales previos y esto permitió concluir qué algoritmos y tipo de modelo se comportaban mejor al generalizar.

Por último, teniendo ya los resultados obtenidos en este trabajo, se realizó una comparativa entre los resultados del framework y los consultados en el estado del arte. Esto permitió identificar contribuciones del trabajo sobre la literatura existente, como el mejor rendimiento de algoritmos no tan usados en la literatura frente a los más típicos, además de exponer que los modelos individuales ofrecen una mejor estimación respecto a los generales y que los conjuntos de características estudiados ofrecieron, en general, buenos resultados.

5 Diseño y resolución del trabajo realizado

Esta sección detalla los procesos de selección, extracción de datos de interés e implementación de las distintas alternativas seguidos durante la elaboración del framework para la detección de la somnolencia. Finalmente se realiza un proceso de comparativa de los resultados obtenidos.

5.1 Escenario de partida

El punto de partida consiste en la adquisición de un conjunto de datos basado en la detección de la somnolencia en escenarios de conducción que alimentará el framework propuesto para este trabajo. Mediante la comparativa realizada en la Tabla 3.1), se concluye que el *dataset* SEED-VIG [48] es el idóneo para este trabajo, ya que contiene datos acerca del EEG y del EOG. En ambos casos, con un buen número de canales, permitiendo así la extracción de varios conjuntos de características de interés que posteriormente alimentarán a los distintos modelos de ML. Adicionalmente, la duración de los experimentos es más que aceptable. Además, posee los datos etiquetados, recogiendo los valores PERCLOS de los sujetos. Cabe destacar que es muy difícil encontrar un buen *dataset* público y etiquetado en el ámbito del estado cognitivo referido a la conducción. Así, este etiquetado de los datos permite entrenar modelos de ML supervisados y, por lo tanto, hacer una mejor comparativa de algoritmos respecto al estado del arte.

Este conjunto de datos es cedido por la universidad Jiao Tong de Shanghai para su uso académico. Para generarlo, se desplegó un simulador de conducción, teniendo como objetivo el recopilar las señales EEG y EOG de forma continua, junto con un registro PERCLOS cada **ocho segundos**. Se realizaron un total de 23 experimentos, repartidos entre 21 sujetos distintos. El escenario de conducción consistía en una carretera de cuatro carriles, en línea recta, durante prácticamente la totalidad del experimento y las mediciones se realizaban generalmente después de la hora de comer. El motivo detrás de estas decisiones es, en ambos puntos, inducir somnolencia en los sujetos. Como se advierte en todos los manuales de conducción [75], dichas situaciones son las más propensas para quedarse dormido al volante. La duración de cada experimento es exactamente de 118 minutos, lo que da lugar a un total de $\frac{118 \times 60(s)}{8(s)} = 885$ registros PERCLOS en cada sesión. Para la medición del EEG y EOG se usó un dispositivo de la

empresa Neuroscan [76]. De forma adicional, se empleó un dispositivo de eye-tracking de la compañía SensoMotoric Instruments con el fin de monitorizar el movimiento de los ojos y medir el PERCLOS [77].

En el *dataset*, se ofrecen los datos en crudo (*raw*) de los distintos experimentos, junto con una serie de datos ya procesados. A continuación, se indican sólo aquellos que se han empleado durante el estudio realizado durante el presente trabajo:

- Datos en crudo de 17 canales EEG. Los electrodos se colocaron siguiendo el Sistema Internacional 10-20. En la Figura 5.1 se puede apreciar, de manera 2D y 3D, la disposición y el nombre de los electrodos que se emplean en el experimento. Los indicadores **azul** y **verde** señalan la zona preauricular derecha y el nasión respectivamente, mientras que los **rojos** indican la posición de los electrodos. Ambas imágenes han sido generadas haciendo uso de la librería MNE de Python.

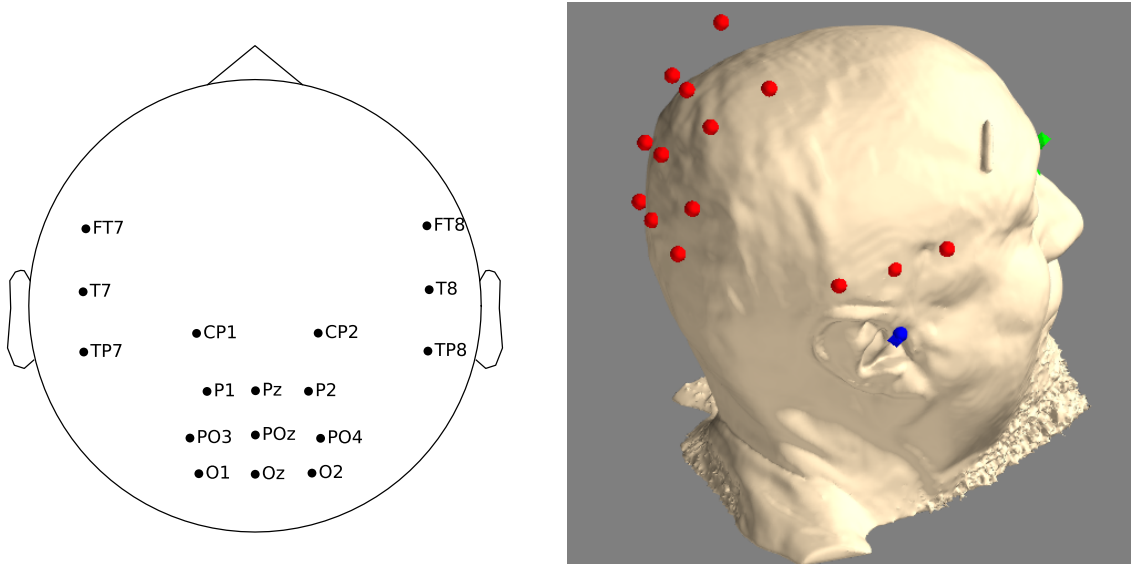


Figura 5.1: Disposición de los electrodos en el experimento.

- Datos en crudo de los canales vertical y horizontal del EOG.
- Datos de PSD calculados con una media móvil [78], relativos a las cinco bandas de frecuencia del EEG que se detallan en la Tabla 5.1.

5.2 Procesamiento y extracción de características

Una vez se tiene el *dataset*, es momento de procesar los datos y generar vectores de características que serán empleadas por el framework. A partir de los datos disponibles,

| Banda | Frecuencia Mínima (Hz) | Frecuencia Máxima (Hz) |
|-------|---------------------------|---------------------------|
| Delta | 1 | 4 |
| Theta | 4 | 8 |
| Alfa | 8 | 14 |
| Beta | 14 | 31 |
| Gamma | 31 | 50 |

Tabla 5.1: Distinción de las bandas frecuenciales dentro del *dataset* SEED-VIG.

se ha decidido extraer *features* de la propia señal EEG, ya que son sencillas de obtener. Sin embargo, este procedimiento no ha sido ampliamente usado en la literatura. Por otro lado, también se ha hecho uso del PSD para extraer la cantidad de energía presente en cada banda de frecuencia. Al contrario que las anteriores *features* extraídas, para estas sí que es ampliamente conocido que ofrecen unos muy buenos resultados. Por último, ya que se dispone del EOG de los sujetos, se ha optado por contabilizar el número de parpadeos del sujeto en cada intervalo de ocho segundos y, con esto, complementar las *features* correspondientes a las cinco bandas de frecuencia. En total, se obtienen tres conjuntos de características. El primero, referente al EEG, mientras que el segundo y el tercero hacen uso del PSD únicamente, o junto al EOG.

Se ha optado por *Python* como lenguaje de programación a la hora de extraer las *features* y clasificarlas posteriormente. Este lenguaje, que es ampliamente usado en investigación sobre BCI, ofrece un gran repertorio de librerías que lo complementan. Particularmente, en la fase de extracción de características se ha usado *MNE* para la representación y procesamiento de las señales EEG y EOG, junto con *NumPy*, librería matemática para obtener las características de interés de la señal. Una vez se tienen las *features*, se ha empleado *sklearn*, concretamente el módulo de preprocesado, con el objetivo de escalar las *features* a un rango concreto de valores. Adicionalmente, se ha usado *Pandas* para exportar e importar ficheros CSV con dichos vectores de características generados. Para terminar, a fin de generar todo tipo de elementos visuales, como gráficas e histogramas, se ha optado por *Matplotlib*.

5.2.1 Sobre el EEG

Dado que se tienen 17 canales muestreados a una frecuencia de 200 Hz, junto con una etiqueta cada ocho segundos, se hace inviable que los vectores de características sean la señal como tal, debido al gran tamaño que estos tendrían. Es por esto que se opta por obtener ciertos valores estadísticos para cada canal. De esta forma, se obtiene un resumen más liviano que caracteriza a ese intervalo de ocho segundos. El paso previo a realizar para obtener estos vectores es, como se especifica en la Sección 2.2, el procesamiento de los datos en crudo del EEG (ver Sección 2.2.2), con el fin de eliminar artefactos de la señal y, por consiguiente, focalizar los datos de interés que se desean

extraer del mismo.

La Figura 5.2 muestra un esquema de los pasos que se han seguido para el procesado de la señal EEG que, adicionalmente, son detallados a lo largo de este apartado. Asimismo, se presenta en la Figura 5.3 un ejemplo del antes y el después del tratamiento que recibe la señal. En dicho ejemplo, es posible apreciar la mejora en claridad que se obtiene en la señal, tanto en el dominio temporal como en el dominio espectral.

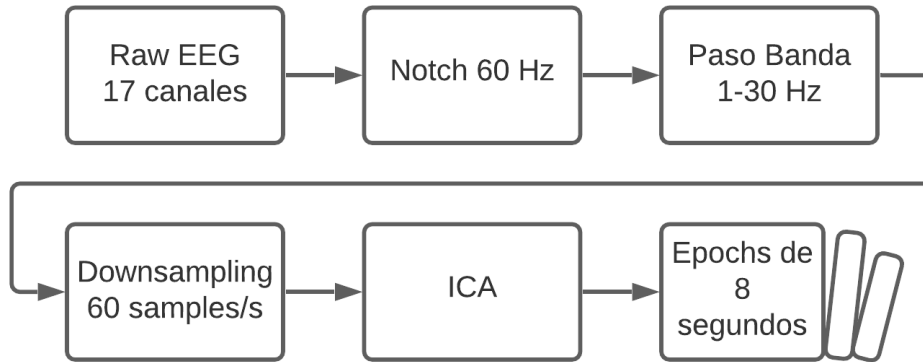


Figura 5.2: Procesamiento y división en épocas de la señal EEG original.

Inicialmente, se realiza un tratamiento de la señal con dos técnicas de filtrado. En primer lugar, se aplica un filtro Notch a 60 Hz con la intención de eliminar los artefactos producidos por la frecuencia de la red eléctrica. En segundo lugar, se emplea un filtro paso banda. En particular, se aplicará en un rango comprendido entre 1 y 30 Hz, ya que es el rango de frecuencias de interés para el estudio de la vigilia, como queda señalado en [10, 12]. Es interesante destacar que, al aplicar un filtro paso banda 1-30 Hz, es posible omitir el primer filtrado con Notch, debido a que posteriormente, la frecuencia de la señal quedará limitada a un máximo de 30 Hz.

Tras este primer filtrado, se obtiene una señal con una frecuencia máxima de 30 Hz. No obstante, inicialmente fue muestreada a 200 Hz, por lo que actualmente se encuentra *sobremuestreada*. Es decir, se tenía más información de la necesaria para describir la onda sin perder calidad. Por lo tanto, se aplica una reducción de muestras, aplicando el Teorema de muestreo de Nyquist-Shannon. Esto quiere decir que, si previamente se realizó un filtrado paso banda 1-30 Hz, el límite de frecuencia de muestreo que conservará la calidad y por lo tanto toda la información original, será de 60 Hz.

Una vez filtrada la señal y reducido el número de muestras de la misma, se aplica ICA. Gracias a este paso, se consigue reducir el ruido de los datos que no se ha podido eliminar con el filtrado previo, como podrían ser latidos del corazón o parpadeos. Tras realizar este procesamiento y, como se puede apreciar en la Figura 5.3, se obtiene una señal con una mayor claridad de los datos respecto a la señal de partida.

Después de esta etapa de procesado, se realiza una división de la señal en porciones o épocas, en inglés *epochs*, de ocho segundos. Ya que, como se ha indicado previamente en los detalles técnicos del *dataset* (ver Sección 5.1), se tiene una etiqueta de PERCLOS

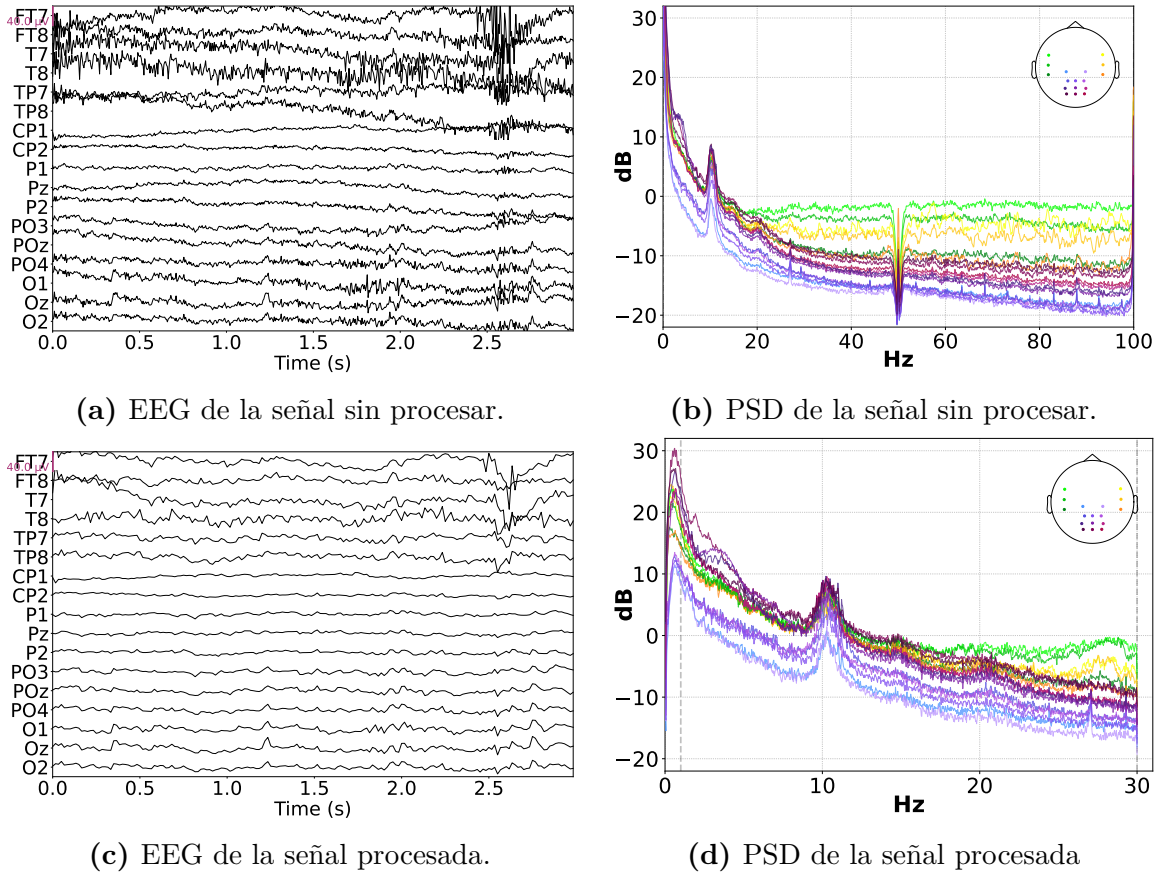


Figura 5.3: Procesamiento completo de la señal de un sujeto.

cada ocho segundos de grabación. En particular, se medirán en cada época las siguientes *features* para cada uno de los 17 canales: media, desviación típica, varianza, percentil 5, primer cuartil, mediana, tercer cuartil y percentil 95.

Como se ha mencionado anteriormente en esta sección, si no se realizase esta extracción de valores estadísticos, se tendrían que manejar vectores de características de tamaño $60(\frac{\text{muestras}}{\text{segundo}}) \times 8(\text{segundo}) \times 17(\text{canales}) = 8160$. Esto sería computacional muy costoso a la hora de trabajar con vectores de características en la fase de clasificación de la señal. Gracias a la simplificación realizada, se tienen vectores de tamaño $17 \times 8 = 136$, que es una cantidad asumible.

5.2.2 Sobre el PSD

Las características extraídas del PSD son cinco, una correspondiente a cada banda de frecuencia descrita en el *dataset* (ver Tabla 5.1). En este caso, el conjunto de datos de partida proporciona la media móvil del PSD (ver Sección 5.1) para cada uno de los 17 canales EEG y, dentro de cada canal, para las cinco bandas de frecuencia. Debido

a esto, no será necesario realizar dicha transformación en la señal, ya que esta viene facilitada.

El procedimiento de extracción, como queda ilustrado en la Figura 5.4, consiste en tomar de cada época los valores PSD de las cinco bandas para los 17 canales. Una vez se tienen estos valores, se calcula el promedio para cada una de las bandas. Posteriormente, se escalan las *features* obtenidas al rango [0-1], haciendo uso de un escalador *MinMax* de *sklearn*, obteniéndose así cinco *features* referentes a la PSD para cada época, una para cada banda de frecuencia.

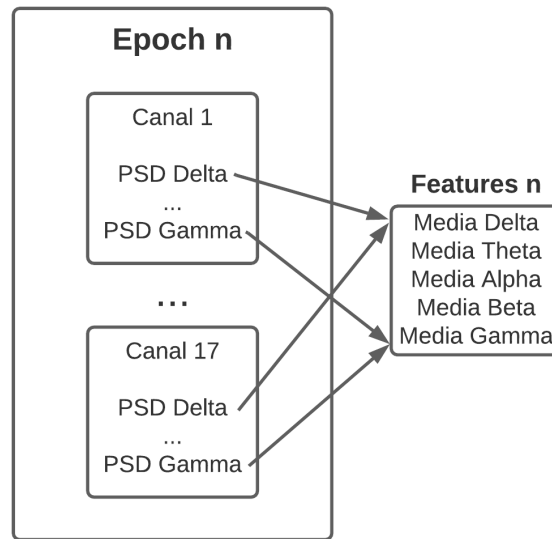


Figura 5.4: Cálculo de las *features* asociadas al PSD.

5.2.3 Sobre el EOG

Se desea extraer una característica de la señal EOG, esta es, el número de parpadeos que realizan los sujetos. Con este objetivo en mente, se comienza con el procesado de la señal. El primer paso a dar es el de cargar en una estructura de la librería MNE los datos EOG. Seguido a esto, se debe hacer uso de la función `find_eog_events` del módulo `preprocessing`. Esta función necesita el parámetro clave `thresh` para detectar de forma correcta cuándo hay un parpadeo, el cual establece el mínimo de amplitud que debe tener la señal para que se contabilice esa porción como un parpadeo. Para calcular este valor límite, se toman los primeros 40 segundos de señal EOG y se calcula el percentil 2.5, el cual se establece como el parámetro `thresh`.

El percentil 2.5 se establece en base a la literatura [27], donde se destaca que una persona promedio parpadea 19.86 veces por minuto mientras conduce, por lo que, en 40 segundos, realizará unos 13.24 parpadeos. Por ello, en el intervalo de 40 segundos, con una frecuencia de muestreo de 125 Hz a la que se tomó el EOG, se tienen 5000

muestras, de las que $\frac{13.24}{5000} = 0.2648\%$ son supuestos parpadeos. Teniendo esto en cuenta y, sabiendo que el parpadeo no es instantáneo, se le añade a este una duración de unas diez muestras para su pico de potencial, obteniendo finalmente como resultado un valor de 2.648%, que finalmente queda simplificado a 2.5%.

En la Figura 5.5 se muestra los parpadeos detectados durante los 24 primeros segundos. La detección es realizada en el canal vertical de la señal EOG. Las líneas verticales delimitan las distintas épocas, mientras que las horizontales discontinuas indican el parámetro **thresh** establecido.

Gracias a este procesamiento, se obtienen los parpadeos que se producen en la señal EOG completa. El siguiente paso a dar consiste en tomar la marca de tiempo asociada a cada parpadeo y mapearla a la época que le corresponde. Obteniendo así, finalmente, el número de parpadeos que los sujetos realizan en cada intervalo de ocho segundos. Este valor es escalado mediante *MinMax*, entre los valores [0-1], de forma similar al PSD. De esta manera, puede ser tratado posteriormente por los distintos modelos de ML de manera correcta.

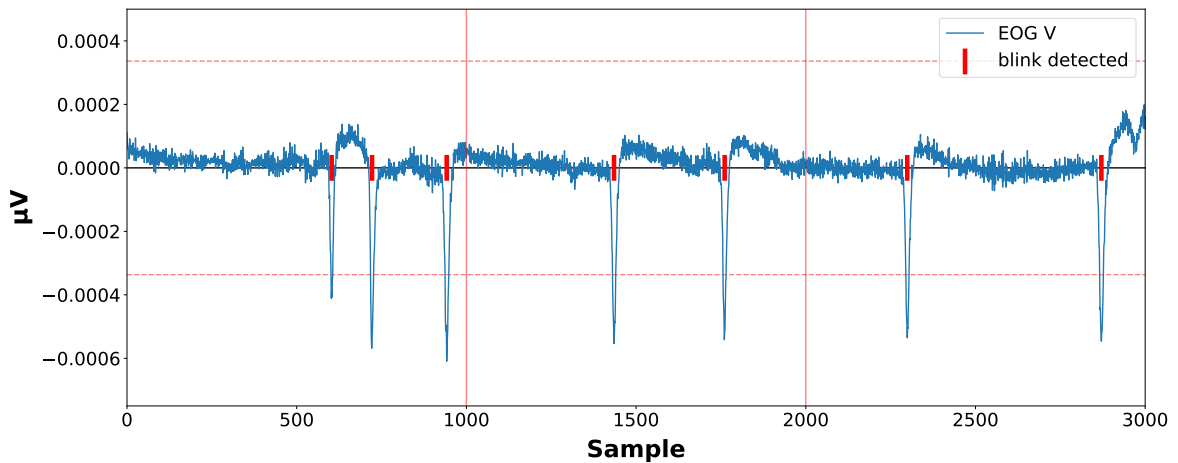


Figura 5.5: Detección de los parpadeos en el canal vertical del EOG una vez establecido el parámetro **thresh**.

5.3 Clasificación de la señal en base a modelos de Machine Learning

Una vez se tienen los distintos conjuntos de *features* que se quieren estudiar, da comienzo la etapa de clasificación de la señal. En esta etapa se procede al entrenamiento de distintos algoritmos de ML, con el objetivo de realizar una comparativa a nivel de modelo (regresión, clasificación o *clustering*) y otra, más precisa, respecto a los distintos algoritmos dentro de un mismo modelo de aprendizaje.

Es necesario establecer de antemano qué tipos de modelos y qué algoritmos de ML serán implementados en el framework. La decisión ha sido tomada en base a lo expuesto en la literatura (ver Sección 3.4). Para el aprendizaje supervisado, se ha optado en primer lugar por SVM ya que es el algoritmo más popular en el estudio de la somnolencia al volante y, en general, ofrece buenos resultados. Adicionalmente, se ha optado por otros algoritmos que no tienen tanto peso en la literatura y que, en principio, no suelen ofrecer unos resultados tan buenos, estos son: kNN, DT y RF. Por último, se ha considerado de interés incluir GP, ya que este se basa en una generalización de la distribución Gaussiana [79], mientras que las señales EEG presentan un comportamiento no gaussiano (esto queda detallado en la Sección 2.2.1), por lo que, a priori, las predicciones no deberían ser buenas. Al contrario que para el aprendizaje supervisado, las opciones disponibles en el no supervisado son más difusas ya que hay menos trabajos en la literatura. El razonamiento es similar al empleado previamente en la elección de los algoritmos. Para empezar, se escoge el algoritmo más destacado de esta rama, en este caso, *k-Means*. A continuación, como opción menos popular, se tiene AP. Para finalizar, se han escogido GM y BGM, que, al igual que GP, se basan en Gaussianidad, como algoritmos que de antemano no ofrecerían resultados óptimos pero vale la pena investigar si esto realmente es así o, en cambio, se pueden obtener también unas buenas predicciones.

El objetivo de partida, que es la estimación del estado cognitivo, se efectúa a través de la predicción del valor PERCLOS (regresión) o a qué rango del mismo pertenece (clasificación). Por contrapartida, los modelos no supervisados deberán realizar una división de las *features* en base a patrones comunes que identifiquen. Es importante señalar que para entrenar cada modelo se ha realizado previamente un ajuste de hiperparámetros junto con validación cruzada. Por una parte, la validación cruzada permite particionar los datos de tal manera que una parte es dedicada al entrenamiento del modelo y la otra a su posterior evaluación [80]; mientras que el ajuste de hiperparámetros posibilita encontrar la mejor configuración de los distintos parámetros de un modelo, consiguiendo el mejor rendimiento posible [81]. Tanto los modelos como la herramienta de configuración de sus parámetros, *RandomizedSearchCV*, son provistos para su uso por la librería *sklearn*.

5.3.1 Discretización de los valores PERCLOS

Si el problema se aborda con algoritmos de clasificación, los valores continuos de PERCLOS se deben discretizar. Para ello, se ha optado por su mapeo a tres clases de somnolencia ya que así lo recomienda la literatura [28, 82]. Para establecer estas clases es necesario definir dos umbrales que las separen. En la Tabla 5.2 se ilustra cómo se mapearía un valor continuo de PERCLOS a una clase en particular.

Para establecer de forma correcta los umbrales de PERCLOS que dividen las clases, se ha realizado un estudio de diferentes experimentos que han trabajado con dicho etiquetado. En gran parte de ellos [83, 84, 85, 86, 87, 88], se establecen de forma fija los

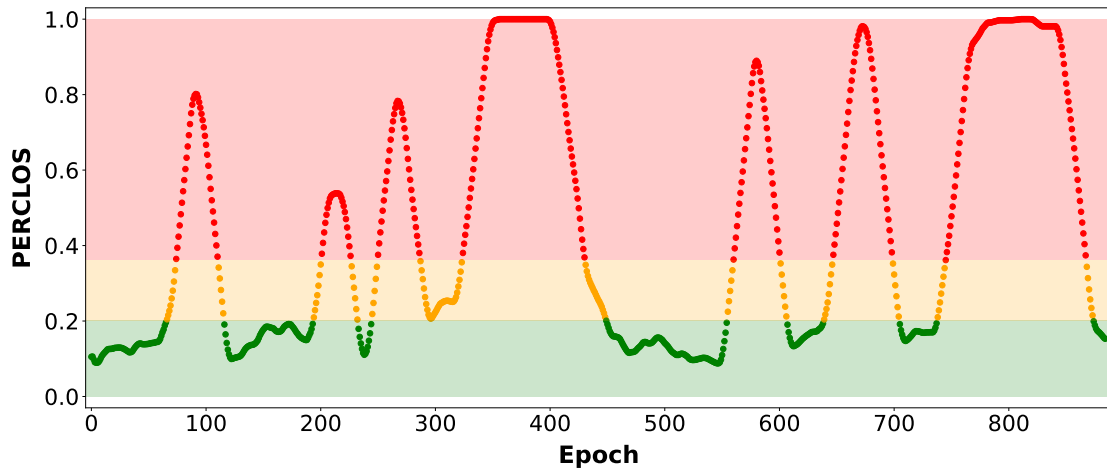
| Clase | Rango de valores |
|---------------------------|--------------------------------------------|
| Alerta o somnolencia leve | $[0, \text{lim_leve}]$ |
| Somnolencia moderada | $(\text{lim_leve}, \text{lim_moderada}]$ |
| Somolencia severa | $(\text{lim_moderada}, 1]$ |

Tabla 5.2: Umbrales de PERCLOS establecidos en este estudio.

valores de lim_leve y lim_moderada para todos los sujetos en base a otros estudios tomados de referencia, y a qué se ajusta mejor a la hora de obtener sus resultados. Por el contrario, en [89] se advierte que estos umbrales están relacionados con los diferentes métodos de detección utilizados por cada investigador, por lo que no es posible considerar directamente los resultados de otro, concluyendo que los umbrales de PERCLOS deben obtenerse a partir de los experimentos. Una vez contemplados ambos puntos de vista y, tras probar ambas alternativas, se ha optado por hacer un cálculo dinámico de los umbrales entre clases para cada sujeto. De esta forma, se logra tener en cuenta las peculiaridades de cada sujeto a nivel fisiológico y, por lo tanto, se obtiene una mejor división por clases de sus valores PERCLOS. El procedimiento seguido para obtener lim_leve es el siguiente: se calcula el rango de valores que comprende el PERCLOS de dicho sujeto y se toma el 12.5% de dicho valor. Mientras que para lim_moderada el cálculo es similar, tomándose en esta ocasión el 30% del valor calculado. Ambas fórmulas se pueden visualizar a continuación.

$$\text{lim_leve} = \min(\text{PERCLOS}) + (\max(\text{PERCLOS}) - \min(\text{PERCLOS})) * 0.125$$

$$\text{lim_moderada} = \min(\text{PERCLOS}) + (\max(\text{PERCLOS}) - \min(\text{PERCLOS})) * 0.30$$

**Figura 5.6:** Discretización del PERCLOS en tres clases.

Como se puede apreciar en Figura 5.6, los umbrales se adaptan de forma correcta

al sujeto, obteniendo así una buena división de clases de somnolencia en base a su valor PERCLOS. La zona marcada en **verde** corresponde a los valores del estado de alerta o somnolencia leve, la zona **naranja**, la descrita para la somnolencia moderada y por último la zona **roja**, que indica somnolencia severa. También se puede apreciar que la amplitud de los intervalos establecidos para cada clase será específica para cada sujeto.

6 Análisis de resultados

Una vez expuesto el diseño de la solución junto con las técnicas que van a ser empleadas, se procede a la implementación de los distintos algoritmos de aprendizaje que compondrán el framework. En particular, se hará una implementación de modelos específicos para cada sujeto y, posteriormente, se implementarán varios modelos generales en base a los mejores resultados obtenidos en los individuales. Para la implementación, se ha configurado la validación cruzada de manera que la proporción de muestras que son reservadas para la posterior evaluación de los modelos sea del 25% del total de las muestras disponibles. Es importante mencionar que la evaluación de los algoritmos se hará siguiendo las métricas expuestas en la Sección 2.2.3.

6.1 Modelos individuales por usuario

Los resultados que se muestran en las diferentes tablas de esta sección se calculan realizando un promedio de los resultados obtenidos por cada modelo en cada uno de los sujetos. Para cada tabla, la fila correspondiente a la mejor combinación queda resaltada en negrita. Por este motivo, los resultados tendrán la forma de $\text{Media} \pm \text{Desviación_Típica}$.

6.1.1 Algoritmos de regresión

Los resultados se muestran en la Tabla 6.1. En general, se observa que el menor RMSE se da en la mayoría de los algoritmos cuando se emplean únicamente los vectores de características referentes al PSD, seguidos muy de cerca de los que emplean PSD junto con el EOG y, por último, los que sólo emplean los datos del EEG. Es interesante destacar el hecho de que, en modelos individuales para cada sujeto, la adición del número de parpadeos en cada intervalo ha hecho que, por lo general, los resultados empeoren respecto a no usar esa *feature*. También cabe señalar que, aunque con el EEG se obtengan los peores resultados en todos los casos, los resultados obtenidos son aceptables a la hora de realizar una predicción sobre el estado de somnolencia más o menos precisa.

Como era de esperar, GP ofrece los peores resultados en cuanto a RMSE para los tres conjuntos de características. Esto se debe a que el algoritmo se basa en la teoría probabilística de la distribución Gaussiana, tal y como se comentó anteriormente. No obstante, incluso siendo el peor, los resultados obtenidos demuestran que es posible

realizar una predicción aceptable incluso si los datos analizados no tienen un comportamiento Gaussiano. Por otra parte, SVM y DT ofrecen unos resultados similares en cuanto al error cometido, mejorando los resultados de GP. Estos resultados son más destacables, sobre todo respecto a SVM, ya que es el método más extendido a la hora de trabajar con la somnolencia al volante y las BCIs. Finalmente, kNN y RF son los algoritmos con los que se ha obtenido la menor tasa de error. La combinación que mejores resultados ofrece, tanto en menor error como en mayor R^2 es RF con PSD más EOG como *features* (resaltado en negrita en la Tabla 6.1). Otras combinaciones que ofrecen también muy buenos resultados son RF haciendo uso únicamente del PSD y kNN con PSD.

| Modelo | Features | Número features | RMSE | R^2 |
|---------------------|----------------|-----------------|-------------------------------------|-------------------------------------|
| SVM-r | EEG | 136 | 0.163 ± 0.049 | 0.414 ± 0.216 |
| | PSD | 5 | 0.116 ± 0.040 | 0.673 ± 0.201 |
| | PSD+EOG | 6 | 0.121 ± 0.040 | 0.645 ± 0.207 |
| k-Nearest Neighbors | EEG | 136 | 0.152 ± 0.048 | 0.488 ± 0.203 |
| | PSD | 5 | 0.086 ± 0.030 | 0.822 ± 0.092 |
| | PSD+EOG | 6 | 0.100 ± 0.035 | 0.745 ± 0.179 |
| Decision Tree | EEG | 136 | 0.206 ± 0.064 | 0.055 ± 0.391 |
| | PSD | 5 | 0.114 ± 0.041 | 0.676 ± 0.191 |
| | PSD+EOG | 6 | 0.115 ± 0.034 | 0.680 ± 0.168 |
| Gaussian Processes | EEG | 136 | 0.210 ± 0.064 | -0.069 ± 0.691 |
| | PSD | 5 | 0.134 ± 0.040 | 0.554 ± 0.264 |
| | PSD+EOG | 6 | 0.167 ± 0.050 | 0.255 ± 0.536 |
| Random Forest | EEG | 136 | 0.142 ± 0.046 | 0.559 ± 0.171 |
| | PSD | 5 | 0.085 ± 0.027 | 0.828 ± 0.087 |
| | PSD+EOG | 6 | 0.084 ± 0.024 | 0.829 ± 0.092 |

Tabla 6.1: Comparativa de los algoritmos regresivos probados.

Para ilustrar estos resultados, en la Figura 6.1 se muestra un predictor RF que hace uso de las *features* del PSD más el EOG. La línea negra hace referencia a los valores reales del PERCLOS que se midieron durante la experimentación mientras que los rojos son los predichos por el modelo. Se aprecia que, en general, los valores predichos se aproximan mucho a los reales, reaccionando de forma adecuada a las diferentes subidas y bajadas, que se corresponderían con los momentos de somnolencia que se desean detectar en este trabajo. Es posible comprobar el *peso* que se le da a cada característica en RF, por lo que se presenta a continuación: número de parpadeos 32%, PSD del ritmo alfa 25%, PSD theta 15%, PSD gamma 10%, PSD delta 9% y PSD beta 8%. En base a ello se aprecia que lo más relevante para el algoritmo es el número de parpadeos y la banda alfa, siendo ambos factores directamente relacionados con la

somnolencia (ver Sección 2.1.3).

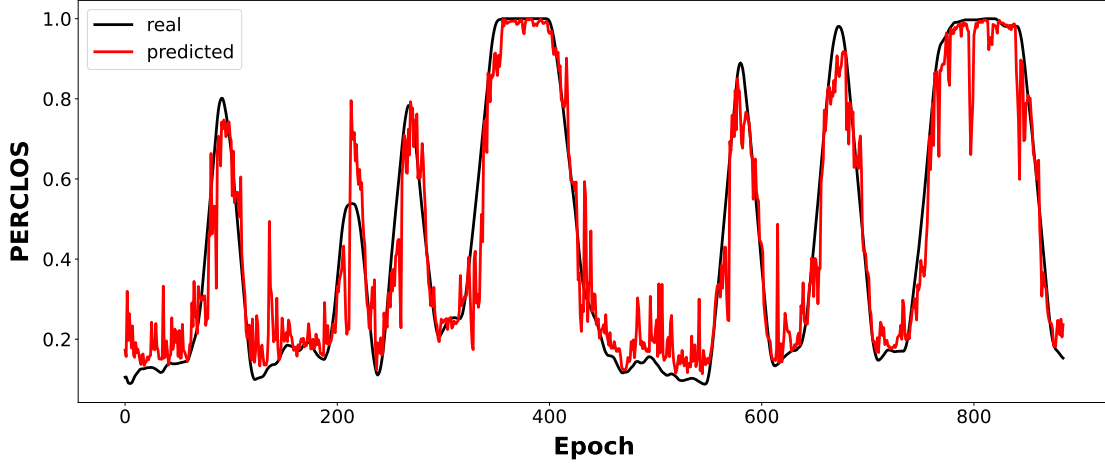


Figura 6.1: Predicción del PERCLOS con Random Forest haciendo uso del PSD y el EOG.

6.1.2 Algoritmos de clasificación

Todos los resultados se pueden visualizar en la Tabla 6.2. En este caso, si tenemos en cuenta los resultados obtenidos previamente, podemos suponer que los correspondientes a esta sección serán parecidos, ya que se emplean los mismos modelos pero en su versión de clasificación. No obstante, es de interés comentar las mejores combinaciones obtenidas ya que las métricas empleadas son distintas y arrojan una serie de valoraciones que no pueden ser estudiadas desde el punto de vista de la regresión.

En términos generales, la tasa de acierto toma valores entre el 67% y el 85%. Por su parte, los valores de *precision* suelen ir de la mano con los de *accuracy* y, tanto *recall* como *f1-score* se dan entre 0.45 y el 0.77 en el mejor de los casos.

De manera particular, en esta ocasión GP también ha producido los peores resultados, aunque DT con los datos EEG ha conseguido una tasa del 67.2%, que es más baja incluso que las de GP, las cuales giran en torno al 70%. Por su parte, los algoritmos de SVM y DT, al igual que sucedía con sus análogos en regresión ofrecen unos resultados mejores que GP pero sin llegar a ser los mejores, siendo la mejor combinación de estos dos DT con el PSD. No es de extrañar que las combinaciones de algoritmos y *features* que ofrecían los mejores resultados previamente, vuelvan a comportarse de manera similar en esta ocasión. De hecho, las tres mejores combinaciones vuelven a ser las mismas, aunque en este caso, en distinto orden. Con el fin de interpretar todas las métricas, se toma como ejemplo la mejor combinación que se tiene, esta es, RF con PSD, con una tasa de acierto del $85.7\% \pm 6\%$. En cuanto a sus otras métricas, su *precision* es del 0.833 ± 0.058 , lo que indica que el modelo producirá muy pocos falsos positivos, mientras que su *recall* de 0.757 ± 0.078 hace referencia a lo mismo que el

anterior pero en cuanto a falsos negativos. En este caso, si que se producirán alguno falsos negativos más que falsos positivos. Por último, el valor de $f1\text{-score}$ 0.776 ± 0.071 refleja qué tan fiable será tanto para falsos positivos como para falsos negativos, siendo ‘uno’ el mejor valor posible.

| Modelo | Features | Número features | Accuracy | Precision | Recall | f1-score |
|---------------------|------------|-----------------|------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| SVM-c | EEG | 136 | $71.9\% \pm 10\%$ | 0.708 ± 0.156 | 0.514 ± 0.120 | 0.502 ± 0.131 |
| | PSD | 5 | $76.4\% \pm 10\%$ | 0.780 ± 0.120 | 0.579 ± 0.130 | 0.575 ± 0.139 |
| | PSD+EOG | 6 | $76.1\% \pm 9\%$ | 0.0744 ± 0.118 | 0.578 ± 0.123 | 0.577 ± 0.131 |
| k-Nearest Neighbors | EEG | 136 | $71.3\% \pm 9\%$ | 0.610 ± 0.127 | 0.515 ± 0.101 | 0.510 ± 0.105 |
| | PSD | 5 | $85.4\% \pm 5\%$ | 0.814 ± 0.049 | 0.766 ± 0.078 | 0.775 ± 0.067 |
| | PSD+EOG | 6 | $78.5\% \pm 8\%$ | 0.730 ± 0.120 | 0.630 ± 0.123 | 0.635 ± 0.120 |
| Decision Tree | EEG | 136 | $67.2\% \pm 11\%$ | 0.627 ± 0.137 | 0.474 ± 0.090 | 0.464 ± 0.101 |
| | PSD | 5 | $80.4\% \pm 8\%$ | 0.726 ± 0.090 | 0.711 ± 0.084 | 0.712 ± 0.088 |
| | PSD+EOG | 6 | $79.8\% \pm 9\%$ | 0.743 ± 0.091 | 0.716 ± 0.073 | 0.715 ± 0.079 |
| Gaussian Processes | EEG | 136 | $70.5\% \pm 10\%$ | 0.703 ± 0.127 | 0.482 ± 0.094 | 0.462 ± 0.092 |
| | PSD | 5 | $69.8\% \pm 11\%$ | 0.790 ± 0.165 | 0.458 ± 0.116 | 0.423 ± 0.125 |
| | PSD+EOG | 6 | $70.7\% \pm 10\%$ | 0.770 ± 0.121 | 0.479 ± 0.108 | 0.453 ± 0.114 |
| Random Forest | EEG | 136 | $73.8\% \pm 9\%$ | 0.740 ± 0.115 | 0.542 ± 0.095 | 0.541 ± 0.098 |
| | PSD | 5 | $85.7\% \pm 6\%$ | 0.833 ± 0.058 | 0.757 ± 0.078 | 0.776 ± 0.071 |
| | PSD+EOG | 6 | $84.6\% \pm 6\%$ | 0.829 ± 0.050 | 0.736 ± 0.080 | 0.756 ± 0.071 |

Tabla 6.2: Comparativa de los algoritmos de clasificación probados.

En la Figura 6.2 se ilustran las predicciones correspondientes a la mejor combinación de algoritmo y *features* para un sujeto del experimento. Los colores, tanto del fondo como de los puntos, denotan cuál sería la clase correspondiente a dicho nivel de PER-CLOS (ver Sección 5.3.1). Una predicción correcta se identifica fácilmente, ya que el color de la franja y del punto serán el mismo; mientras que si son distintos, la predicción será incorrecta. Un ejemplo de esto sería un punto verde sobre franja roja, reflejando una predicción de *leve* ante un valor discretizado como *severa*. Se puede apreciar qué tan buenas son las predicciones realizadas al comparar estos resultados con los de la propia discretización realizada en la Figura 5.6.

Otra manera de analizar los resultados obtenidos para el modelo anterior de RF con PSD, es haciendo uso de una matriz de confusión. La matriz de confusión correspondiente a las predicciones anteriores se muestra en la Figura 6.3, pudiendo analizar de un rápido vistazo las predicciones correctas e incorrectas. Los elementos diagonales representan el número de puntos para los cuales la predicción es correcta, mientras que los elementos fuera de la diagonal son aquellos que están mal etiquetados por el clasificador. Cuanto más altos sean los valores diagonales de la matriz de confusión, mayor número de predicciones correctas.

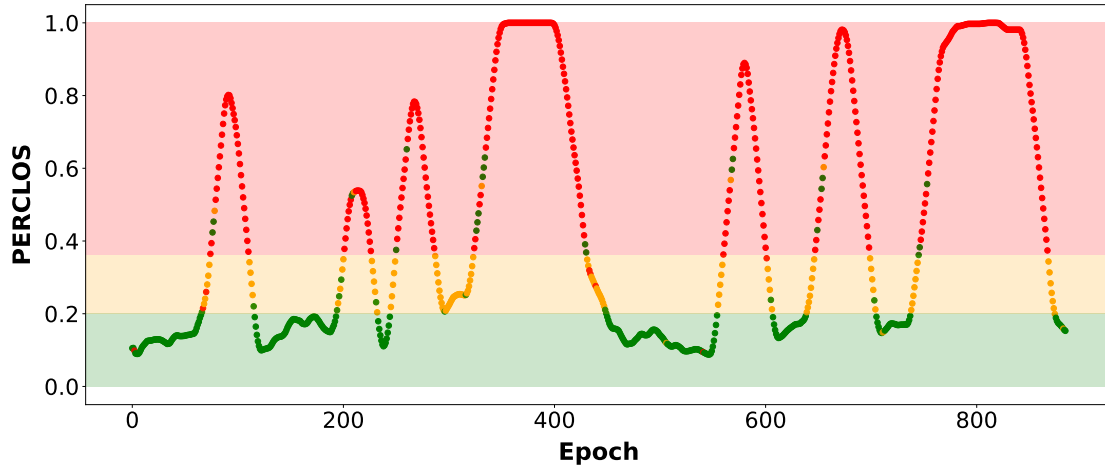


Figura 6.2: Predicción de las clases PERCLOS.

6.1.3 Algoritmos de agrupamiento

En este caso, el problema pertenece a la categoría de aprendizaje no supervisado, por lo que no se precisa del etiquetado de los datos. Dados los vectores de características, el algoritmo indicará a qué grupo (*clúster*) pertenece cada uno, en base a patrones que él mismo ha detectado dentro de esos datos. Los resultados obtenidos se muestran en la Tabla 6.3. Al evaluar estos modelos se persigue que su SC sea lo más cercano a uno a la vez que el valor del DBI lo sea a cero.

De manera general, los mejores resultados son obtenidos con el PSD, siendo *k-Means* y AP los algoritmos con mejores prestaciones en ambas métricas. Por su parte, tanto GM como BGM ofrecen unos resultados peores, debido a que estos algoritmos suponen que todos los puntos de datos se generan a partir de una mezcla de distribuciones Gaussianas [90] y, al igual que pasaba con GP, el comportamiento de las señales EEG es no Gaussiano, resultando en unas predicciones peores en comparación a los otros algoritmos tal y como se había teorizado durante la selección de estos tres algoritmos. La combinación que mejores resultados ofrece es AP con únicamente el PSD, donde los valores obtenidos son de 0.613 ± 0.210 para el SC, lo que indica que los clústers están bastante bien definidos; y de 0.593 ± 0.322 en DBI que, por su parte, indica que se tiene una separación aceptable entre los clústers.

A pesar de que los resultados obtenidos para AP con PSD han sido aceptables, el algoritmo produce agrupaciones de los datos que no son capaces de diferenciar entre estados de somnolencia. Esto se ve reflejado en la Figura 6.4. Las líneas roja y naranja indican los umbrales entre niveles de somnolencia para ese sujeto y, los puntos, hacen referencia a la agrupación por parte del algoritmo, correspondiendo cada color con un clúster. Como se observa en la imagen, las agrupaciones no guardan relación con las subidas y bajadas del PERCLOS.

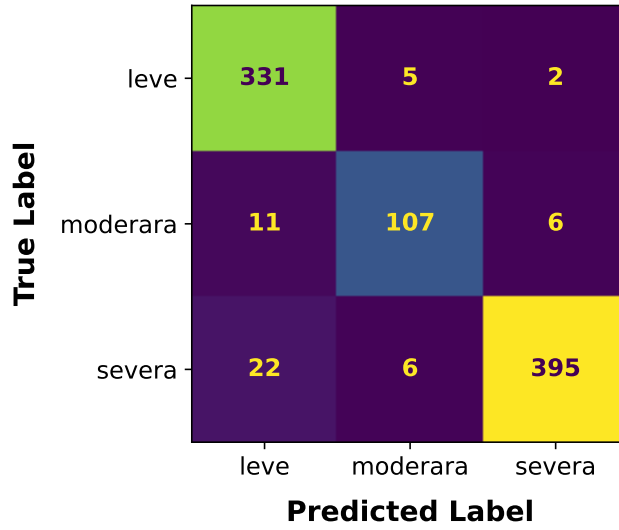


Figura 6.3: Matriz de confusión relativa a las predicciones del ejemplo con RF y PSD.

6.2 Modelo general para todos los usuarios

Un vez se tienen los resultados de los modelos específicos, se toman los dos algoritmos que mejor resultados han obtenido en la fase de modelos individuales (kNN y RF) junto con su respectivo conjunto de *features*. Además, se incluye también el algoritmo SVM debido a su gran presencia en la literatura. Estos tres algoritmos se usan para evaluar un único modelo general, el cual contiene toda la información de los usuarios existentes en el *dataset*. Debido a ello, estos modelos servirán para hacer predicciones sobre cualquiera de los sujetos. En este caso, al evaluar el modelo solo se obtendrá un único valor por cada métrica, por lo que los resultados se presentan directamente sin necesidad de calcular el promedio ni la desviación típica. Por último, cabe destacar que se ha descartado el estudio de modelos generales con *clustering* ya que los resultados obtenidos en los modelos individuales no han sido capaces de diferenciar clases que puedan derivar en una diferenciación entre estados de somnolencia.

De las tres combinaciones probadas para los modelos generales regresivos, cuyos resultados se ofrecen en la Tabla 6.4, tanto kNN con PSD como RF con PSD más EOG presentan un RMSE y un R^2 bastante buenos. Por otro lado, SVM con PSD ofrece unos resultados inferiores en comparación a las otras opciones. Esto sucede de igual forma con este algoritmo en los modelos individuales.

Por parte de los modelos generales de clasificación y, en la misma sintonía que los modelos generales regresivos, las mejores opciones vuelven a ser kNN y RF, en este caso, ambas combinaciones empleando el PSD. De la misma manera, SVM ha sido el que peor rendimiento ha ofrecido. Todos estos datos se muestran en la Tabla 6.5.

| Modelo | Features | Número features | Silhouette Coefficient | Davies-Bouldin Index |
|---------------------------|------------|-----------------|-------------------------------------|-------------------------------------|
| k-Means | EEG | 136 | 0.261 ± 0.094 | 1.580 ± 0.277 |
| | PSD | 5 | 0.573 ± 0.156 | 0.711 ± 0.195 |
| | PSD+EOG | 6 | 0.447 ± 0.123 | 0.781 ± 0.217 |
| Gaussian Mixture | EEG | 136 | 0.126 ± 0.141 | 4.848 ± 3.635 |
| | PSD | 5 | 0.397 ± 0.200 | 1.367 ± 0.344 |
| | PSD+EOG | 6 | 0.322 ± 0.137 | 1.703 ± 0.364 |
| Bayesian Gaussian Mixture | EEG | 136 | 0.320 ± 0.127 | 2.044 ± 0.761 |
| | PSD | 5 | 0.439 ± 0.191 | 1.311 ± 0.450 |
| | PSD+EOG | 6 | 0.384 ± 0.143 | 1.586 ± 0.438 |
| Affinity Propagation | EEG | 136 | 0.096 ± 0.054 | 1.615 ± 0.345 |
| | PSD | 5 | 0.613 ± 0.210 | 0.593 ± 0.322 |
| | PSD+EOG | 6 | 0.400 ± 0.144 | 0.800 ± 0.228 |

Tabla 6.3: Comparativa de los algoritmos de agrupamiento probados.

| Modelo | Features | Número features | RMSE | R^2 |
|----------------------|----------------|-----------------|--------------|--------------|
| SMV-r | PSD | 5 | 0.206 | 0.428 |
| k-Nearest Neighbors | PSD | 5 | 0.137 | 0.748 |
| Random Forest | PSD+EOG | 6 | 0.122 | 0.798 |

Tabla 6.4: Comparativa de los algoritmos regresivos probados para el modelo general.

6.3 Otros modelos estudiados

La ultima alternativa que se ha explorado es el entrenamiento de un modelo general a partir de un subconjunto de los sujetos del experimento que, posteriormente, sea probado y evaluado con los sujetos que no formaron parte del entrenamiento. Se plantean dos versiones de este enfoque. El primero presenta una versión en la que el 90% de los sujetos serán usados para el entrenamiento y el 10% restante para la evaluación, nombrando a este modelo como *Modelo 90-10*. En el segundo enfoque, se empleará únicamente el 70% de los sujetos para el experimento. Siguiendo en este razonamiento, se le denominará *Modelo 70-30*. Las combinaciones de algoritmos y conjuntos de características empleadas para su implementación serán los mismos que en la sección anterior. Dado que en este caso los resultados son medidos para los sujetos ajenos a la fase de entrenamiento, es necesario calcular de forma particular las métricas correspondientes al tipo de modelo y expresar los resultados con el formato *Media \pm Desviación_Típica*.

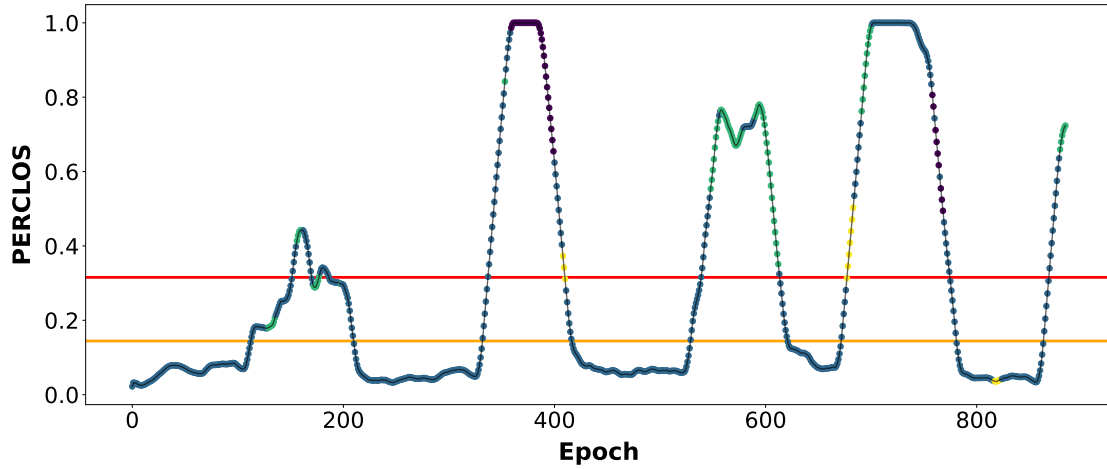


Figura 6.4: Agrupamiento de los diferentes vectores de características realizado por AP con PSD.

| Modelo | Features | Número features | Accuracy | Precision | Recall | f1-score |
|----------------------|------------|-----------------|--------------|--------------|--------------|--------------|
| SVM-c | PSD | 5 | 63.1% | 0.585 | 0.510 | 0.519 |
| k-Nearest Neighbors | PSD | 5 | 79.4% | 0.758 | 0.761 | 0.760 |
| Random Forest | PSD | 5 | 82.5% | 0.799 | 0.783 | 0.790 |

Tabla 6.5: Comparativa de los algoritmos de clasificación probados para el modelo general.

6.3.1 Modelo 90-10

Dado que se tiene un total de 23 experimentos en el conjunto de datos (ver Sección 5.1), se reservarán dos sujetos, correspondiente al $\sim 10\%$ del total, para la evaluación del modelo, que será entrenado con los datos de los 21 sujetos restantes.

Los resultados proporcionados haciendo uso de regresión siguen la misma tendencia que en el modelo general (Sección 6.2). SVM presenta las peores métricas de las tres combinaciones estudiadas. Tras ello, y en orden de mejoría, se tiene a kNN y, por último, a RF, siendo este algoritmo junto a los datos PSD y EOG la combinación que mejores resultados ofrece. La Tabla 6.6 puede ser consultada para ver esta información con más detalle.

Respecto a los modelos 90-10 de clasificación (ver Tabla 6.7), es importante destacar que kNN con PSD como conjunto de *features* supera esta vez a RF con PSD, no en tasa de acierto, siendo ambos prácticamente idénticos, pero sí en el resto de métricas (*precision*, *recall* y *f1-score*). Por su parte, SVM junto a PSD ofrece unos resultados casi similares a estas dos últimas combinaciones mencionadas pero siempre un peldaño

| Modelo | Features | Número features | RMSE | R^2 |
|----------------------|----------------|-----------------|-------------------------------------|-------------------------------------|
| SMV-r | PSD | 5 | 0.258 ± 0.056 | -1.908 ± 0.996 |
| k-Nearest Neighbors | PSD | 5 | 0.203 ± 0.080 | -0.655 ± 0.038 |
| Random Forest | PSD+EOG | 6 | 0.158 ± 0.074 | 0.033 ± 0.184 |

Tabla 6.6: Comparativa de los algoritmos regresivos probados para el modelo 90-10.

por debajo.

| Modelo | Features | Número features | Accuracy | Precision | Recall | f1-score |
|----------------------------|------------|-----------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| SVM-c | PSD | 5 | $57.2\% \pm 23\%$ | 0.374 ± 0.144 | 0.407 ± 0.181 | 0.382 ± 0.150 |
| k-Nearest Neighbors | PSD | 5 | $59.7\% \pm 17\%$ | 0.455 ± 0.124 | 0.484 ± 0.118 | 0.458 ± 0.120 |
| Random Forest | PSD | 5 | $61.2\% \pm 19\%$ | 0.438 ± 0.181 | 0.452 ± 0.186 | 0.433 ± 0.167 |

Tabla 6.7: Comparativa de los algoritmos de clasificación probados para el modelo 90-10.

6.3.2 Modelo 70-30

De manera análoga al razonamiento de la sección previa, en este caso, se destinan 16 sujetos al entrenamiento del modelo y se reservan los siete restantes para su evaluación.

En ambos modelos (clasificación y regresión), SVM siempre ha sido la peor de las tres combinaciones. En regresión, RF con PSD y EOG se sigue manteniendo como la mejor alternativa, seguida de kNN con PSD. Mientras que para clasificación, tanto kNN como RF (ambos con PSD) son alternativas a considerar, ya que RF ofrece una mejor tasa de acierto en comparación a kNN pero este último le supera ligeramente en el resto de métricas. Es interesante mencionar que la tasa de acierto media ha caído en los tres casos por debajo del 50%, lo que hace que este conjunto de modelos no sea tan interesante como otros presentados anteriormente. La Tabla 6.8 presenta los resultados de regresión anteriormente comentados mientras que los respectivos a clasificación se muestran en Tabla 6.9.

6.4 Comparativa entre métodos

Esta sección está dedicada a comparar cómo de bueno es usar modelos individuales frente a generales, realizando una comparativa independiente para regresión y para

| Modelo | Features | Número features | RMSE | R^2 |
|----------------------|----------------|-----------------|-------------------------------------|--------------------------------------|
| SMV-r | PSD | 5 | 0.256 ± 0.065 | -1.897 ± 3.024 |
| k-Nearest Neighbors | PSD | 5 | 0.220 ± 0.047 | -0.697 ± 0.648 |
| Random Forest | PSD+EOG | 6 | 0.183 ± 0.046 | -0.166 ± 0.454 |

Tabla 6.8: Comparativa de los algoritmos regresivos probados para el modelo 70-30.

| Modelo | Features | Número features | Accuracy | Precision | Recall | f1-score |
|----------------------|----------|-----------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| SVM-c | PSD | 5 | $39.6\% \pm 20\%$ | 0.379 ± 0.072 | 0.406 ± 0.085 | 0.289 ± 0.136 |
| k-Nearest Neighbors | PSD | 5 | $44.8\% \pm 16\%$ | 0.436 ± 0.074 | 0.442 ± 0.113 | 0.368 ± 0.109 |
| Random Forest | PSD | 5 | $45.9\% \pm 15\%$ | 0.405 ± 0.059 | 0.411 ± 0.069 | 0.354 ± 0.075 |

Tabla 6.9: Comparativa de los algoritmos de clasificación probados para el modelo 70-30.

clasificación. Como en el resto de modelos generales, *clustering* ha sido excluido al no presentar resultados significativos en la estimación de la somnolencia

6.4.1 Sobre regresión

Dentro del abanico de algoritmos que se han estudiado, los modelos individuales han sido siempre superiores a los modelos generales con el inconveniente de que sólo son capaces de hacer una correcta estimación para el usuario con el que han sido entrenados. En particular, las mejores combinaciones de modelos generales (RF con PSD y EOG, y kNN con PSD) presentan unas métricas un poco por debajo que sus análogas en los modelos individuales. Concretamente, se ha producido un aumento de RMSE del 0.04 y un descenso del 0.035 en R^2 . En base a las métricas de estos modelos generales es posible apreciar que las predicciones que se realicen serán bastante buenas aunque, si los comparamos con las mejores combinaciones de los modelos individuales (ver Figura 6.1), serán algo peores. No obstante, ambas opciones se pueden considerar perfectamente de cara a evaluar la somnolencia de un conductor. Por su parte, los modelos 90-10 y 70-30 ofrecen unos resultados peores que el modelo general, cosa que cabría esperar. Sin embargo, el modelo 90-10 ha conseguido hacer unas predicciones aceptables, teniendo en cuenta que los sujetos con los que han sido evaluados son totalmente desconocidos para este modelo.

6.4.2 Sobre clasificación

De manera más o menos similar a como ha sucedido en la comparación previa de los modelos regresivos, los modelos individuales han tenido un mejor rendimiento, aunque con algunas excepciones. En este caso, la tasa de acierto media para la mejor combinación de modelo general (RF con PSD) bajó tan solo un 3.2% respecto al valor medio medido en la mejor combinación de los modelos individuales, que también es RF con PSD. Por su parte, las métricas de *recall* y *f1* se **incrementaron** respecto a la versión individual ligeramente, lo que indica que este modelo general produce menos falsos positivos y negativos, aunque también hay que recordar que tiene una menor tasa de acierto.

Se concluye que tanto el modelo general como los individuales son válidos para realizar predicciones y que estas son de gran utilidad. Para terminar, las alternativas 90-10 y 70-30 sí que ven reducidas sus métricas respecto a los otros modelos. Esto es debido a que no se consiguen adaptar bien a todos los nuevos sujetos con los que son sometidas a evaluación (al igual que sucede para regresión). Hay ciertos sujetos con los que estos modelos sí que realizan una predicción aceptable, mientras que se tienen otros en los que el modelo no consigue adaptarse correctamente. También cabe mencionar que los resultados obtenidos de los modelos 90-10 y 70-30 son preliminares y requieren de más evaluación.

7 Conclusiones y vías futuras

Las Interfaces Cerebro-Máquina (BCIs) son sistemas que han contribuido en el desarrollo del bienestar de las personas en gran variedad de ámbitos como la salud, la prevención de riesgos e incluso el ocio. Además, destaca su apoyo en el desarrollo de la neurociencia cognitiva. Todo esto gracias a que brindan la posibilidad de realizar estudios de las ondas cerebrales a muchos equipos de investigación sin la necesidad de una infraestructura como de la que se dispone en grandes centros. Una de las ramas de estudio que surgen gracias a las BCIs es la estimación del nivel de somnolencia que presenta un sujeto mientras conduce.

Para resolver el problema de identificar el nivel de somnolencia de un conductor se ha revisado el estado del arte. Posteriormente, se ha diseñado un framework inteligente capaz de predecir el estado de somnolencia de un conductor. Para ello, se han explorado algoritmos de ML y, concretamente, sistemas de regresión, clasificación y *clustering* que hacen uso de ciertas características presentes en la señal para efectuar las predicciones referentes al PERCLOS del sujeto, siempre y cuando se emplee aprendizaje supervisado. El repertorio de algoritmos y características se han extraído de la literatura en base a su grado de popularidad y los resultados que producen. Se ha estudiado la eficacia de modelos individuales para cada usuario además de un modelo general incluyendo los datos de todos los usuarios. Dentro de este último conjunto, también se han explorado dos alternativas en las que se reserva cierta parte de los sujetos para realizar la evaluación del algoritmo, denominados como 90-10 y 70-30.

A partir de la experimentación realizada, se ha observado que el algoritmo más popular en la literatura (SVM) se ha visto superado en este framework por otros no tan nombrados en la misma. En particular, los mejores resultados se obtuvieron con RF tanto para clasificación como regresión, no solo en los modelos individuales sino también en los generales. Por otra parte, los modelos individuales de *clustering* no han sido capaces de agrupar los datos de forma que se obtenga una diferenciación entre estados de somnolencia. En ambos tipos de aprendizaje, se observa que algoritmos basados en Gaussianidad son los que peores resultados obtienen, debido a la propiedades intrínsecas de la señal electroencefálica, pero aún así, consiguen hacer estimaciones válidas en los modelos individuales. En cuanto a las características estudiadas, los datos referentes a la Densidad Espectral de Potencia (PSD) ofrecen, en general, las mejores prestaciones de cara a la estimación de la somnolencia. Al margen de esto, usar tanto el EEG como la combinación PSD más EOG ha dado igualmente buenos resultados y serían válidos para la predicción de la somnolencia. Por último, respecto a los diferentes tipos de modelos entrenados, los modelos individuales ofrecen las mejores prestaciones,

con la limitación de ceñirse a un único sujeto. Estos son seguidos muy de cerca por los modelos generales, mientras que los modelos 90-10 y 70-30 realizaban unas predicciones aceptables para parte de los sujetos pero sin conseguir adaptarse de manera correcta a todos ellos, aunque se debe tener en cuenta que son unos resultados preliminares y se requiere de más evaluación.

Las vías futuras que se plantean en este trabajo consisten, en primer lugar, en generar un conjunto de datos haciendo uso de una BCI propia con el fin de cotejar los resultados obtenidos en este estudio con los de este nuevo *dataset*. El siguiente punto a contemplar sería aplicar algoritmos de DL para la estimación de la somnolencia, ya que cada vez son más populares en la literatura y pueden llegar a ofrecer mejores resultados. Es preciso señalar que se han realizado pruebas preliminares con estos sistemas, ofreciendo resultados prometedores. No obstante, debido a limitaciones temporales durante la realización del trabajo no se ha podido realizar una experimentación suficiente. En este sentido, estamos interesados en aplicar técnicas de DL tanto sobre el *dataset* usado en el trabajo, como sobre el que se desea generar usando una BCI. Finalmente, se contempla la posibilidad de añadir el trabajo realizado en el presente TFG como módulo de detección de la somnolencia al volante en el framework implementado en trabajos previos del equipo de investigación [16].

Bibliografía

- [1] Elena Martinez. El telencéfalo: anatomía y función, May 2021. URL <https://www.psicoactiva.com/blog/el-telencefalo/>.
- [2] Jaakko Malmivuo and Robert Plonsey. *Bioelectromagnetism*. 13. *Electroencephalography*, pages 247–264. 01 1995. ISBN 978-0195058239.
- [3] Explicación de la máquina de vectores de soporte (svm), Feb 2021. URL <https://ichi.pro/es/explicacion-de-la-maquina-de-vectores-de-soporte-svm-97743104690915>.
- [4] Leonardo José Gómez Figueroa. *Análisis de señales EEG para detección de eventos oculares, musculares y cognitivos*. PhD thesis, Industriales, 2016.
- [5] Edmundo Rosales Mayor and Jorge Rey De Castro Mujica. Somnolencia: Qué es, qué la causa y cómo se mide. *Acta médica peruana*, 27(2):137–143, 2010.
- [6] Dirección General de Tráfico (DGT). Otros factores de riesgo: El sueño. URL https://www.dgt.es/multimedia/educacion_vial/contenidos-profesores/El_sue%C3%B1o/cap-sueno/index.html?bid=14.
- [7] Sergio Amadoz. La somnolencia al volante (sobre todo en julio) causa los accidentes con mayor mortalidad, Jul 2017. URL <https://motor.elpais.com/actualidad/somnolencia-volante-accidentes-graves/>.
- [8] Cristina Castro. Quedarse dormido al volante duplica el riesgo de morir en un accidente de tráfico, Jul 2017. URL <https://www.elindependiente.com/vida-sana/2017/07/10/la-somnolencia-duplica-el-riesgo-de-morir-en-un-accidente-de-trafico/>.
- [9] Tecnología para no dormirse al volante, Apr 2021. URL <https://canalsalud.imq.es/blog/tecnologia-para-no-dormirse-al-volante>.
- [10] Diego Redolar Ripoll. *Neurociencia cognitiva*. QUITO/UIDE/2015, 2015.
- [11] Jamie Ward. *The student's guide to cognitive neuroscience*. Routledge, 2019.

-
- [12] Izzat A. Akbar, Arthur M. Rumagit, Mitaku Utsunomiya, Takamasa Morie, and Tomohiko Igasaki. Three drowsiness categories assessment by electroencephalogram in driving simulator environment. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, July 2017. doi: 10.1109/embc.2017.8037464.
- [13] Gordon M Shepherd. *Neurobiology*. Oxford University Press, 1988.
- [14] Neil R Carlson and David P Clark. *Fisiología de la conducta*. Pearson Educación Madrid, Spain:, 2014.
- [15] Citlali Helenes Gonzalez. Lóbulos cerebrales y sus funciones principales: Neuroméxico: Investigación, divulgación y difusión científica, Oct 2020. URL <https://www.neuromexico.org/2020/08/20/lobulos-cerebrales-y-sus-funciones-principales/>.
- [16] Enrique Tomás Martínez Beltrán. Framework para detección de ondas p300 y ciberataques basados en ruido en interfaces cerebro-máquina. Master's thesis, Universidad de Murcia, 2020.
- [17] Benny Mwengue Gimbada and Daniel Rodenstein. Evaluación de la somnolencia. *Archivos de Bronconeumología*, 45(7):349–351, July 2009. doi: 10.1016/j.arbres.2008.10.002.
- [18] David F. Dinges and John W. Powell. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers*, 17(6):652–655, November 1985. doi: 10.3758/bf03200977.
- [19] Azmeh Shahid, Kate Wilkinson, Shai Marcu, and Colin M. Shapiro. *Karolinska Sleepiness Scale (KSS)*, pages 209–210. Springer New York, New York, NY, 2012. ISBN 978-1-4419-9893-4. doi: 10.1007/978-1-4419-9893-4_47.
- [20] Thomas A. Dingus, H. Lenora Hardee, and Walter W. Wierwille. Development of models for on-board detection of driver impairment. *Accident Analysis & Prevention*, 19(4):271–283, August 1987. doi: 10.1016/0001-4575(87)90062-5.
- [21] United States. Federal Motor Carrier Safety Administration. Technology Division. Perclos: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance, Oct 1998.
- [22] Iván Daza, Luis Bergasa, Sebastián Bronte, J. Yebes, Javier Almazán, and Roberto Arroyo. Fusion of optimized indicators from advanced driver assistance systems (ADAS) for driver drowsiness detection. *Sensors*, 14(1):1106–1131, January 2014. doi: 10.3390/s140101106.
-

-
- [23] Udo Trutschel, Bill Sirois, David Sommer, Martin Golz, and Dave Edwards. PERCLOS: An alertness measure of the past. In *Proceedings of the 6th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design : driving assessment 2011*. University of Iowa, 2011. doi: 10.17077/drivingassessment.1394.
- [24] Lawrence Barr, Heidi Howarth, Stephen Popkin, and Robert J Carroll. A review and evaluation of emerging driver fatigue detection measures and technologies. *National Transportation Systems Center, Cambridge. US Department of Transportation, Washington*, 2005. URL <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.508.8409&rep=rep1&type=pdf>.
- [25] Control de fatiga y somnolencia de conductores para buses y camiones, Jan 2020. URL <https://www.movilflix.com/control-de-fatiga-y-somnolencia/>.
- [26] Lawrence Barr, Stephen Popkin, Heidi Howarth, et al. An evaluation of emerging driver fatigue detection measures and technologies. Technical report, United States. Department of Transportation. Federal Motor Carrier Safety Administration, 2009.
- [27] Noa Quevedo López. Estudio del parpadeo durante la conducción de vehículos (aspectos cognitivos y de flujo de información). Master's thesis, Universitat Politècnica de Catalunya, 2012.
- [28] Byung-Chan Chang, Jung-Eun Lim, Hae-Jin Kim, and Bo-Hyeok Seo. A study of classification of the level of sleepiness for the drowsy driving prevention. In *SICE Annual Conference 2007*. IEEE, September 2007. doi: 10.1109/sice.2007.4421521.
- [29] Luis Fernando Nicolas-Alonso and Jaime Gomez-Gil. Brain computer interfaces, a review. *Sensors*, 12(2), January 2012. doi: 10.3390/s120201211.
- [30] Sergio López Bernal, Alberto Huertas Celdrán, Gregorio Martínez Pérez, Michael Taynnan Barros, and Sasitharan Balasubramaniam. Security in brain-computer interfaces. *ACM Computing Surveys*, 54(1):2–3, April 2021. doi: 10.1145/3427376.
- [31] Mario Quiles Pérez. Estudio de ciberataques con estímulos visuales subliminales que pretenden obtener información confidencial usando interfaces cerebro-máquina. Master's thesis, Universidad de Murcia, 2020.
- [32] Pablo Ballarin Usieto and Javier Minguez. La importancia de la ciberseguridad en brain-computer interfaces, Feb 2019. URL <https://www.bitbrain.com/es/blog/ciberseguridad-cerebro-computadora>.
- [33] Electroencefalograma: qué es, síntomas, causas, prevención y tratamiento. URL <https://www.topdoctors.es/diccionario-medico/electroencefalograma>.
-

-
- [34] Difference between stationary and non-stationary signals (with table). URL <https://askanydifference.com/difference-between-stationary-and-non-stationary-signals/>.
- [35] Tugce Balli and Ramaswamy Palaniappan. Nonlinear approach to brain signal modeling. In *Encyclopedia of Information Science and Technology, Second Edition*, pages 2834–2839. IGI Global. doi: 10.4018/978-1-60566-026-4.ch453.
- [36] Dan Russell. Acoustics and vibration animations, Jul 1996. URL <https://www.acs.psu.edu/drussell/demos/superposition/superposition.html>.
- [37] A. Hyvärinen and E. Oja. Independent component analysis: algorithms and applications. *Neural Networks*, 13(4):411–430, 2000. ISSN 0893-6080. doi: 10.1016/S0893-6080(00)00026-5.
- [38] Shawhin Talebi. Independent component analysis (ica), Mar 2021. URL <https://towardsdatascience.com/independent-component-analysis-ica-a3eba0ccec35>.
- [39] Aurélien Géron. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, 2019.
- [40] Jose Martinez Heras. Máquinas de vectores de soporte (svm), May 2019. URL <https://www.iartificial.net/maquinas-de-vectores-de-soporte-svm>.
- [41] Jason Brownlee. What is deep learning?, Aug 2020. URL <https://machinelearningmastery.com/what-is-deep-learning/>.
- [42] Kevin Gurney. *An introduction to neural networks*. CRC press, 1997.
- [43] Metrics and scoring: quantifying the quality of predictions. URL https://scikit-learn.org/stable/modules/model_evaluation.html.
- [44] Clustering. URL <https://scikit-learn.org/stable/modules/clustering.html#clustering-performance-evaluation>.
- [45] Alberto Cianda, Nélida Melek, Nora Oliveri, and Horado Garcia. Electrooculograma de movimiento: técnica y metodología. *Arquivos Brasileiros de Oftalmologia*, 54(5), 1991. doi: 10.5935/0004-2749.19910016.
- [46] Zehong Cao, Michael Chuang, J.T. King, and Chin-Teng Lin. Multi-channel EEG recordings during a sustained-attention driving task (raw dataset), 2 2019.
- [47] Jianliang Min, Ping Wang, and Jianfeng Hu. The original eeg data for driver fatigue detection, Jul 2017.
-

-
- [48] Wei-Long Zheng and Bao-Liang Lu. A multimodal approach to estimating vigilance using eeg and forehead eeg, 2017. URL <http://stacks.iop.org/1741-2552/14/i=2/a=026017>.
- [49] Mastaneh Torkamani-Azar, Sumeyra Demir Kanik, Serap Aydin, and Mujdat Cetin. Prediction of reaction time and vigilance variability from spatio-spectral features of resting-state eeg in a long sustained attention task, 2020.
- [50] G. Borghini, G. Vecchiato, J. Toppi, L. Astolfi, A. Maglione, R. Isabella, C. Caltagirone, W. Kong, D. Wei, Z. Zhou, L. Polidori, S. Vitiello, and F. Babiloni. Assessment of mental fatigue during car driving by using high resolution eeg activity and neurophysiologic indices. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6442–6445, 2012. doi: 10.1109/EMBC.2012.6347469.
- [51] Mervyn V.M. Yeo, Xiaoping Li, Kaiquan Shen, and Einar P.V. Wilder-Smith. Can svm be used for automatic eeg detection of drowsiness during car driving? *Safety Science*, 47(1):115–124, 2009. ISSN 0925-7535. doi: 10.1016/j.ssci.2008.01.007.
- [52] Michael Simon, Eike A. Schmidt, Wilhelm E. Kincses, Martin Fritzsche, Andreas Bruns, Claus Aufmuth, Martin Bogdan, Wolfgang Rosenstiel, and Michael Schrauf. EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions. *Clinical Neurophysiology*, 122(6):1168–1178, June 2011. doi: 10.1016/j.clinph.2010.10.044.
- [53] Eike A. Schmidt, Michael Schrauf, Michael Simon, Martin Fritzsche, Axel Buchner, and Wilhelm E. Kincses. Drivers’ misjudgement of vigilance state during prolonged monotonous daytime driving. *Accident Analysis & Prevention*, 41(5):1087–1093, September 2009. doi: 10.1016/j.aap.2009.06.007.
- [54] Izzat Aulia Akbar and Tomohiko Igasaki. Drowsiness estimation using electroencephalogram and recurrent support vector regression. *Information*, 10(6), 2019. ISSN 2078-2489. doi: 10.3390/info10060217.
- [55] Natasha Padfield, Jaime Zabalza, Huimin Zhao, Valentin Masero, and Jinchang Ren. Eeg-based brain-computer interfaces using motor-imagery: Techniques and challenges. *Sensors*, 19(6), 2019. ISSN 1424-8220. doi: 10.3390/s19061423.
- [56] Li-Chen Shi, Ying-Ying Jiao, and Bao-Liang Lu. Differential entropy feature for eeg-based vigilance estimation. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 6627–6630, 2013. doi: 10.1109/EMBC.2013.6611075.
- [57] Agustina Garcés Correa, Lorena Orosco, and Eric Laciari. Automatic detection of drowsiness in eeg records based on multimodal analysis. *Medical Engineering &*
-

- Physics*, 36(2):244–249, 2014. ISSN 1350-4533. doi: 10.1016/j.medengphy.2013.07.011.
- [58] La densidad espectral de potencia: qué es y cómo medirla, Jan 2021. URL <https://www.safeloadtesting.com/es/la-densidad-espectral-de-potencia-que-es-y-como-medirla/>.
- [59] Yu-Ting Liu, Shang-Lin Wu, Kuang-Pen Chou, Yang-Yin Lin, Jie Lu, Guangquan Zhang, Wen-Chieh Lin, and Chin-Teng Lin. Driving fatigue prediction with pre-event electroencephalography (eeg) via a recurrent fuzzy neural network. In *2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 2488–2494, July 2016. doi: 10.1109/FUZZ-IEEE.2016.7738006.
- [60] Kai-Quan Shen, Xiao-Ping Li, Chong-Jin Ong, Shi-Yun Shao, and Einar P.V. Wilder-Smith. Eeg-based mental fatigue measurement using multi-class support vector machines with confidence estimate. *Clinical Neurophysiology*, 119(7):1524–1533, 2008. ISSN 1388-2457. doi: 10.1016/j.clinph.2008.03.012.
- [61] Chin-Teng Lin, Chun-Hsiang Chuang, Chih-Sheng Huang, Shu-Fang Tsai, Shao-Wei Lu, Yen-Hsuan Chen, and Li-Wei Ko. Wireless and wearable eeg system for evaluating driver vigilance. *IEEE Transactions on Biomedical Circuits and Systems*, 8(2):165–176, April 2014. ISSN 1940-9990. doi: 10.1109/TBCAS.2014.2316224.
- [62] B.J. Wilson and T.D. Bracewell. Alertness monitor using neural networks for eeg analysis. In *Neural Networks for Signal Processing X. Proceedings of the 2000 IEEE Signal Processing Society Workshop (Cat. No.00TH8501)*, volume 2, pages 814–820 vol.2, 2000. doi: 10.1109/NNSP.2000.890161.
- [63] Lan-lan Chen, Yu Zhao, Jian Zhang, and Jun-zhong Zou. Automatic detection of alertness/drowsiness from physiological signals using wavelet-based nonlinear features and machine learning. *Expert Syst. Appl.*, 42(21):7344–7355, November 2015. ISSN 0957-4174. doi: 10.1016/j.eswa.2015.05.028.
- [64] Charlotte Jacobé de Naurois, Christophe Bourdin, Clément Bougard, and Jean-Louis Vercher. Adapting artificial neural networks to a specific driver enhances detection and prediction of drowsiness. *Accident Analysis & Prevention*, 121: 118–128, 2018. ISSN 0001-4575. doi: 10.1016/j.aap.2018.08.017.
- [65] Shuyan Hu and Gangtie Zheng. Driver drowsiness detection with eyelid related parameters by support vector machine. *Expert Systems with Applications*, 36(4): 7651–7658, 2009. ISSN 0957-4174. doi: 10.1016/j.eswa.2008.09.030.
- [66] Eike A. Schmidt, Michael Schrauf, Michael Simon, Axel Buchner, and Wilhelm E. Kincses. The short-term effect of verbally assessing drivers’ state on vigilance
-

- indices during monotonous daytime driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 14(3):251–260, May 2011. doi: 10.1016/j.trf.2011.01.005.
- [67] Tomohiko Igasaki, Kazuki Nagasawa, Nobuki Murayama, and Zhencheng Hu. Drowsiness estimation under driving environment by heart rate variability and/or breathing rate variability with logistic regression analysis. In *2015 8th International Conference on Biomedical Engineering and Informatics (BMEI)*, pages 189–193, Oct 2015. doi: 10.1109/BMEI.2015.7401498.
- [68] Yuqi Cuui, Yifan Xu, and Dongrui Wu. Eeg-based driver drowsiness estimation using feature weighted episodic training, 2019.
- [69] Atsuo Murata, Yoshito Fujii, and Kensuke Naitoh. Multinomial logistic regression model for predicting driver’s drowsiness using behavioral measures. *Procedia Manufacturing*, 3:2426–2433, 2015. ISSN 2351-9789. doi: 10.1016/j.promfg.2015.07.502. 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015.
- [70] Jongseong Gwak, Motoki Shino, and Akinari Hirao. Early detection of driver drowsiness utilizing machine learning based on physiological signals, behavioral measures, and driving performance. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, November 2018. doi: 10.1109/itsc.2018.8569493.
- [71] Varun Bajaj, Sachin Taran, Smith K. Khare, and Abdulkadir Sengur. Feature extraction method for classification of alertness and drowsiness states EEG signals. *Applied Acoustics*, 163:107224, June 2020. doi: 10.1016/j.apacoust.2020.107224.
- [72] Roman Rosipal, Björn Peters, Göran Kecklund, Torbjörn Åkerstedt, Georg Gruber, Michael Woertz, Peter Anderer, and Georg Dorffner. EEG-based drivers’ drowsiness monitoring using a hierarchical gaussian mixture model. In *Foundations of Augmented Cognition*, pages 294–303. Springer Berlin Heidelberg, 2007. doi: 10.1007/978-3-540-73216-7_33.
- [73] Anqi Bi, Wenhao Ying, and Lu Zhao. Fast enhanced exemplar-based clustering for incomplete eeg signals. *Computational and Mathematical Methods in Medicine*, 2020:4147807, May 2020. ISSN 1748-670X. doi: 10.1155/2020/4147807.
- [74] Xiao Shuyuan, Wang Bei, Zhang Jian, Zhang Qunfeng, Zou Junzhong, and Matsatoshi Nakamura. Notice of removal: An improved k-means clustering algorithm for sleep stages classification. In *2015 54th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, pages 1222–1227, 2015. doi: 10.1109/SICE.2015.7285326.
-

-
- [75] Tema 10: Estado físico del conductor, 2021. URL <https://vialtest.com/manual-del-conductor/permiso-B/estado-fisico-del-conductor/20>.
- [76] 64-channel quik-cap. URL <https://compumedicsneuroscan.com/product/64-channels-quik-cap-synamps-2-rt/>.
- [77] SMI eye tracking glasses - imotions. URL <https://imotions.com/hardware/smi-eye-tracking-glasses/>.
- [78] Krzysztof Kudrynski and Paweł Strumiłło. An algorithm for parametric modelling of a series of time intervals. In Ryszard S. Romaniuk and Krzysztof S. Kulpa, editors, *Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2009*, page 5. SPIE, June 2009. doi: 10.1117/12.837821.
- [79] Jason Brownlee. Gaussian processes for classification with python, Aug 2020. URL <https://machinelearningmastery.com/gaussian-processes-for-classification-with-python/>.
- [80] Cross-validation: evaluating estimator performance. URL https://scikit-learn.org/stable/modules/cross_validation.html.
- [81] Aniththa Umamahesan. Ajuste de hiperparámetros de un modelo - azure machine learning, Feb 2021. URL <https://docs.microsoft.com/es-es/azure/machine-learning/how-to-tune-hyperparameters>.
- [82] Leonard J. Trejo, Kevin Knuth, Raquel Prado, Roman Rosipal, Karla Kubitz, Rebekah Kochavi, Bryan Matthews, and Yuzheng Zhang. Eeg-based estimation of mental fatigue: Convergent evidence for a three-state model. In Dylan D. Schmorrow and Leah M. Reeves, editors, *Foundations of Augmented Cognition*, pages 201–211, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg. ISBN 978-3-540-73216-7.
- [83] W W Wierwille, S S Wreggit, C L Kirn, L A Ellsworth, and R J Fairbanks. Research on vehicle-based driver status/performance monitoring; development, validation, and refinement of algorithms for detection of driver drowsiness, final report, Dec 1994. URL <https://rosap.ntl.bts.gov/view/dot/2578>. Tech Report.
- [84] S. Lin, Y. Tan, P. Chua, L. Tey, and C. Ang. Perclos threshold for drowsiness detection during real driving. *Journal of Vision*, 12:546–546, 08 2012. doi: 10.1167/12.9.546.
- [85] Burcu Kir Savas and Yasar Becerikli. Real time driver fatigue detection system based on multi-task ConNN. *IEEE Access*, 8:12491–12498, 2020. doi: 10.1109/access.2020.2963960.
-

-
- [86] Qianyang Zhuang, Zhang Kehua, Jiayi Wang, and Qianqian Chen. Driver fatigue detection method based on eye states with pupil and iris segmentation. *IEEE Access*, 8:173440–173449, 2020. doi: 10.1109/access.2020.3025818.
- [87] Feng Zhou, Areen Alsaid, Mike Blommer, Reates Curry, Radhakrishnan Swaminathan, Dev Kochhar, Walter Talamonti, Louis Tijerina, and Baiying Lei. Driver fatigue transition prediction in highly automated driving using physiological features. *Expert Systems with Applications*, 147:113204, June 2020. doi: 10.1016/j.eswa.2020.113204.
- [88] Fang Zhang, Jingjing Su, Lei Geng, and Zhitao Xiao. Driver fatigue detection based on eye state recognition. In *2017 International Conference on Machine Vision and Information Technology (CMVIT)*. IEEE, February 2017. doi: 10.1109/cmvit.2017.25.
- [89] Wang Huan Gu, Yu Zhu, Xu Dong Chen, Lin Fei He, and Bing Bing Zheng. Hierarchical CNN-based real-time fatigue detection system by visual-based technologies using MSP model. *IET Image Processing*, 12(12):2319–2329, November 2018. doi: 10.1049/iet-ipr.2018.5245.
- [90] Gaussian mixture models. URL <https://scikit-learn.org/stable/modules/mixture.html>.
- [91] Magnetoencefalografía: qué es, síntomas, causas, prevención y tratamiento. URL <https://www.topdoctors.es/diccionario-medico/magnetoencefalografia>.
- [92] Calculating psd from a time history file - vibration testing, Jun 2019. URL <https://vru.vibrationresearch.com/lesson/calculating-psd-time-history/>.
- [93] Priyanka A. Abhang, Bharti W. Gawali, and Suresh C. Mehrotra. Chapter 2 - technological basics of eeg recording and operation of apparatus. In Priyanka A. Abhang, Bharti W. Gawali, and Suresh C. Mehrotra, editors, *Introduction to EEG- and Speech-Based Emotion Recognition*, pages 19–50. Academic Press, 2016. ISBN 978-0-12-804490-2. doi: 10.1016/B978-0-12-804490-2.00002-6.
-

Lista de Acrónimos y Abreviaturas

| | |
|----------|-----------------------------------------|
| R^2 | Coeficiente de Determinación. |
| ANN | Artificial Neural Network. |
| AP | Affinity Propagation. |
| AVT | Tarea de Vigilancia Auditiva. |
| BGM | Bayesian Gaussian Mixture. |
| BLE | Bluetooth Low Energy. |
| CC | Coeficiente de correlación. |
| CNN | Convolutional Neural Network. |
| DBI | Índice Davies-Bouldin. |
| DL | Deep Learning. |
| DT | Decision Tree. |
| DWT | Transformación de Onda Discreta. |
| EEG | Electroencefalografía. |
| ELM | Extreme Learning Machine. |
| EOG | Electrooculografía. |
| ESS | Escala de Somnolencia de Epworth. |
| FFT | Transformada Rápida de Fourier. |
| GM | Gaussian Mixture. |
| GP | Gaussian Process. |
| ICA | Análisis de Componentes Independientes. |
| kNN | k-Nearest Neighbors. |
| KSS | Escala de Somnolencia de Karolinska. |
| LSL | Lab Streaming Layer. |
| LSTM | Long Short-Term Memory. |
| MEG | Magnetoencefalografía. |
| ML | Machine Learning. |
| MSLT | Test de Latencias Múltiples del Sueño. |
| MWT | Test de Mantenimiento de la Vigilia. |
| NASA-TLX | NASA Task Load Index. |
| PERCLOS | PERcentage of eye CLOSure. |
| PSD | Densidad Espectral de Potencia. |
| PVT | Tests de Vigilancia Psicomotora. |
| RBF | Radial Basis Function. |

| | |
|---------------|-----------------------------------------------|
| RF | Random Forest. |
| RMSE | Error Cuadrático Medio. |
| RSEFNN | Recurrent Self-Evolving Fuzzy Neural Network. |
| SC | Coeficiente Silhouette. |
| SSS | Escala de Somnolencia de Stanford. |
| SVM | Support Vector Machine. |
| SWAI | Inventario de Actividad vigilia-sueño. |
| TFG | Trabajo Final de Grado. |
| UDP | User Datagram Protocol. |
