

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería



**Estudio del Impacto de un Sistema Integrado de Detección de
Patrones de Señales Biomédicas y de Generación de Pulsos
Binaurales en la Calidad del Sueño de las Personas**

Trabajo de graduación presentado por Diego Alejandro Alegría Ochoa
para optar al grado académico de Licenciado en Ingeniería Electrónica

Guatemala,

2022

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería



**Estudio del Impacto de un Sistema Integrado de Detección de
Patrones de Señales Biomédicas y de Generación de Pulsos
Binaurales en la Calidad del Sueño de las Personas**

Trabajo de graduación presentado por Diego Alejandro Alegría Ochoa
para optar al grado académico de Licenciado en Ingeniería Electrónica

Guatemala,

2022

Vo.Bo.:



(f)

Dr. Luis Alberto Rivera Estrada

Tribunal Examinador:



(f)

Dr. Luis Alberto Rivera Estrada



(f)

MSc. Carlos Esquit



(f)

Ing. Andrés Monterroso

Fecha de aprobación: Guatemala, 6 de enero de 2022.

El desarrollo de este trabajo de investigación es una suma de esfuerzos. Durante varias generaciones el Departamento de Ingeniería Electrónica y Mecatrónica de la Universidad del Valle ha liderado una línea de investigación en el estudio de la neuro-retroalimentación con estímulos para distintas aplicaciones. En este trabajo de exploración, diseño y evaluación, se aborda el problema de la calidad de sueño en las personas siendo este un factor que influye en el desempeño diario, la calidad de vida y la salud en general.

Uno de los desafíos más grandes es salir de la zona de confort, pensar fuera de la caja y tener la capacidad de trabajar con disciplinas ajenas a las ya conocidas según un plan de estudios determinado. Para esta investigación hubo un trabajo exploratorio bastante denso, una curva de aprendizaje sobre los aspectos del sueño, las etapas de sueño y como existen patrones con los que se pueden predecir, utilizando técnicas de predicción innovadoras como el aprendizaje automático. También, un esfuerzo de investigación y diseño sobre como ayudar a las personas con trastornos de sueño utilizando pulsos auditivos, una tecnología de bajo costo que promete ayudar en los trastornos de sueño y conseguir un sueño reparador.

Mi profundo agradecimiento a todas aquellas personas que estuvieron presentes en el camino, a mis padres Rodolfo y Jacqueline que no dudaron nunca en mi potencial, que creyeron en mí aún en días en los que yo no lo hice y me entregaron todos los medios de apoyo que estuvieron en su posibilidad para lograr mis sueños. A mis hermanos Sebastián, Fátima, Ignacio y Rebeca por su amor incondicional y el eterno apoyo que me han otorgado.

Mi gratitud a las familias que me recibieron como uno más en su mesa y me hicieron sentir que, aunque mi familia biológica estaba lejos, tenía un lugar dónde encontrar un hogar. La familia Leal Cahuex, Díaz Rubio, Gómez Werner, Ochoa Maldonado y especialmente a la familia López Aldana quienes representaron un fuerte apoyo en los últimos pasos de esta parte de mi vida, María José, Marlen y Carlos Roberto, estoy eternamente agradecido.

Los amigos que la vida me dio, Pablo Leal, Pablo Yee, Guillermo Aldana, Ronel Argueta, Rodrigo Sandoval, Manolo y José Capilla, Alejandro Paz, Federico Ordoñez, Julio Samayoa, Jorge Súchite, Juan Carlos Morales, Mariano Rubio, Kenneth Catún, Andrés Villafuerte, Aldo y Jaime López, Erick Ríos y a todos aquellos que cuando la situación lo ameritó, me tendieron su mano.

Por último y no menos importante, un reconocimiento a todas aquellas personas que representaron un impacto en mi vida académica y profesional, desde el colegio Marylin de Antonio, Sabrina Milián, Salvador Lemus, Américo Ordoñez y Alfredo Lobato. Los excelentes profesores universitarios Carlos Esquit, Pablo Oliva, Miguel Zea por enseñarme a disfrutar de la vida académica. A Paulo De León, Ricardo Rodríguez y Charles Hess por su apoyo y mentoría en el área profesional. Una mención especial al Dr. Luis Alberto Rivera por su admirable disciplina, la pasión con la que enseña y el apoyo que me dio durante esta investigación.

Me gusta pensar en esta línea de investigación como la idea de un transductor, pequeños dispositivos que convierten energías de otro tipo a señales eléctricas analizables, si encontramos el transductor correcto podemos ser capaces de comprender otros dominios.

Prefacio	IV
Lista de figuras	IX
Lista de cuadros	X
Resumen	XI
Abstract	XII
1. Introducción	1
2. Antecedentes	2
3. Justificación	4
4. Objetivos	6
4.1. Objetivo general	6
4.2. Objetivos específicos	6
5. Alcance	7
6. Marco teórico	8
6.1. Señales Electroencefalográficas (EEG)	8
6.2. Variables utilizadas en la clasificación de actividad de señales electro- encefalográficas (EEG)	9
6.2.1. Frecuencia	9
6.2.2. Amplitud	9
6.2.3. Morfología	10
6.2.4. Sincronía	10
6.2.5. Periodicidad	10
6.3. Obtención de señales electroencefalográficas (EEG)	10

6.3.1.	Electrodos	10
6.3.2.	Electro Cap	11
6.3.3.	Electro Gel	12
6.3.4.	Impedancia	12
6.3.5.	Posicionamiento de electrodos (Sistema 10/20)	12
6.4.	Open BCI	13
6.4.1.	Cython Board	13
6.5.	Arquitectura de las etapas del sueño	14
6.5.1.	Ciclos de sueño para NREM y REM	15
6.5.2.	Las cuatro etapas del NREM	15
6.5.3.	Etapas 1	15
6.5.4.	Etapas 2	16
6.5.5.	Etapas 3 y 4	16
6.5.6.	Sueño REM	16
6.6.	Aprendizaje automático	17
6.6.1.	Aprendizaje supervisado	18
6.6.2.	Aprendizaje no supervisado	19
6.6.3.	Aprendizaje por redes neuronales	20
6.6.4.	Principales clasificadores de aprendizaje automático según tipo de aprendizaje	21
6.6.5.	Algoritmos clasificadores por árboles	21
6.7.	Características de las señales	23
6.7.1.	Dimensión Fractal Petrosiana	24
6.8.	Pulsos binaurales	26
6.9.	Pruebas estadísticas	27
6.9.1.	Diseño completamente aleatorizado	28
6.9.2.	Diseño de bloques aleatorizado	28
6.9.3.	Experimento multifactorial	29
7.	Descarga, extracción y almacenamiento de datos	30
7.1.	Cambios en versiones anteriores y consideraciones	30
7.2.	Obtención de datos de entrenamiento	31
7.2.1.	Descarga de datos	31
7.2.2.	Implementación final de herramienta de descarga	32
7.2.3.	Extracción de épocas	36
7.2.4.	Agrupación y almacenamiento de datos	36
7.2.5.	Gráficos de ejemplo con herramienta de procesamiento	37
8.	Clasificador de etapas de sueños	39
8.1.	Características de las señales	39
8.1.1.	Selección de características de forma automática	40
8.2.	Sobremuestreo de datos de entrenamiento	41
8.3.	Modelos de aprendizaje automático	41
8.3.1.	<i>eXtreme Gradient Boosting (XGBoost)</i>	42
8.3.2.	Bosques aleatorios (<i>Random Forests</i>)	43

9. Pulsos binaurales	45
9.1. Cambios en versiones anteriores y consideraciones	45
9.2. Módulo generador de pulsos binaurales	47
10. Integración de módulos	50
10.1. Cambios en versiones anteriores y consideraciones	50
10.2. Obtención de datos electroencefalográficos	51
10.2.1. Consideraciones y configuración de ambiente	51
10.2.2. Librería BrainFlow	52
10.3. Extracción de características y clasificación	53
10.4. Arquitectura de sistema con plan de sueño	53
10.5. Almacenamiento de datos	54
11. Diseño de estudio estadístico y documentos protocolarios	55
11.1. Consentimiento informado	55
11.2. Protocolo de investigación	56
11.3. Evaluación inicial de impacto	58
12. Conclusiones	60
13. Recomendaciones	61
14. Bibliografía	62
15. Anexos	65

Lista de figuras

1.	Generador de pulsos binaurales en <i>Raspberry Pi</i> [8].	3
2.	Actividad eléctrica cerebral [13].	8
3.	Electrodos de acero inoxidable [13].	11
4.	Electro - Cap [13].	11
5.	Electro - Gel [13].	12
6.	Sistema de posicionamiento 10/20 [13].	13
7.	Cython Board y USB Dongle [14].	14
8.	REM Y NREM [15].	14
9.	Etapas de NREM [15].	15
10.	Aprendizaje de máquinas [17]	18
11.	Aprendizaje supervisado de máquina. [17]	19
12.	Neurona [17].	20
13.	Red neuronal artificial [17].	20
14.	Distribución según género	33
15.	Distribución según edades	33
16.	Gráfico de distribución de datos por clase.	37
17.	Gráfico de polisomnografía.	37
18.	Gráfico de épocas en polisomnografía	38
19.	Gráfico de densidad espectral de potencia	38
20.	Distribución de datos por clase	41
21.	Matriz de confusión para algoritmo XGBoost.	43
22.	Matriz de confusión para algoritmo Random Forest.	44
23.	Flujo de migración a generador de pulsos binaurales en Raspberry Pi [8].	45
24.	Interfaz gráfica de generador de pulsos binaurales [8].	46
25.	Pulso binaural de 12 Hz.	47
26.	Transformada de Fourier de pulso binaural de 12 Hz.	48
27.	Pulso binaural modulado con pista portadora.	48
28.	Transformada de Fourier de pulso binaural modulado con pista portadora.	48

29.	Proceso de creación de pulso binaural.	49
30.	Sistema integrado de neuro-retroalimentación.	50
31.	Retroalimentación en ciclo infinito.	51
32.	Módulos de <i>BrainFlow</i>	52
33.	Relación de coeficiente de Higuchi con etapas de sueño [33].	53
34.	Distribución de etapas de sueño en plan de sueño.	54
35.	Impacto en sueño de terapia en sujeto 1.	58
36.	Impacto de sueño de terapia en sujeto 2.	59

1.	Diferencias fisiológicas NREM-REM [15]	17
2.	Diccionario de etapas de sueño a clases	34
3.	Diccionario de unificación de clases	34
4.	Datos de ejemplo	35
5.	Información de archivo EDF	35
6.	Datos con épocas	36
7.	Características en <i>PyEEG</i> [20]	40
8.	Métricas de precisión en algoritmo XGBoost.	42
9.	Métricas generales de precisión por etapa en algoritmo XGBoost.	42
11.	Métricas generales de precisión por etapa en algoritmo Random Forest.	44
10.	Métricas de precisión en algoritmo Random Forest.	44
12.	Distribución del sueño [34]	54

Este trabajo de investigación consistió en la integración de fases previas para el desarrollo de un sistema de clasificación automática de las etapas de sueño, que además produce pulsos binaurales para disminuir las afecciones producidas por trastornos del sueño. A estos módulos se les realizaron cambios fundamentales para el mejoramiento de desempeño y escalabilidad del sistema.

En el entrenamiento del clasificador de etapas de sueño se utilizaron vectores de características de las señales correspondientes a la complejidad, potencia espectral y entropía para determinar patrones y asociarlos con las etapas de sueño. Estos vectores de características fueron extraídos de datos de noches de sueño para un total de 60 personas sanas, provenientes de una base de datos pública. Un entrenamiento de tipo entre-sujetos permitió adquirir suficientes datos para mejorar la precisión sin necesidad de un estudio intra-sujeto debido a las restricciones producto de la pandemia del virus COVID-19.

Se realizaron modificaciones en el generador de pulsos binaurales añadiendo una pista de fondo que nos de un sonido agradable y estimule el sueño. Un proceso de rediseño fue necesario para ser reproducidos en la computadora, esto con la motivación de poder conectar audífonos inalámbricos buscando comodidad para el sujeto.

En la sección de experimentación estadística para la evaluación de impacto de pulsos binaurales se solicitó apoyo a la dirección de la maestría de neuropsicología de la Universidad del Valle. A través de este proceso se delimitó el grupo a estudiar, se elaboró un consentimiento informado con las buenas prácticas éticas de investigación con un humanos y el respectivo protocolo de investigación para lectura y aprobación del Comité de Ética.

This research consisted of the integration of previous phases for the development of an automatic classification system of sleep stages, which also produces binaural pulses to reduce conditions caused by sleep disorders. Fundamental changes were made to these modules to improve the performance and scalability of the system.

In the training process of the sleep stage classifier, features like complexity, spectral power and entropy were used to determine patterns and associate them with sleep stages. These feature vectors were extracted from data on nights of sleep for a total of 60 healthy people, from a public database. An inter-subject approach of training allowed to acquire enough data to improve precision without the need for an within-subject study due to the restrictions caused by the COVID-19 virus pandemic.

A redesign process was necessary for the binaural pulse generator module adding a background track that gives us a pleasant sound and stimulates sleep. In addition, a reengineering in the generation of pulses was necessary to be able to be reproduced on the computer, this with the motivation of being able to connect wireless headphones seeking comfort for the person.

In the statistical experimentation section of this investigation, for the evaluation of the impact of binaural pulses. Support was requested from the direction of the master's degree in neuropsychology at the Universidad del Valle. Through this process, the group to be studied was delimited, an informed consent was drawn up with good ethical research practices with a human and the respective research protocol for reading and approval by the Ethics committee.

Durante el sueño ocurren eventos importantes, cambios fisiológicos en los órganos y sistemas, se consolida la memoria y el cuerpo entra en un estado de recuperación no solo física sino también psicológica. El sueño es una parte fundamental de la vida cotidiana, una necesidad básica del ser humano y muchas personas sufren de condiciones médicas que no les permiten obtener un descanso reparador al dormir.

Motivado con ofrecer soluciones para las afecciones relacionadas al sueño y mejorar la calidad de vida de las personas, este trabajo tiene la finalidad de presentar resultados sobre la integración y mejora de módulos planteados en fases pasadas de esta investigación. Un módulo de reconocimiento de patrones de señales biomédicas que realiza el trabajo de clasificador para etapas de sueño y un módulo generador de pulsos binaurales que produce como su nombre lo indica pulsos auditivos asociados a las etapas de sueño.

Las técnicas utilizadas en este trabajo de investigación son innovadoras, los sistemas de clasificación automática de etapas de sueño y la terapia de neuro-retroalimentación con pulsos auditivos son tecnologías relativamente nuevas y necesitan validación, es por ello que dentro de esta investigación también se establecen los procedimientos técnicos y éticos para la medición del impacto de los pulsos binaurales en la calidad de sueño de las personas.

Por último, se espera que esta investigación no represente el final del sistema integrado. Los resultados presentados en este trabajo constituyen una base para un proceso continuo de mejora hasta alcanzar una versión final del sistema que incluya más información de sensores, reforzando así el proceso de clasificación y que esta cuente con su respectiva validación científica, el impacto de pulsos binaurales en la calidad de sueño de las personas.

La estimulación cerebral por pulsos binaurales ha sido un tema disruptivo en las últimas décadas por los hallazgos de como a través de procesos de sincronización el cerebro llega a estados cognitivos o de humor solo con escuchar pulsos de sonido que presentan distintas frecuencias para cada oído como lo plantea la investigación en [1]. Sumado a esto, la evolución tecnológica ha permitido monitorear en tiempo real la reacción que el cerebro está teniendo a estos estímulos a través de interfaces de cerebro computador. Estas interfaces permiten tener una idea más cercana de como este proceso de sincronización sucede y las mejorías que se pueden realizar a los métodos. Estas investigaciones sobre inducción de estados cognitivos han ido desde mejorar los procesos de creatividad y consolidación de memoria como en [2] hasta inducir estados de vigilancia como en [3].

En años anteriores se ha realizado un esfuerzo de investigación en la Universidad del Valle de Guatemala relacionado a las señales bioeléctricas asociadas a distintos estados neurológicos y como estas explican su comportamiento y calidad. En 2018 se realizan los primeros estudios [4] y [5], donde se buscaba desarrollar sistemas en donde primero se recolectan las señales producidas por el cerebro a través del procedimiento médico de colección de señales bioeléctricas llamado encefalograma. Luego, a través de estas señales, retro-alimentar al cerebro del sujeto de estudio para mejorar su respuesta a ciertos estados neurológicos como el estrés o cuando existe una necesidad de atención profunda.

En 2019 se continúan las investigaciones sobre cómo mejorar los procesos neuronales en [6]. Esta vez ya se enfocan específicamente en el estudio del sueño y se realiza análisis exploratorios de cómo combatir los problemas de sueño en atletas, estudio totalmente justificado en la dolencia que sufren para poder conciliar el sueño y esto provocando un déficit en su desempeño profesional. Las investigaciones empiezan a tomar un sendero relacionado a lo antes mencionado usando técnicas de neuro-retroalimentación con pulsos binaurales para poder centrar la frecuencia de

operación de ciertas ondas cerebrales asociadas al sueño. Para esta fase del proyecto se empezaron a diseñar distintos módulos como lo son la interpretación de ondas cerebrales con algoritmos de clasificación de aprendizaje automático así como un dispositivo que pueda emitir estos pulsos que en aquel entonces fueron probados desde la computadora.

Para el año de 2020 se unen dos estudiantes más al proyecto y dividen el trabajo. La primera parte consistió en el dispositivo clasificador, el cual fue entrenado con algoritmos de aprendizaje automático alimentado a través de un proceso intra-sujeto [7]. La otra parte del trabajo fue el desarrollo de un dispositivo de producción de pulsos binaurales en una *Raspberry Pi* como se puede observar en la Figura 1, la cual permite movilidad y además una aproximación a un sistema embebido el cual puede ser usado en clínicas de sueño para facilitar los análisis [8], pero con las restricciones de la arquitectura de computadora que este dispositivo tiene. En [7] y [8] podemos notar que son trabajos complementarios lo que abre espacio a esta investigación para un proyecto integrador y además de estudio de sus efectos.

El trabajo de las fases anteriores concluye con avances importantes, pero dejando aún incógnitas como el diseño de un estudio cuantitativo o estadístico que valide la eficacia de los pulsos binaurales, la población a evaluar y las mejores prácticas para hacerlo. No se determinó en fases anteriores cuáles son las mejores frecuencias para las distintas fases de sueño, seleccionar el mejor algoritmo clasificador según las necesidades. Por último y más importante realizar una integración del generador de pulsos binaurales con el módulo clasificador.

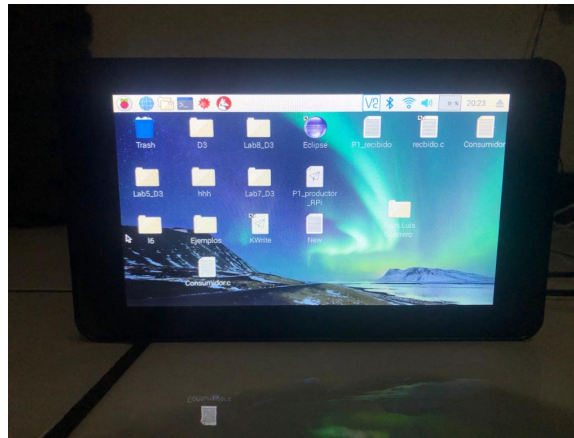


Figura 1: Generador de pulsos binaurales en *Raspberry Pi* [8].

La deficiencia crónica de sueño, definida como el estado inadecuado o mínimo de sueño, es un factor determinante y infravalorado en lo que concierne al estado de salud según [9]. La privación de sueño contribuye en una gran parte al desarrollo de trastornos a niveles moleculares, inmunes y neurológicos, todos estos factores antes mencionados sirve para el desarrollo y exacerbación de enfermedades cardíacas y del metabolismo por lo tanto reduciendo la calidad de vida y la longitud de esta. He aquí una de las motivaciones fundamentales que conducen esta investigación, ya que con el sistema a desarrollar se estaría mejorando la calidad de sueño y por lo tanto de vida de las personas con trastornos de sueño.

En fases anteriores se realizaron importantes avances en esta línea de investigación, desarrollando los módulos que conforman a este sistema de estimulación por pulsos binaurales como lo son el generador de pulsos [8] y el clasificador con aprendizaje automático [7] por separado. Es entonces tarea de esta investigación completar el sistema realizando un trabajo integrador de módulos y validarlo a través de pruebas estadísticas.

Motivado por la integración final de este sistema, es de suma importancia evaluar el trabajo de los clasificadores estudiados en [7] y de otros algoritmos no estudiados para escoger el que presente el mejor desempeño, además que el clasificador anterior fue entrenado con datos intra-sujeto, es decir que este presenta una buena precisión de predicción pero para datos del mismo sujeto, es por ello necesario implementar un entrenamiento más variado. Esta evaluación se realizará no solo a través de datos para testear sino también comparar con estándares médicos sobre la clasificación que sugieren para los datos colectados.

En el estudio del generador de pulsos binaurales [8] se realizó el prototipo del módulo generador de pulsos binaurales mas no se definen frecuencias específicas para inducir las fases de sueño en el orden correcto. No existe una función para reproducir sin necesidad de antes guardar, esto generando una pérdida en el tiempo de

procesamiento. A través de investigación y pruebas se realizarán recomendaciones acompañadas de una opinión profesional para luego diseñar un estudio cuantitativo que evalúe si existe un impacto significativo en la calidad de sueño a través de la terapia de neuro-retroalimentación.

4.1. Objetivo general

Integrar el módulo de reconocimiento de patrones de señales biomédicas con el módulo de generación de pulsos binaurales, desarrollados en fases previas, y realizar un estudio del impacto del sistema en la calidad del sueño de las personas.

4.2. Objetivos específicos

- Realizar un trabajo de integración del módulo de reconocimiento de patrones de señales biomédicas relacionadas a las etapas de sueño con el módulo de implementación de un sistema generador de pulsos binaurales para el estudio del impacto de los pulsos en la calidad de sueño propios de la fase anterior del proyecto.
- Mejorar los clasificadores de etapas de sueño basados en algoritmos de aprendizaje automático desarrollados en fases previas del proyecto, y validar los resultados mediante asesoría médica y comparación con estudios similares.
- Definir perfiles de individuos a estudiar y obtener señales bioeléctricas utilizadas en el análisis de etapas de sueño en una clínica.
- Realizar un análisis estadístico para determinar si hay un impacto significativo de los pulsos binaurales en las señales bioeléctricas asociadas a las etapas de sueño.

Este proyecto representa la tercera fase de una línea de investigación que pretende desarrollar y estudiar herramientas que mejoren la calidad de sueño en las personas y por consiguiente la calidad de vida. En las fases anteriores [7] y [8] se estudiaron componentes claves para desarrollar un dispositivo generador de pulsos binaurales activado según la etapa de sueño que un clasificador automático determine al recibir las señales bioeléctricas provenientes de un casco con electrodos.

Motivado por la necesidad de una integración de etapas anteriores, en este trabajo se abordó la mejora del algoritmo clasificador de etapas de sueño utilizando técnicas de aprendizaje automático entrenando a este con la bases de datos públicas como Sleep-EDFX [10]. Se delimitó el uso de un solo canal para realizar la extracción de características y el entrenamiento de la red como en [11]. Esto se debe a que el enfoque multicanal es afectado por el movimiento del paciente en cuestión provocando así una distorsión en los datos y por ende en la clasificación. Una de las ventajas que presenta utilizar un enfoque de canal único también es el bajo costo computacional de cálculos y la velocidad de estos.

Según [3] los pulsos binaurales conducen al cerebro a alinearse en las frecuencias del pulso, para ello entonces se realizó un trabajo de investigación sobre las frecuencias asociadas a las etapas del sueño en adultos jóvenes. Para ello se escogió una frecuencia transportadora y parámetros modificables según la necesidad.

Por último, cabe mencionar que dentro de los objetivos se planteó no solo el diseño del experimento estadístico sino también la ejecución de estos. Se desarrolló el protocolo de investigación, un consentimiento informado y una metodología para medir el impacto pero, debido a la pandemia del COVID-19 y las restricciones de movilidad y medidas de distanciamiento social no fue posible la ejecución del experimento y la recolección de datos en la escala pretendida originalmente.

6.1. Señales Electroencefalográficas (EEG)

El método de comunicación de nuestro cerebro es químico, pero también eléctrico, lo cual genera una gran ventaja debido a que esta actividad puede ser monitorizada desde métodos no invasivos tomando la información a través de electrodos que registren la actividad eléctrica y a través de distintos canales poder asociar las frecuencias de cada onda a estados del cerebro humano y las actividades relacionadas a este estado. Esta frecuencia de las ondas es el principal indicador del estado mental y se mide en Hercios (Hz) o ciclos por segundo con variaciones desde 0.5 hasta 38 Hz y una amplitud asociada a la intensidad de la onda. El procedimiento para la medición de estas ondas se denomina encefalograma o por sus siglas ECG [12].

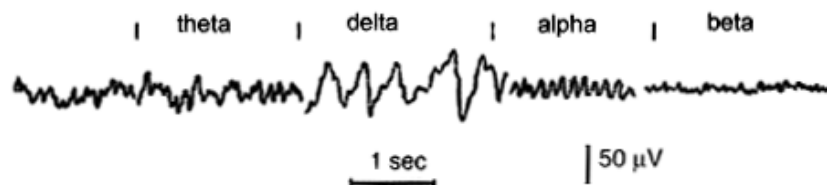


Figura 2: Actividad eléctrica cerebral [13].

Las ondas cerebrales se clasifican según los rangos de frecuencia donde se encuentran contenidas y las cuatro categorías son: Beta, Alfa, Theta y Delta.

- Ondas Beta: Estas se encuentran contenidas entre los 14 a 30 Hz, comúnmente se asocian a estados con una alta actividad mental en situaciones como estudio,

resolución de problemas, ataques de ansiedad o concentración plena. Se asocia a una actividad intensa del cerebro donde se encuentra despierto, con extroversión y una alta percepción por parte de los sentidos [12].

- Ondas Alfa: Se encuentran entre los 8 a 13.9 Hz, se pueden encontrar en momentos antes de dormir, relajación mental y no actividad. Presentan una mayor amplitud [12].
- Ondas Theta: Una mayor amplitud en relación a las alfa, estas se encuentran entre los 4 a 7.9 Hz y están asociadas con un estado de sueño, meditación profunda, con actividades creativas y que requieren inspiración [12].
- Ondas Delta: Son las ondas más lentas y además con mayor amplitud de todas, en estas se encuentran estados como el sueño profundo o una concentración plena en donde se requiere de una operación integral del cerebro [12].

6.2. Variables utilizadas en la clasificación de actividad de señales electroencefalográficas (EEG)

6.2.1. Frecuencia

Frecuencia se refiere a la tasa de repetición de un ciclo por unidad de tiempo, como antes se menciona para las señales electroencefalográficas se presentan distintas señales con distintos comportamientos, entre ellos podemos listar [13]:

- Rítmico: Actividad electroencefalográfica con señales de frecuencia constante.
- Arrítmico. Señales electroencefalográficas donde no hay un ritmo constante (frecuencia).
- Disrítmico. Ritmos especiales no presentes en sujetos normales sino en pacientes con algún tipo de trastorno presente.

6.2.2. Amplitud

Amplitud, medido en voltios, es la cantidad de actividad eléctrica presente en una onda, esto puede depender de la técnica de medición [13].

- Atenuación (también supresión o depresión). Es la reducción de la actividad eléctrica, cuando esta es disminuida a través de un estímulo se le llama señal bloqueada o que presenta un bloqueo [13].
- Hiper sincronía. Es cuando se presenta un aumento en la actividad eléctrica o en el ritmo, entre las señales alpha, beta, o theta. El fenómeno se asocia a la cantidad de neuronas participando en la tarea [13].

- Paroxístico. Actividad en donde se alcanza una alta amplitud en un corto período de tiempo para luego caer en una depresión abrupta a niveles normales. No necesariamente debe de ser una anormalidad pero la mayoría de anormalidades si las presenta [13].

6.2.3. Morfología

Se refiere a la forma de la señal u onda. La forma de la señal se determina por la frecuencia y el voltaje en conjunto. Los patrones se pueden definir como los siguientes [13]:

- Monomórfica. Patrón dominado por una actividad en específico.
- Polimórfica. Composición de varias formas para formar un patrón complejo constante.
- Sinusoidal. Con forma de función matemática seno. La actividad monomórfica es usualmente sinusoidal.
- Trasiente. Patrón aislado y completamente diferente a la actividad anterior.
 - Spike: Subida abrupta de amplitud que dura entre 20 hasta 70 msec.
 - Sharp wave: Subida abrupta de larga duración entre 70 hasta 200 msec.

6.2.4. Sincronía

Sincronía se refiere a la aparición simultánea y rítmica de varios patrones sobre distintas regiones, para este caso de las ondas cerebrales en distintas regiones siendo estas unilaterales o bilaterales [13].

6.2.5. Periodicidad

Periodicidad se refiere a la distribución de elementos de un patrón en una unidad de tiempo y si este se repite o es actividad inusual [13].

6.3. Obtención de señales electroencefalográficas (EEG)

6.3.1. Electrodo

Pequeños discos de un material metálico, usualmente acero inoxidable como se observan en la Figura 3, pero también se pueden encontrar de oro, plata o estaño. Los

electrodos usualmente se cubren con materiales como polímeros o recubrimientos de gel, cada uno de estos se encuentra en una posición definida que se estandariza en el sistema internacional 10/20, la cual será explorada en la sección 6.3.5. Se utilizan las letras y números para etiquetar áreas del cerebro. El sistema 10/20 establece que la letra F se utiliza para reconocer el lóbulo frontal y la letra T para el lóbulo temporal, los números para denotar el lado del cráneo, es decir los números pares son para el lado derecho del cráneo y los números impares para el lado izquierdo.



Figura 3: Electrodo de acero inoxidable [13].

6.3.2. Electro Cap

El *Electro Cap* (Gorro de electrodos en español) es una técnica utilizada para realizar electroencefalogramas con una mayor densidad de señales puesto que se colocan más electrodos ordenados según el sistema 10/20 y obteniendo un bus de salida para analizar las señales [13].



Figura 4: Electro - Cap [13].

6.3.3. Electro Gel

Utilizado como una extensión maleable de los electrodos, esto ayuda a maximizar el contacto con la piel y a tener una menor resistencia en la lectura de las señales mejorando así la calidad de la medición [13].



Figura 5: Electro - Gel [13].

6.3.4. Impedancia

La impedancia es la medida de oposición a la corriente que pasa a través de un elemento a una frecuencia determinada, esta se mide en Ohms. A medida que la impedancia aumenta, esta presenta mayor resistencia a un flujo de corriente. Para el estudio de las señales biomédicas se desea una impedancia baja en los electrodos para que no exista una atenuación en la amplitud de las señales obtenidas y por ende estas sean de mejor calidad [13].

6.3.5. Posicionamiento de electrodos (Sistema 10/20)

El sistema se basa en dividir el cráneo medido desde la región *Nasión* que es el punto de intersección del hueso frontal y de dos huesos nasales llegando hasta la región *Inión* la cual es la proyección más prominente del hueso occipital en la parte posteroinferior del cráneo. A estas divisiones de los porcentajes se les asigna un nombre y un código empezando por el Polo Frontal (Fp), Central (C), Parietal (P), Occipital (O) y Temporal (T). Los electrodos ubicados en el centro del cráneo se les da la letra Z (Por el número cero, en inglés *Zero*. A los puntos en el hemisferio derecho se les asignan números pares y a los puntos del hemisferio izquierdo se les asignan números impares [13].

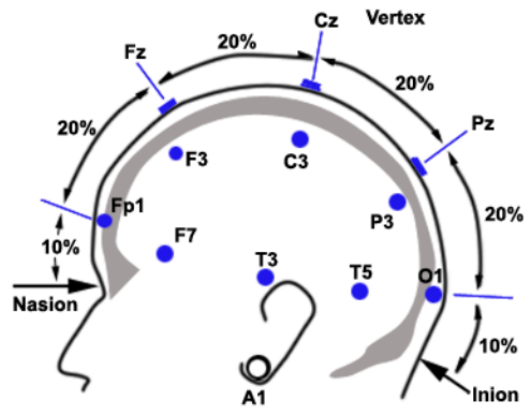


Figura 6: Sistema de posicionamiento 10/20 [13].

6.4. Open BCI

BCI por sus siglas en inglés *Brain Computer Interface* o Interfaz de cerebro computador, es un término acuñado en el estudio de señales electroencefalográficas y su aplicación en distintas ramas. Open BCI entonces es una plataforma de fuente libre la cuál se dedica a diseñar *Software* y *Hardware* especializado para tomar mediciones a señales biomédicas y luego visualizarlas en la computadora o utilizarlas para posterior control en aplicaciones embebidas, posee una gama de productos variada pero se han centrado en desarrollar tarjetas con circuitos embebidos especializados en procesamiento de señales bioeléctricas con amplificadores (puesto que estas señales son de baja amplitud) [14].

6.4.1. Cython Board

La *Cython Board* es una tarjeta con un microcontrolador compatible con la suite de Arduino, posee 8 canales de recepción de señales con procesador de arquitectura de 32 bits implementando un PIC32MX250F128B dando una apertura al procesamiento rápido de 250 Hz, una memoria local que satisface las necesidades y que además ya viene pre-instalado el *bootloader* con la última versión del *software* que esta organización ofrece. Con esta tarjeta podemos monitorear actividades electroencefalográficas, musculares y cardíacas, posee una comunicación inalámbrica gracias a su integración con un dispositivo *USB* llamado *Dongle* que implementa comunicación por radiofrecuencia. Además de ser compatible con el protocolo *Bluetooth Low Energy* o BLE abriendo la posibilidad de comunicación con cualquier otro dispositivo que tenga habilitado este tipo de comunicación [14].

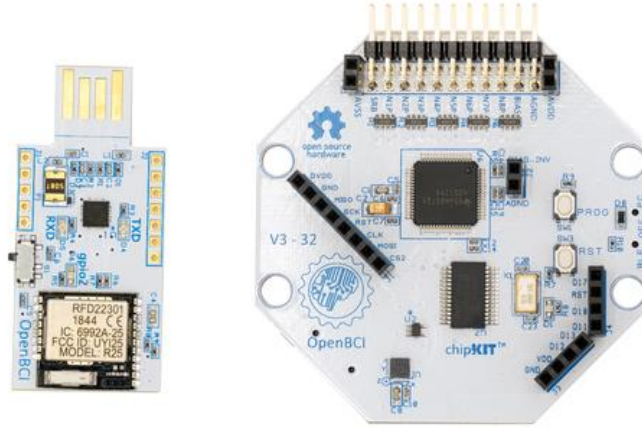


Figura 7: Cython Board y USB Dongle [14].

6.5. Arquitectura de las etapas del sueño

Arquitectura del sueño se refiere a la organización estructural básica de un sueño normal. Hay dos tipos de sueño, NREM por sus siglas en inglés (*Non-Rapid Eye Movement*), y REM por sus siglas en inglés (*Rapid Eye Movement*) la cual se divide en cuatro etapas. Cada una de estas presenta patrones característicos en las ondas cerebrales, movimiento de los ojos, actividad muscular, etc.

Conforme va evolucionando el período de sueño, las etapas NREM y REM se alternan de forma cíclica, el motivo de esto aún no está bien explicado pero se cree que la intercalación irregular de estas etapas está asociado a desordenes de sueño. Como ejemplo podemos mencionar que una persona con narcolepsia en lugar de entrar en la etapa NREM como es usual, entran en una etapa REM [15].

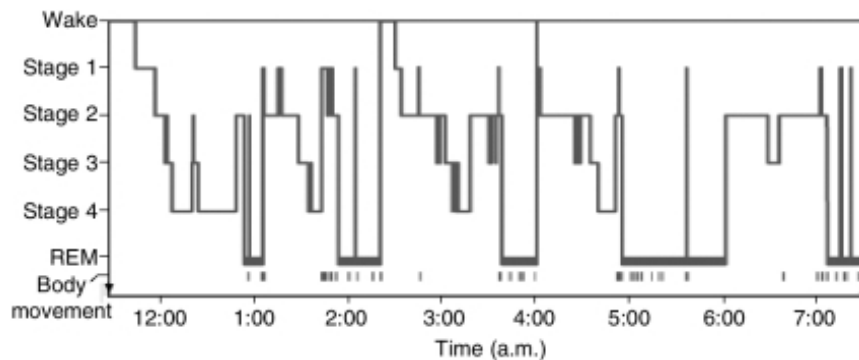


Figura 8: REM Y NREM [15].

6.5.1. Ciclos de sueño para NREM y REM

Un episodio de sueño comienza con un pequeño período de NREM en la etapa 1 progresando a la etapa 2 y por consecuente 3 y 4 hasta llegar a la etapa REM. Sin embargo, una persona no se queda en el ciclo REM el resto de la noche sino que comienza una intercalación entre las etapas de NREM y REM durante la noche como lo vemos en la Figura 8. La etapa NREM constituye un 75 a 70 por ciento del total de la noche mientras que la etapa REM el resto de esta. El promedio de la primera fase NREM-REM dura alrededor de 70 a 100 minutos, las siguientes tienen un tiempo promedio de 90 a 120 minutos. [15]

6.5.2. Las cuatro etapas del NREM

Cada una de estas etapas está asociada con una distinta actividad cerebral como lo podemos observar en la Figura 9 en donde se muestran patrones característicos de cada etapa. También existen instrumentos para seguir la actividad de los ojos y músculos para complementar el etiquetado y caracterización de cada etapa [15].

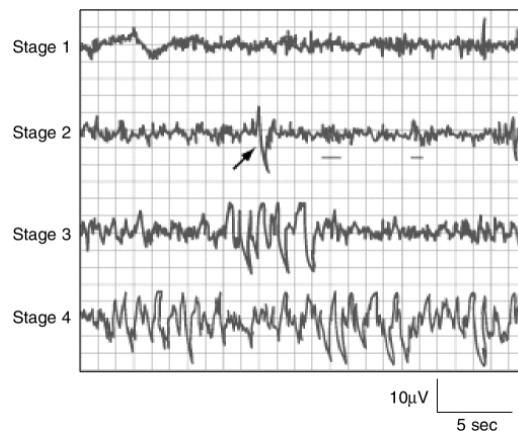


Figura 9: Etapas de NREM [15].

6.5.3. Etapa 1

La etapa 1 de la fase NREM juega un papel fundamental puesto que es la transición hacia dormir, sin contar a aquellas personas que presentan desórdenes de sueño y recién nacidos, esta etapa es la primera etapa de todo el ciclo de sueño y dura de 1 a 7 minutos aproximadamente. Esta etapa es fácilmente interrumpida por ruidos fuertes y dado su naturaleza transitoria nos puede llevar a estar despiertos, la actividad cerebral en esta etapa 1 comienza en patrones que marcan la vigilia como lo son las ondas alfa con patrones rítmicos llegando a amplitudes bajas con frecuencias mezcladas. Las ondas alfa asociadas con el estado de una vigilia en relajación están caracterizadas por frecuencias de 8 a 13 ciclos por segundo (Hz) [15].

6.5.4. Etapa 2

Esta etapa dura aproximadamente de 10 a 25 minutos en su ciclo inicial y se va expandiendo con cada ciclo exitoso de sueño, eventualmente va a constituir de 45 a 55 por ciento del total del tiempo en el episodio de sueño completo. En esta etapa despertar es más difícil y se requieren estímulos más fuertes que en la etapa 1. La actividad de esta etapa muestra amplitudes bajas, frecuencias mezcladas caracterizadas por la aparición de husos, se considera que esta parte es importante para la consolidación de la memoria. Las personas que durante su día tuvieron la oportunidad de aprender una nueva tarea o habilidad tienden a presentar una mayor densidad de husos en los patrones de esta etapa [15].

6.5.5. Etapa 3 y 4

La etapa 3 y 4 se tienden a agrupar en un estado denominado SWS por sus siglas en inglés *Slow-Wave Sleep* o sueño de ondas lentas, la mayoría de veces esta se presenta durante el primer tercio del episodio de sueño. Ambas etapas tienen patrones que les caracterizan, la etapa 3 dura solo unos minutos y constituye de 3 a 8 por ciento de todo el sueño, en esta etapa se presenta un incremento en la amplitud de las ondas y se presenta SWS.

La última etapa de la fase NREM, la etapa 4, se constituye de un tiempo de 20 a 40 minutos en el primer ciclo y representa un 10 a 15 por ciento de todo el episodio de sueño, es la etapa en donde se presenta una mayor actividad en la amplitud de las ondas SWS [15].

6.5.6. Sueño REM

El sueño REM es definido por la presencia de ondas cerebrales de-sincronizadas con una amplitud baja y frecuencias mezcladas, movimientos rápidos de los ojos, presencia de patrones “Diente de Sierra”, una actividad fuerte en las ondas teta (3 a 7 Hz) además de una actividad lenta de ondas alfa que caracterizan al sueño REM.

Hay diferencias fisiológicas importantes entre las etapas REM y NREM y una de ellas es que durante el sueño REM se desactiva el movimiento del cuerpo (Atonía muscular) debido a que se asocia la fase REM con los sueños y para evitar que un individuo se mueva durante un sueño o pesadilla. Aproximadamente el 80 por ciento de los sueños vívidos resultan durante esta etapa del sueño, además la consolidación de la memoria se asocia durante el sueño REM [15].

Proceso fisiológico	NREM	REM
Actividad cerebral	Disminuye respecto a la vigilia	Aumenta respecto a vigilia
Pulso del corazón	Realentiza respecto a la vigilia	Aumenta respecto NREM
Presión sanguínea	Disminuye respecto a la vigilia	Aumenta respecto a NREM
Actividad del nervio simpático	Disminuye respecto a la vigilia	Aumenta respecto a vigilia
Tono muscular	Similar a la vigilia	Ausencia
Flujo sanguíneo al cerebro	Disminuye respecto a la vigilia	Aumenta respecto NREM
Respiración	Disminuye respecto a la vigilia	Aumenta respecto NREM
Airway resistance	Aumenta respecto a la vigilia	Aumenta respecto a la vigilia
Temperatura corporal	Menor que a la vigilia	No regulada,
Excitación sexual	No ocurre con frecuencia	Mayor que en NREM

Cuadro 1: Diferencias fisiológicas NREM-REM [15]

6.6. Aprendizaje automático

El paradigma original del método científico plantea que si se introduce datos y reglas este dará como salida una respuesta esperada, esto según la evolución de la tecnología ha sido desafiado por el paradigma de la inteligencia artificial el cual plantea que si introducimos datos y respuestas podamos entender las reglas del comportamiento de un fenómeno, es así como las ramas de la inteligencia como el aprendizaje automático (*Machine Learning*) es usado para enseñarle a máquinas (computadores) como manejar los datos de una forma más eficiente. Muchas veces los patrones contenidos en los datos no son observables desde un solo punto de vista, es por ello que aplicamos aprendizaje automático para observar la “foto completa”. Dentro de este campo del aprendizaje automático encontramos varias divisiones, para delimitar las que son importantes para esta investigación se consideran tres de las más grandes, aprendizaje supervisado, aprendizaje no supervisado y redes neuronales. Estos algoritmos se han organizado según su taxonomía [16].

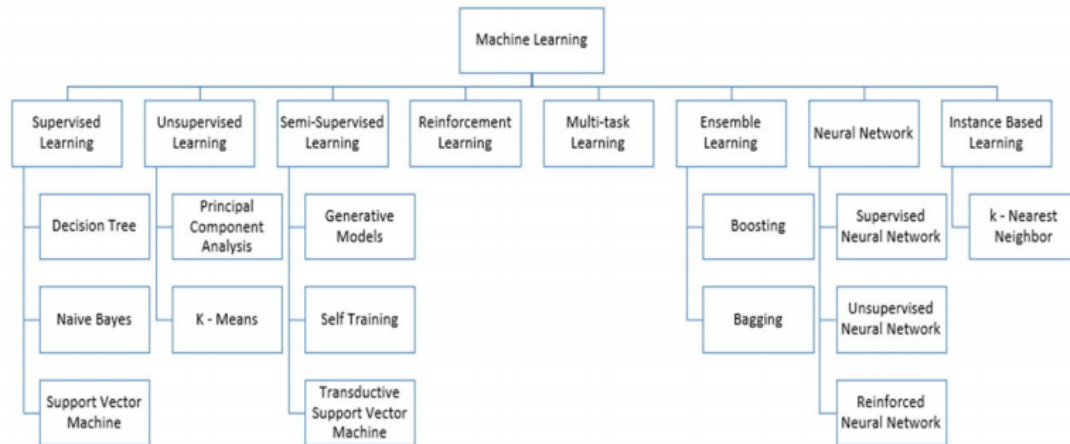


Figura 10: Aprendizaje de máquinas [17]

6.6.1. Aprendizaje supervisado

La principal característica de este tipo de algoritmos es que este genera una función de salida que mapea las entradas en salidas deseadas. El principal problema que se formula en los algoritmos no supervisados es el de clasificación, la unidad de aprendizaje tiene como meta aprender a mapear una función que mapea información sobre una entrada hacia una categoría. [17]

Se le llama supervisado porque usualmente esta función de mapeo a clasificaciones implementa categorías que se han determinado por el diseñador del algoritmo, es importante resaltar que cuando se diseñan este tipo de algoritmos tiene que ser fácil determinar las categorías dentro de los datos que obtenemos para realizar predicciones y mapear de forma correcta. Entre los principales algoritmos presentes en el aprendizaje supervisado tenemos las redes neuronales y los árboles de decisiones, podemos notar que en este tipo de algoritmos dependemos de forma completa de las entradas y de las categorías previamente diseñadas

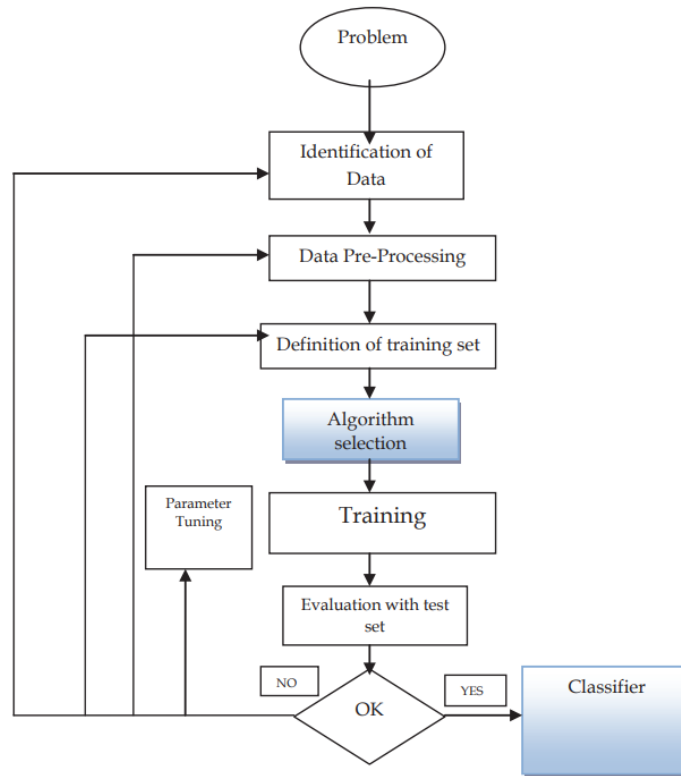


Figura 11: Aprendizaje supervisado de máquina. [17]

6.6.2. Aprendizaje no supervisado

Como su nombre lo sugiere, en este tipo de algoritmos se pretende que la computadora aprenda a realizar algo que no se le indica explícitamente, retomando el ejemplo de los clasificadores uno de los principales problemas en los cuáles se enfoca el aprendizaje automático, a los algoritmos de aprendizaje no se le va a dotar de categorías ya establecidas, en cambio se le va a indicar éxito con un sistema de recompensa haciendo referencia al mundo real donde de igual forma se recompensa el trabajo bien hecho o se penaliza el trabajo que no obtuvo los resultados esperados [17].

Uno de los principales algoritmos es el denominado *Clustering*, este algoritmo trata de buscar similitudes en los datos de entrenamiento del modelo. El presunto es que estos *clusters* o agrupaciones descubiertas van a coincidir con clasificaciones intuitivas sin embargo este no nos va a devolver nombres para las categorías, es aquí donde el diseñador examina los resultados y analiza si satisface sus necesidades. Es importante hacer mención que para utilizar este tipo de algoritmos se requiere la cantidad de datos para que nos de resultados fidedignos [17].

6.6.3. Aprendizaje por redes neuronales

El aprendizaje por redes neuronales (Comúnmente llamadas redes neuronales artificiales o ANN por sus siglas en inglés *artificial neural networks*) se deriva del concepto biológico de las neuronas. Una neurona es una estructura que funciona como celda en el cerebro, para entender como funciona una red primero debemos entender como funciona su unidad básica. Las neuronas tienen cuatro partes básicas que se pueden apreciar en la siguiente Figura 12 [17].

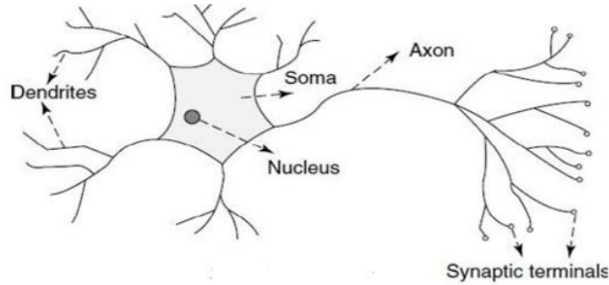


Figura 12: Neurona [17].

Las dendritas (*Dendrites*) reciben el impulso eléctrico, el soma lo procesa y la salida es guiada por el axón hacia la dendrita de otra neurona, la interconexión de varias neuronas es llamada red neuronal.

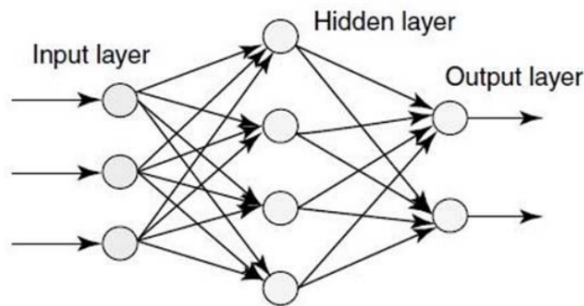


Figura 13: Red neuronal artificial [17].

Una red neuronal artificial se comporta de la misma forma. Existen tres tipos de capas, las de entrada (similares a las dendritas), las capas ocultas (similares al soma y el axón) y la capa de salida (similar a las terminales de las dendritas). Cada una de estas neuronas tiene una función que dependiendo de la entrada se activará o no y pasará esta información a la siguiente capa como se puede observar en la Figura 13.

6.6.4. Principales clasificadores de aprendizaje automático según tipo de aprendizaje

- Supervisados según [16]
 - Árbol de decisiones
 - *Naïve Bayes*
 - Máquina de vector de soporte
 - *Random Forest*
 - *Gradient Boosting*
- No supervisados según [16]
 - *K-Means Clusters*
 - Análisis de componente principal
- Redes neuronales artificiales según [16]
 - Redes supervisadas
 - Redes no supervisadas
 - Redes neuronales reforzadas

6.6.5. Algoritmos clasificadores por árboles

Algoritmo de *XGBoost*

Sea un vector de entrenamiento $\{(x_i, y_i)\}_{i=1}^N$, una función diferenciable de costo $L(y, F(x))$, un número M de aprendices débiles y una tasa de aprendizaje α

Se inicializa el modelo con un valor constante:

$$\hat{f}_{(0)}(x) = \arg \min_{\theta} \sum_{i=1}^N L(y_i, \theta).$$

Para $m=1$ hasta M se calculan los gradientes y *hessianos*:

$$\hat{g}_m(x_i) = \left(\left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right] \right)_{f(x)=\hat{f}_{(m-1)}(x)}.$$

$$\hat{h}_m(x_i) = \left(\left[\frac{\partial^2 L(y_i, f(x_i))}{\partial f(x_i)^2} \right] \right)_{f(x)=\hat{f}_{(m-1)}(x)}.$$

Entrenar un aprendiz básico (o aprendiz débil, por ejemplo un árbol) utilizando el vector de entrenamiento $\{x_i, -\frac{\hat{g}_m(x_i)}{\hat{h}_m(x_i)}\}_{i=1}^N$ resolviendo el siguiente problema de optimización:

$$\hat{\phi}_m = \arg \min_{\phi \in \Phi} \sum_{i=1}^N \frac{1}{2} \hat{h}_m(x_i) \left(\left[-\frac{\hat{g}_m(x_i)}{\hat{h}_m(x_i)} - \phi(x_i) \right] \right)^2.$$

$$\hat{f}_m(x) = \alpha \hat{\phi}_m(x).$$

Actualizar el modelo:

$$\hat{f}_{(m)}(x) = \hat{f}_{(m-1)}(x) + \hat{f}_m(x).$$

Vector de salida:

$$\hat{f}(x) = \hat{f}_{(M)}(x) = \sum_{m=0}^M \hat{f}_m(x).$$

Algoritmo de Bosques Aleatorios

Para definir un clasificador que implementa Bosques Aleatorios, primero se debe definir los elementos que lo componen, los árboles de decisión, cuáles son las métricas para las divisiones y como se realiza luego el proceso de votación.

El algoritmo comienza con la construcción de un árbol de arriba hacia abajo, donde cada separación se va a realizar dependiendo de la variable y si esta nos otorga una mayor ganancia de información o reduce la entropía, dependiendo de la métrica de medición, los arboles van a cambiar la dinámica en como se genera la siguiente división. Para el caso de esta investigación se utilizó la métrica de Impureza de Gini. [18]

Nombrado por el matemático Corrado Gini, esta métrica mide que tantas veces un elemento aleatorio va a ser clasificado de forma incorrecta si este se selecciona de forma aleatoria según la distribución de clases en la porción de datos utiliza para entrenar. Esta métrica se puede considerar una suma de probabilidades p_i de un elemento con clasificación i multiplicado por la probabilidad $\sum_{k \neq i} p_k = 1 - p_i$ de ser escogido de la forma incorrecta. [18]

Para calcular la impureza de Gini para un juego de datos con J clases, suponemos un valor $i \in \{1, 2, \dots, J\}$ tal que p_i sea la fracción de elementos etiquetados con la clase i [18]

$$I_G(p) = \sum_{i=1}^J ((p_i \sum_{k \neq i} p_k)) = \sum_{i=1}^J p_i(1 - p_i) = \sum_{i=1}^J (p_i - p_i^2) = \sum_{i=1}^J p_i - \sum_{i=1}^J p_i^2 = 1 - \sum_{i=1}^J p_i^2$$

Este algoritmo de clasificación por árboles es ahora utilizado en un grupo, para el llamado Aprendizaje en Conjunto (*Ensemble Learning*) y aplicar el algoritmo de bosques aleatorios a través de una técnica llamada *bagging*.

Dando un conjunto de datos de entrenamiento $X \in \{X_1, X_2, \dots, X_n\}$ con salidas $Y \in \{Y_1, Y_2, \dots, Y_n\}$, agrupadas un número B de veces en la cuál se selecciona de forma aleatoria los segmentos del conjunto de aprendizaje para entrenar distintos árboles con estas muestras.

Para cada iteración desde b=1 hasta B realizamos el siguiente procedimiento:

- Muestrear dentro del conjunto de entrenamiento una porción para entrenar a los árboles, se define ahora como X_b y Y_b .
- Se realiza la clasificación f_b para X_b y Y_b .

Una vez obtenida la predicciones del grupo total de árboles en el bosque se procede a seleccionar la clase que la mayor cantidad de árboles seleccionó.

6.7. Características de las señales

Las características son fundamentales cuando se habla de análisis de datos y algoritmos de aprendizaje automático. En los trabajos de clasificación y regresión los valores de las características extraídas de la información son las variables sobre las cuales se basan las predicciones [19]. Dependiendo de el tipo de información que se está investigando se necesitan distintos tipos de características, para esta investigación se utilizaron aquellas que extraen información de series de tiempo.

Densidad espectral de potencia

Sea una serie de tiempo $\{x_1, x_2, \dots, x_n\}$, denotamos su transformada rápida de Fourier (FFT) teniendo como resultado $\{X_1, X_2, \dots, X_n\}$. Se tiene también una banda continua con frecuencias desde f_{baja} hasta f_{alta} dividiendo el vector en K contenedores. Los límites para cada contenedor están dados por un vector $\{f_1, f_2, \dots, f_n\}$ tal que las frecuencias bajas y altas del i-ésimo contenedor son f_i y f_{i+1} respectivamente. Comúnmente se utilizan en las aplicaciones biomédicas las bandas δ (0.5–4 Hz), θ (4–7

Hz), α (8–12 Hz), β (12–30 Hz), y γ (30–100 Hz). Con estos contenedores tenemos un vector $\{0.5, 4, 7, 12, 30, 1\}$ [20].

Entonces, la densidad espectral de potencia para el K -ésimo contenedor se evalúa como:

$$PSI_k = \sum_{[i=N(f_k/f_s)]}^{[N(f_{k+1}/f_s)]} |X_i|, \quad k = 1, 2, \dots, K - 1$$

Donde f_s es la frecuencia de muestreo y N el largo de la serie.

La intensidad relativa para este vector nos sirve para entender la distribución de la potencia, especialmente para caracterizar señales, se define como:

$$RIR_j = \frac{PSI_j}{\sum_{k=1}^{K-1} PSI_k}, \quad j = 1, 2, \dots, K - 1$$

6.7.1. Dimensión Fractal Petrosiana

Para una serie de tiempo, el coeficiente PDF se define como:

$$PFD_k = \frac{\log_{10} N}{\log_{10} N + \log_{10}(N/(N + 0.4N_\delta))}$$

Donde N es el largo de la serie y N_δ es el número de cambio de signos que tuvo la señal derivada [20]. Es similar al *Zero-Crossing* utilizado en [7].

Dimensión Fractal de Higuchi

Este algoritmo construye k nuevas series a partir de la serie original $\{x_1, x_2, \dots, x_N\}$ de la siguiente forma [20]:

$$x_m, x_{m+k}, x_{m+2k}, \dots, x_{m+[(N-m)/k]k}$$

Donde $m = 1, 2, \dots, k$

Para cada iteración en la construcción de la serie, el largo $L(m, k)$ es calculado como:

$$L(m, k) = \frac{\sum_{i=2}^{[(N-m)/k]} |x_{m+ik} - x_{m+(i-1)k}|}{[(N-m)/k]k}$$

El largo promedio es calculado cuando $L(k) = [\sum_{i=1}^k L(i, k)]/k$. Este procedimiento se repite k_{max} veces para cada k desde 1 hasta k_{max} , y luego se utiliza el método de mínimos cuadrados para determinar la pendiente que mejor se ajusta a $\ln(L(k))$ contra $\ln(1/k)$. La pendiente es el coeficiente de Higuchi [20].

Parámetros de Hjorth

Sea una serie de tiempo $\{x_1, x_2, \dots, x_N\}$, el coeficiente de movilidad y complejidad de Hjorth son respectivamente, definidos como

$$\text{Mobility} = \sqrt{M2/TP}$$

$$\text{Complexity} = \sqrt{(M4 \times TP)/(M2 \times M2)}$$

Donde $TP = \sum x_i/N$, $M2 = \sum d_i/N$, $M4 = \sum (d_i - d_{i-1})^2/N$ y $d_i = x_i - x_{i-1}$ [20].

Entropía SVD

Define una medida de entropía utilizando Descomposición por Valores Singulares (SVD). Sea $\{x_1, x_2, \dots, x_N\}$ una serie de tiempo de entrada [20]. Se construyen vectores con retraso de la siguiente forma:

$$Y = \{x_i, x_{i+\tau}, \dots, x_{i+(d_E-1)\tau}\}$$

Donde τ es el retraso y d_E es la dimensión de incrustación. Para esta aplicación se utilizó $d_E = 20$ y $\tau = 2$. El espacio de incrustación se define como:

$$Y = [y(1), y(2), \dots, y(N - (d_E - 1)\tau)]^T$$

El coeficiente SVD se ejecuta en la matriz Y para producir M valores singulares, $\sigma_1, \dots, \sigma_M$ conocidos como el espectro singular [20]. Se define la entropía SVD como:

$$H_{\text{SVD}} = - \sum_{i=1}^M \bar{\sigma} \log_2 \bar{\sigma}_i$$

Donde M es el número de valores singulares y $\sigma_1, \dots, \sigma_M$ son valores singulares normalizados tal que $\bar{\sigma}_i = \frac{\sigma_i}{\sum_{j=1}^M \sigma_j}$

Coefficiente de Información de Fisher

Para calcular el coeficiente de Información de Fisher se utilizan el espectro singular normalizado que se calcula en la entropía SVD [20] y se define como:

$$I = - \sum_{i=1}^{M-1} \frac{(\sigma_{i+1}^- - \bar{\sigma}_i)^2}{\bar{\sigma}_i}$$

Análisis de fluctuación sin tendencia

Sea $\{x_1, x_2, \dots, x_N\}$ una serie de tiempo, el procedimiento para calcular el coeficiente DFA se define de la siguiente forma [20]:

1. Integrar la serie de entrada a una nueva serie $y = \{y(1), \dots, y(N)\}$ donde $y(k) = \sum_{i=1}^k (x_i - \bar{x})$ y \bar{x} es la media de la serie de entrada.
2. La serie integrada luego se divide en segmentos de igual largo n , una línea determinada por mínimos cuadrados es ajustada por los datos, representando la tendencia en ese segmento. Las coordenadas y de la línea de tendencia son determinadas por $y_n(k)$.
3. La media cuadrática de la fluctuación de la serie integrada es calculada por $F(n) = \sqrt{(1/N) \sum_{k=1}^N [y(k) - y_n(k)]^2}$, donde $y(k) - y_n(k)$ es un factor para eliminar tendencia.
4. Las fluctuaciones son determinadas como la pendiente de $\log(F(n))$ para $\log(n)$.

6.8. Pulsos binaurales

Los pulsos binaurales han sido utilizados en investigaciones [2] buscando mejorar diversos procesos cognitivos como la creatividad, concentración y también para mejorar la calidad de sueño, en algunos sitios de internet [2] se le ha llamado una “droga digital” bajo el supuesto que induce un estado cognitivo y de humor especial activando patrones en el cerebro. Un pulso binaural es una ilusión auditiva que se activa cuando dos tonos a distintas frecuencias son presentados en los oídos, por ejemplo exponer un oído a un tono de 250 Hz y el otro a 260 Hz, por resultado obtenemos que el cerebro solo percibe un pulso binaural de 10 Hz que fluctua de oído a oído en lugar de dos a frecuencias más alta.

La forma exacta en la que el cerebro percibe estos pulsos aún está indefinido, pero parece que el sistema reticular de activación y el *colliculus* inferior juegan un rol principal [1]. En animales, el estímulo producido por pulsos binaurales ha mostrado una producción de patrones neuronales en ciertas frecuencias o una sincronización a la frecuencia diferenciada que percibe el cerebro. Cabe mencionar que las frecuencias que pueden presentar interacción e inducir sincronización con estos pulsos son solo aquellas que son detectables al oído del individuo y que también se presenten en sus ondas electroencefalográficas, se ha demostrado que en los humanos esta interacción ha inducido estados cognitivos y de ánimo. [3]

Si los pulsos binaurales tienen un efecto sobre los procesos cognitivos a través de una sincronización neuronal, es entonces importante considerar que dependiendo de la frecuencia a la que se exponga tendrá efecto en distintos patrones neuronales y por consiguiente en estados cognitivos. Por ejemplo, la comunicación de corto rango en las áreas del cerebro están asociadas a la sincronización con frecuencias gamma mientras que las comunicaciones de rango mayor están asociadas con frecuencias más lentas. Entonces tenemos “frecuencias mensajeras” para control cognitivo. Por ejemplo la sincronización en frecuencias gamma tienen un papel en el control de la consolidación de memoria y que es utilizada para personas que desear estudiar y obtener una mejor retención. Otro ejemplo interesante es la sincronización en las frecuencias de la banda alfa la cual presenta una excitación cortical produciendo estados de cognición que aumentan el rendimiento en tareas de pensamiento divergente, entonces podemos notar que en la banda alfa y gamma se mejora el rendimiento de tareas de creatividad [1].

6.9. Pruebas estadísticas

Los estudios estadísticos se pueden clasificar como experimentales u observacionales. En un estudio experimental se realiza un procedimiento para generar datos, desde aquí se define la variable de interés y luego se definen las variables que pueden ser de interés para el objetivo y como influyen sobre la variable de interés. En cambio en un estudio observacional los datos se obtienen a través de encuestas por muestreo y no por un procedimiento controlado y aunque siempre se consideran las mejores prácticas para realizar este análisis en este tipo de investigación no es posible el control riguroso que se tiene en un estudio experimental que se diseña desde cero. [21]

Para esta investigación se consideran tres tipos de diseños de experimentos: diseño completamente aleatorizado, diseño de bloques aleatorizado y experimento factorial. Para cada uno de ellos se precisa la herramienta estadística conocida como análisis de varianza (ANOVA, por su acrónimo en inglés) y con ella analizar los datos disponibles, cabe mencionar que esta herramienta estadística también se puede utilizar con datos obtenidos a partir de un estudio observacional. Por ejemplo, se verá que en un diseño de experimentos completamente aleatorizado este procedimiento también sirve para probar la igualdad de tres o más medias poblacionales cuando se obtienen los datos mediante un estudio observacional. [21]

Para los análisis de varianza en donde deseamos comparar si la media de una población es igual a la de otra tenemos una premisa principal y se refleja en el planteamiento de una hipótesis nula y una hipótesis alternativa de la siguiente forma:

$$H_o : \mu_1 = \mu_2 = \mu_3 \dots = \mu_k$$

$$H_a : \text{no todas las medias poblacionales son iguales}$$

Donde k es el número de poblaciones presentes en el estudio y μ su media poblacional

El criterio de rechazo es si se obtiene un valor-p menor que el nivel de significancia (usualmente de 0.05) [21]

Existen supuestos requeridos que debemos de considerar para que las pruebas de varianza puedan aplicar en un experimento:

- En cada población, la variable de respuesta está normalmente distribuida
- La varianza de la variable de respuesta, denotada como σ^2 , es la misma en todas las poblaciones
- Las observaciones deben ser independientes

6.9.1. Diseño completamente aleatorizado

En este tipo de experimento se seleccionan los niveles de un factor como variable a las cuales les vamos a llamar tratamientos y a cada una de las veces que ese experimento con los distintos niveles de tratamientos fue realizado le llamaremos replicación. Cabe mencionar que la asignación de tratamientos es de forma aleatoria para cada unidad experimental. Este método nos ofrece ventajas como la flexibilidad de poder seleccionar la cantidad de tratamientos y replicaciones siempre y cuando cumplamos con una cota mínima para que el experimento sea significativo además que a través del análisis de varianza podemos obtener de una forma rápida y precisa los resultados del experimento. [21]

6.9.2. Diseño de bloques aleatorizado

Este experimento es construido a partir de un diseño completamente aleatorizado pero esta vez se toma en consideración el error que puede inducir el orden de corrida del experimento así como también el error que inducen en las replicaciones asociado a la singularidad de cada individuo. Es un experimento más robusto puesto que al error se le extrajo los valores de los bloques (filas de datos o replicaciones). [21]

6.9.3. Experimento multifactorial

Los diseños antes mencionados solo permiten analizar y formular conclusiones estadísticas acerca de un factor. Sin embargo, existe la necesidad muchas veces de formular conclusiones acerca de más de una variable o factor, entonces el experimento factorial resuelve este problema siendo un diseño que permite obtener conclusiones simultáneas acerca de dos o más factores. Se utiliza el término factorial porque también se considera la interacción entre factores. [21]

Descarga, extracción y almacenamiento de datos

7.1. Cambios en versiones anteriores y consideraciones

En la fase anterior se estableció un primer acercamiento a un algoritmo diseñado para clasificar etapas de sueño con datos de la actividad electroencefalográfica de personas [7]. El clasificador en cuestión fue desarrollado en el lenguaje de programación Matlab y se hicieron dos versiones:

- Con datos públicos provenientes de la base de datos pública Sleep-EDFX [10].
- Con datos del investigador.

Originalmente el proyecto fue planteado para que el entrenamiento del clasificador fuera preferiblemente desarrollado con datos recolectados de pacientes, pero en esta investigación se plantea la utilización de una base de datos pública sumado a esfuerzos de investigadores de la Universidad del Valle para recolectar datos por las siguientes razones:

- La utilización de un acercamiento inter-sujetos nos permite captar la varianza que existe entre distintos tipos de individuos y así lograr una precisión en el clasificador que no dependa del sujeto a aplicar el tratamiento.
- La cantidad de datos necesaria para mejorar un modelo de clasificación es importante, debido a que no se cuenta con los recursos necesarios para armar un conjunto de datos lo suficientemente grande se refuerza la idea de sumar datos públicos a la investigación.

- Debido a la pandemia del COVID-19 y a las restricciones de movilidad y distanciamiento social es recomendado no exponer a pacientes a clínicas en donde más personas puedan asistir, motivado por priorizar la salud y bienestar de pacientes y el equipo investigador.

Para esta nueva fase del proyecto se escogió el lenguaje de programación Python en un entorno en línea llamado Google Colab, una plataforma en línea que cuenta con librerías para utilizar modelos ya diseñados, una fuerte comunidad en línea para soporte y herramientas de visualización para poder entender las métricas y resultados.

7.2. Obtención de datos de entrenamiento

La base de datos Sleep-EDFX [10] es una base de datos que recoge información de hasta 83 sujetos, teniendo dos noches por cada uno de los pacientes sanos. Esta base de datos fue escogida debido a que contiene una gran cantidad de información no solo de los datos de polisomnografías como tal sino que también aspectos de los experimentos como la edad y género de los pacientes, frecuencia de muestreo, canales de datos y la información más importante que son anotaciones de los especialistas donde determinan de forma técnica las etapas del sueño. La parte de tener los datos etiquetados nos permite utilizar modelos de aprendizaje automático supervisado.

7.2.1. Descarga de datos

La descarga de los datos fue un procedimiento bastante complicado, esto debido a que los datos que se utilizan como estándar en el estudio del sueño son de tipo EDF (*European Data Format* [22] o Formato de Datos Europeo). Este tipo de archivo tiene la característica que cuenta no solo con los datos de interés sino de anotaciones, frecuencia, fecha y nombre de los canales además que cuenta con funciones propias para este tipo de archivos que permiten sumar más fácil otros cuadros o la extracción de su información.

En el primer esfuerzo de descarga se intentó utilizar la herramienta de Matlab que fue implementada en [7] y si bien fue posible una descarga parcial de los datos, su instalación y utilización fue bastante complicada. La documentación por parte del desarrollador de la herramienta no cubre la totalidad de la instalación y es necesario modificar la herramienta para poder comprender el proceso. Se perdió entonces la idea de ser un módulo que ahorra tiempo.

Se desarrolló una nueva herramienta de descarga y procesamiento en el lenguaje de programación Python. Basada en la investigación [23] en donde se descargan los datos sin ninguna librería y luego se procesan sin la utilización de librerías más que para extraer las anotaciones. Este procedimiento no fue eficiente y se descartó por las siguientes razones:

- La base de datos sin procesar presenta un tamaño grande respecto a la información que aporta (10 GBs aproximadamente).
- Los datos se cargan a memoria por lo que reduce la velocidad y aumenta costo computacional.
- Los métodos no permiten recortar el tiempo de análisis para cada conjunto de datos de una forma eficiente.

Debido a las limitaciones mencionadas con anterioridad, se utilizó la librería MNE [24] para producir el modelo de descarga final.

7.2.2. Implementación final de herramienta de descarga

La librería MNE [24] nos provee de herramientas para el análisis de datos relacionados a eventos. La base de datos Sleep-EDFX [10] es una de las bases de datos que están disponibles para descargar y analizar a través de esta librería, especialmente porque sus datos son relacionados a eventos (etapas de sueño). Entre las principales operaciones que se pueden realizar son las siguientes.

- Descarga de datos.
- Cargar los datos a entorno sin cargar a memoria.
- Utilización eficiente de datos cuando se necesitan.
- Agrupación por épocas.
- Preprocesamiento de datos.
- Gráficas especializadas a datos biomédicos.
- Estimación de actividad cortical desde un sensor.

Descarga

La librería MNE-Python [24] nos provee una función llamada *Fetch Data* que se traduce a obtención de datos, la cual permite descargar la base de datos Sleep Physionet Dataset [10].

En esta base de datos se encontraron dos noches de sueño por cada sujeto con promedio de 45 MBs de tamaño por cada noche. Todos los pacientes cuentan con una condición de salud sana y cada grabación contiene los siguientes archivos:

- PSG.edf contiene la polisomnografía y los datos de cómo fue realizada.

- Hypnogram.edf contiene las anotaciones de expertos clasificando cada etapa de sueño; Se utilizaron épocas (*epochs*) de 30 segundos.

Es importante resaltar que los sujetos 39, 68, 69, 78 y 79 no están disponibles por lo que fue necesario eliminarlos en el proceso.

Análisis exploratorio de sujetos de estudio

Antes de usar los datos, se realizó un proceso de exploración de la base de datos y se encontraron los siguientes resultados respecto a las características de los sujetos:

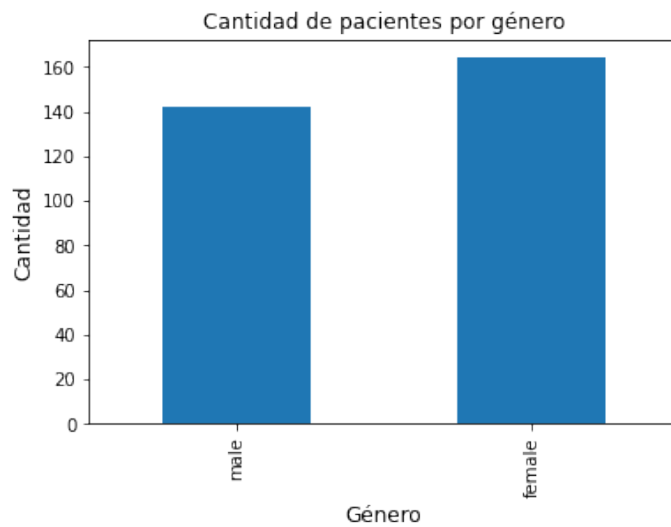


Figura 14: Distribución según género

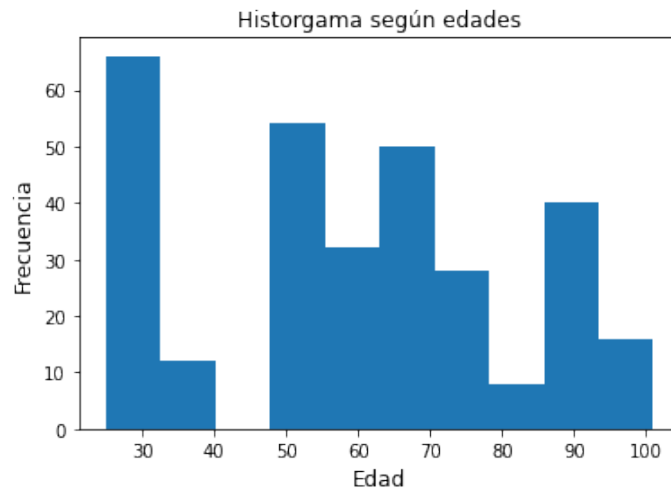


Figura 15: Distribución según edades

La Figura 14 ilustra que se tiene una distribución bastante equilibrada en lo que respecta a género de los sujetos de estudio. Para la edad se puede observar en la Figura 15 que existe una clara mayoría en los sujetos con una edad de 30 años, pero también de personas de la tercera edad. Esto añade variabilidad y robustez al modelo debido a que se consideran distintos casos cuando se miden pacientes externos a esta base de datos.

Mapeo de clases

Una de las primeras consideraciones de diseño fue mapear los eventos a números como se puede observar en el Cuadro 2:

Etapa	Clase
Despierto	1
Etapa 1	2
Etapa 2	3
Etapa 3	4
Etapa 4	4
REM	5

Cuadro 2: Diccionario de etapas de sueño a clases

Se puede observar que las etapas 3 y 4 fueron unidas en una misma clasificación, esto con la motivación que comparten características y en investigaciones previas [23] [24] se ha demostrado efectividad en el proceso de clasificación.

Luego de haber mapeado los valores se construyó un diccionario que mapea de regreso las etapas desde valores para una interpretación más fácil como se puede ver en el Cuadro 3.

Clase	Etapa
1	Despierto
2	Etapa 1
3	Etapa 2
4	Etapas 3/4
5	REM

Cuadro 3: Diccionario de unificación de clases

Preprocesamiento de datos

La carga de archivos fue sin duda uno de los procedimientos más importantes en esta parte de la investigación puesto que en esfuerzos pasados fue un cuello de botella en el preprocesamiento de los datos. La librería MNE [24] cuenta con funciones denominadas *read_raw_edf()* y *read_annotations()* que como su nombre indica se utiliza

para cargar los archivos de tipo EDF que contiene los cuadros de información y los archivos PSG que contiene las anotaciones. Es por ello que el parámetro que reciben es el nombre del archivo que contiene la noche de sueño del sujeto en cuestión. En el Cuadro 4 y 5 se puede notar la estructura de los datos crudos y además información sobre estos proveniente de la lectura del archivo.

Cabe resaltar que debido a que esta investigación se basa en un solo canal del electroencefalograma se eliminaron otros canales de sensores con los que no se cuenta. Los canales contienen la siguiente información:

- EOG
- Respiración nasal
- EMG de barbilla
- Temperatura

Índice	Tiempo (s)	EEG Fpz-Cz (μV)
0	0	5.016850
1	1	-2.578755
2	2	1.359707
3	3	-2.391209
4	4	-5.204396
...
7949995	7949995	11.299634
7949996	7949996	3.610256
7949997	7949997	5.298168
7949998	7949998	-4.641758
7949999	7949999	6.610989

Cuadro 4: Datos de ejemplo

Measurement date	April 24, 1989 16:13:00 GMT
Experimenter	Unknown
Participant	Unknown
Digitized points	Not available
Good channels	0 magnetometer, 0 gradiometer, and 7 EEG channels
Bad channels	
EOG channels	Not available
ECG channels	Not available
Sampling frequency	100.00 Hz
Highpass	0.50 Hz
Lowpass	100.00 Hz
Filenames	SC4001E0-PSG.edf
Duration	22:04:59 (HH:MM:SS)

Cuadro 5: Información de archivo EDF

7.2.3. Extracción de épocas

Se definen las épocas como la agrupación de datos en una serie de tiempo [25]. Para ello es importante denotar que en el estudio del sueño se utilizan épocas de 30 segundos y que dentro de cada una de estas épocas se encuentran los patrones necesarios para realizar la clasificación de la etapa de sueño.

Dentro de el conjunto de los datos se encontró un problema. Había un claro sesgo en la distribución de las clases orientado hacia la clase asociada al estar despierto puesto que en el conjunto de datos de Sleep-EDFX [10]. Se tomaron datos desde antes que el sujeto comenzara el proceso de sueño por lo que uno de los procesos principales fue delimitar el tiempo a solo media hora antes y media hora después del período de sueño.

Utilizando los datos provenientes de las anotaciones se procedió entonces a realizar el proceso de extracción de épocas que se puede ver reflejado en el Cuadro 6 donde se observa el tiempo, etapa del sueño, el número de época y los valores del canal de electroencefalograma en ese instante. Este proceso es de mucha importancia puesto que luego de aplicar las funciones para extracción de características se obtuvo el conjunto final de datos para entrenamiento.

Índice	Tiempo(s)	Etapa	Época	EEG Fpz-Cz (μV)
0	0	Sleep stage W	0	8.111355
1	1	Sleep stage W	0	17.488645
2	2	Sleep stage W	0	21.239560
3	3	Sleep stage W	0	6.892308
4	4	Sleep stage W	0	9.424176
...
2522995	29950	Sleep stage W	840	-17.863736
2522996	29960	Sleep stage W	840	-17.676190
2522997	29970	Sleep stage W	840	-17.582418
2522998	29980	Sleep stage W	840	-10.455678
2522999	29990	Sleep stage W	840	-13.456410

Cuadro 6: Datos con épocas

7.2.4. Agrupación y almacenamiento de datos

Una vez que ya se cuenta con los datos y su respectiva época se tiene la necesidad de agrupar toda la información posible de una época en los vectores de características. Se aplicaron funciones a las series de tiempo y se extrajeron a través de funciones de agrupación como lo es *Group By*. En esta función se itera sobre cada una de las épocas, se extrae la información en números escalares y se forma el vector final para caracterizar una época. Por último, se almacenaron los datos agrupados en archivos de *Excel* para su posterior análisis en el proceso de entrenamiento del clasificador.

7.2.5. Gráficos de ejemplo con herramienta de procesamiento

Dentro de las metas a lograr en el desarrollo de la herramienta de descarga y procesamiento fue agregar herramientas visuales para extraer información de los datos utilizando la librería MNE [24] con los datos obtenidos de la base de datos Sleep-EDFX [10].

La Figura 16 es un gráfico de distribución de eventos. Esta gráfica fue de suma importancia para poder determinar el desequilibrio que existía en la cantidad de clases y la necesidad de recortar la ventana de análisis en el período de sueño.

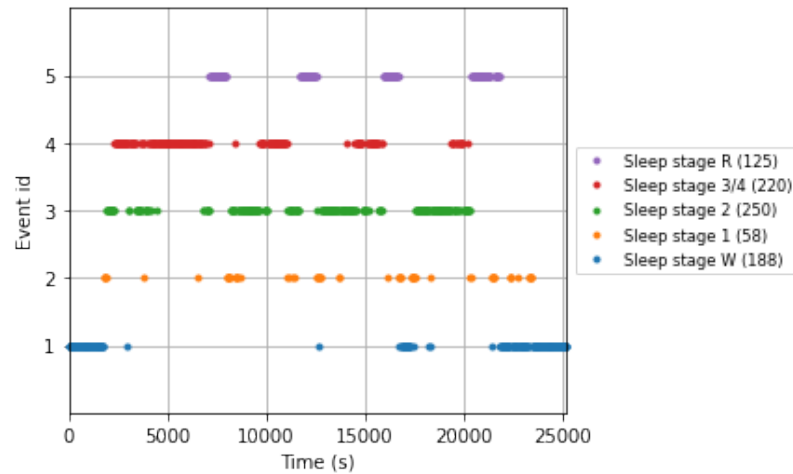


Figura 16: Gráfico de distribución de datos por clase.

Por otra parte, la Figura 17 es útil cuando se cuenta con los datos suficientes para visualizar una polisomnografía completa, en la barra de abajo se encuentra una codificación de colores para clasificar las etapas de sueño.

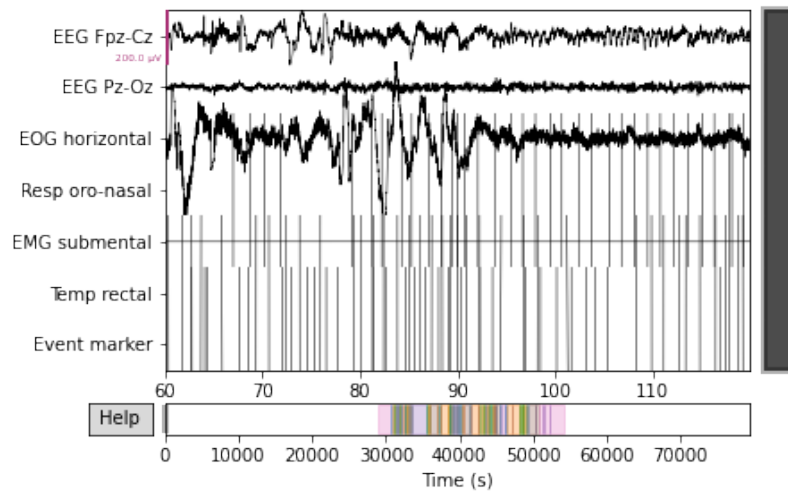


Figura 17: Gráfico de polisomnografía.

En las Figura 18 se puede observar en el eje x la duración en segundos de una época, en el eje y las épocas y en la codificación de colores el valor en micro voltios para cada segundo por época. Esta visualización es importante para entender la estructura de las épocas y como se relacionan entre sí.

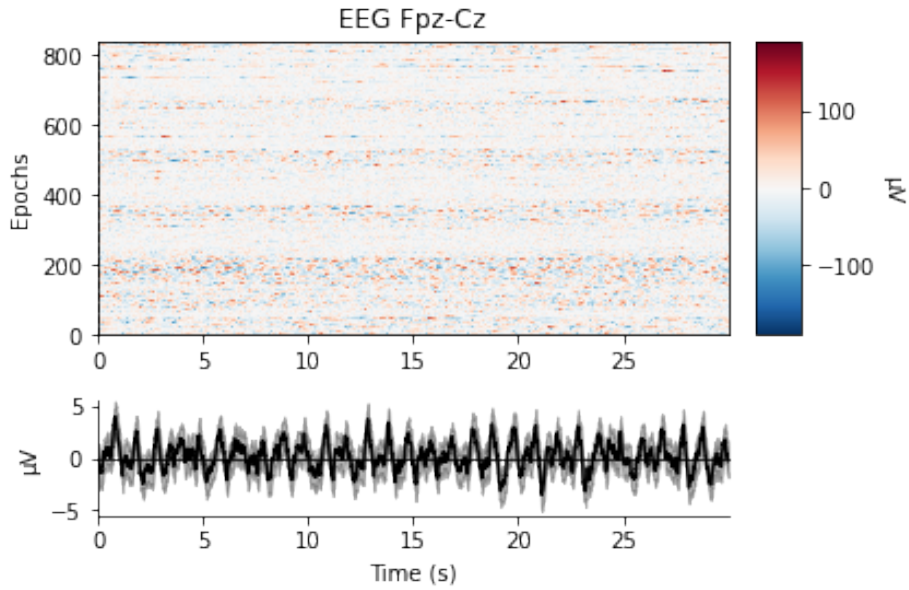


Figura 18: Gráfico de épocas en polisomnografía

Por último se observa en la Figura 19 un gráfico de densidad espectral de potencia, cada curva corresponde a una etapa de sueño. Este gráfico es importante puesto que esta característica es una de las más significativas para el clasificador.

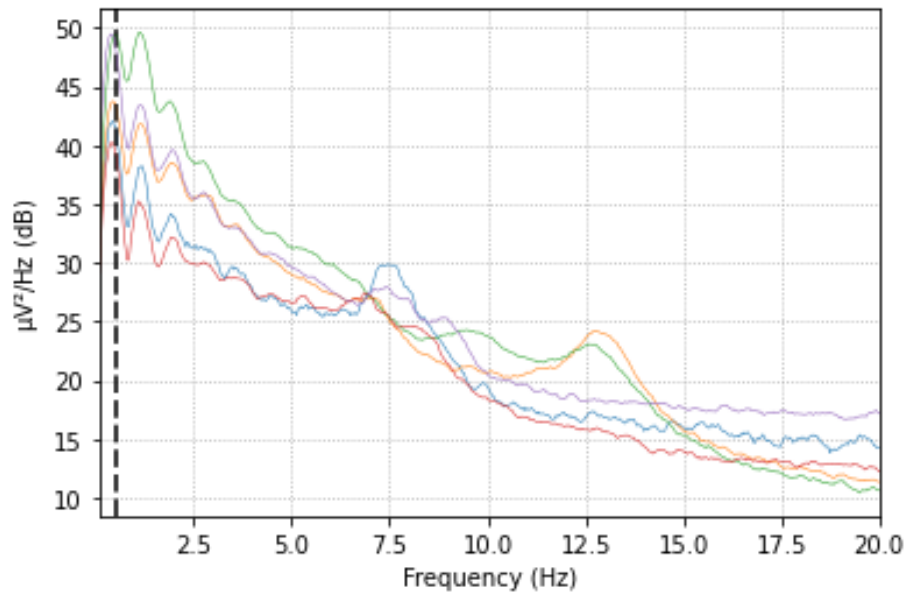


Figura 19: Gráfico de densidad espectral de potencia

8.1. Características de las señales

En la investigación [7] se planteó la utilización de las siguientes características para entrenamiento y clasificación:

- Mean Absolute Value
- Zero Crossing
- Max Min Distance
- Slow Wave Activity
- Kurtosis

Y se obtuvieron resultados prometedores más no un buen porcentaje promedio en la clasificación de cada clase según las matrices de confusión. Motivado con la mejora del modelo se realizó una investigación sobre las características utilizadas en investigaciones de otras universidades y se encontraron vectores y coeficientes que fueron fundamentales en este trabajo de graduación.

Estas características fueron derivadas de la librería *PyEEG* planteada en [20], una librería que contiene los cálculos necesarios para la extracción de características en señales o series de tiempo como podemos observar en el Cuadro 7.

Característica	Función	Tipo de dato
Densidad espectral de potencia	bin_power()	2 vectores de 1-D
Dimensión fractal Petrosiana (PFD)	pdf()	Escalar
Dimensión fractal de Higuchi (HFD)	hfd()	Escalar
Movilidad y complejidad de Hjorth	hjorth()	Dos escalares
Entropía SVD	svd_entropy()	Escalar
Coefficiente de información de Fisher	fisher_info()	Escalar
Análisis de fluctuación sin tendencia (DFA)	dfa()	Escalar

Cuadro 7: Características en *PyEEG* [20]

8.1.1. Selección de características de forma automática

La cantidad de características con las que se entrena un modelo no necesariamente debe de ser alta, sino más bien cada característica debe de aportar información que en otras no se encuentra. Motivado con la idea de encontrar la mejor combinación de característica se utilizó una técnica llamada *Random Feature Elimination* o (RFE) [26].

RFE es una técnica de eliminación de características en donde genera varios modelos y los entrena con distintas combinaciones de características para luego proponer la mejor combinación. En los resultados se encontró que para todos los modelos estudiados en esta investigación la mejor combinación incluye las siguientes características:

- Densidad espectral de potencia
- Dimension fractal de Higuchi
- Análisis de fluctuación sin tendencias

Esto concuerda con lo estudiado en otras investigaciones [7] y [15], en donde se asocia frecuencias específicas a las etapas del sueño. Las demás características se agregaron y mejoraron solo en pequeña parte la clasificación, pero de igual forma es una mejora a cuando se utilizan solo las tres que el optimizador indica.

8.2. Sobremuestreo de datos de entrenamiento

Debido a que los datos con los que se cuenta no tienen una distribución homogénea según las clases como se puede observar en la Figura 20

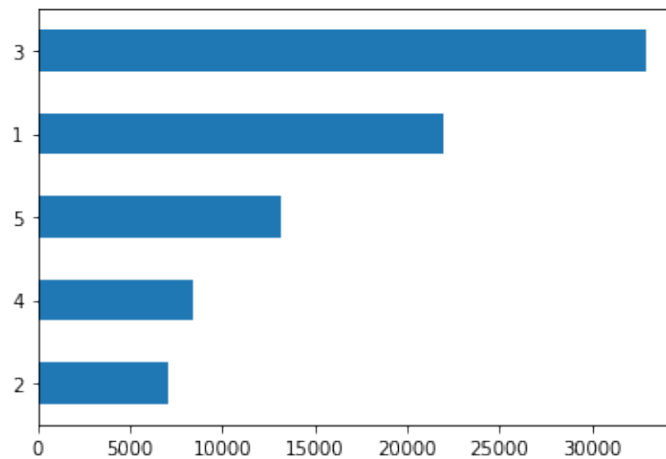


Figura 20: Distribución de datos por clase

Se procedió a utilizar una técnica de sobremuestreo llamada *Random Under-Sampling* de la librería *Imbalance Learning* [27] donde muestrea de forma aleatoria hasta conseguir una distribución de datos por clase igual al máximo de datos en la clase que menos datos tiene.

Si bien se creía que esto era necesario para mejorar la clasificación, los resultados mostraron lo contrario consiguiendo una precisión un 15% menor en la mayoría de iteraciones con respecto a dejar los datos iguales. Es posible que este fenómeno se deba a que la arquitectura del sueño tiene esa estructura.

8.3. Modelos de aprendizaje automático

En esta investigación se utilizaron modelos de aprendizaje automático debido a que se desea evaluar también, las mejores características para clasificar patrones en las etapas de sueño. Al diseñar con técnicas de aprendizaje automático profundo no se puede acceder a las características puesto que las redes determinan con convoluciones (Si se utiliza un modelo CNN o RNN) de forma automática lo que crean más convenientes llevando esto a una caja negra donde no se puede acceder a entender el comportamiento.

Se utilizaron dos modelos en esta investigación debido a su naturaleza y los resultados obtenidos en investigaciones similares, son el modelo *XGBoost* y el modelo de Bosques Aleatorios o *Random Forest*, para ambos modelos se utilizaron las características mencionadas en el Cuadro 7.

8.3.1. *eXtreme Gradient Boosting (XGBoost)*

XGBoost es la abreviación para el algoritmo *Extreme Gradient Boosting*, una implementación en donde se utiliza un *solver* que utiliza modelos lineales en conjunto con un algoritmo de aprendizaje por arboles. Se puede implementar con distintas funciones objetivos tales como clasificación, regresión y *ranking* [28]. Entre sus principales características podemos encontrar:

- La velocidad, este algoritmo puede realizar operaciones paralelas aumentando la velocidad hasta 10 veces respecto al tradicional *Gradient Boosting*.
- Soporta varios tipos de matrices de entrada permitiendo una apertura a varios tipos de aplicaciones.
- Funciona bien con bases de datos que presenten escasez en la cantidad de datos.
- Permite una personalización del algoritmo variando las funciones objetivo y sus parámetros.

Para este algoritmo se utilizaron los siguientes hiper parámetros:

- Función objetivo: Clasificación Múltiple con *Softmax*.
- Número de estimadores de 100.
- Métricas de evaluación *LogLoss* y error cuadrático en la versión multivariable para ambos.

Resultados

Métrica	Precisión	Exhaustividad	Valor-F
Precisión	0.717361	0.717361	0.717361
Promedio	0.641677	0.600380	0.602408
Promedio ponderado	0.691819	0.717361	0.695243

Cuadro 8: Métricas de precisión en algoritmo XGBoost.

Etapas	Precisión	Exhaustividad	Valor-F
Despierto	0.757357	0.853828	0.802705
Etapas 1	0.370690	0.091684	0.147009
Etapas 2	0.743385	0.839994	0.788742
Etapas 3/4	0.742475	0.658754	0.698113
REM	0.594476	0.557642	0.575470

Cuadro 9: Métricas generales de precisión por etapa en algoritmo XGBoost.

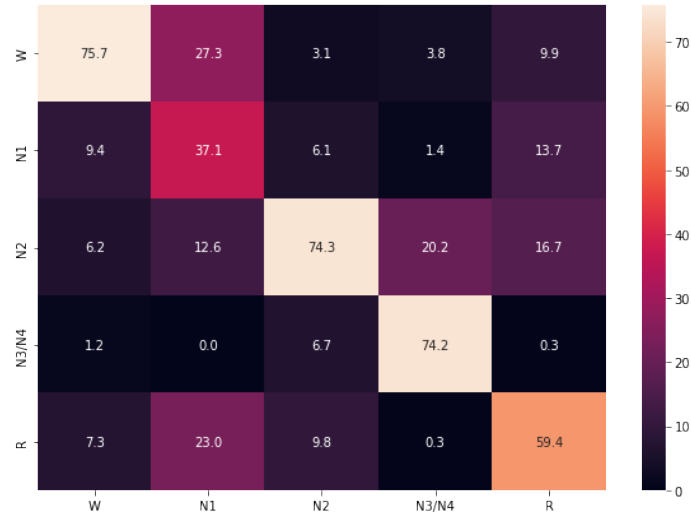


Figura 21: Matriz de confusión para algoritmo XGBoost.

Análisis

Hay un comportamiento bastante claro que se puede inferir de los Cuadros 8 y 9 y en la matriz de confusión en la Figura 21. Existe un problema en la clasificación de la etapa 2 debido a que presenta un comportamiento bastante parecido entre las clases REM y la clase Despierto. Estas dificultades en la clasificación originalmente se pensó que estarían asociadas con la falta de balance en la cantidad de clases en el conjunto de entrenamiento, pero luego de realizar operaciones de sobre muestreo para tener números más homogéneos la predicción disminuía en valores entre 10 % y 15 % respecto al conjunto original. Uno de los principales factores para mejorar esta clasificación sería implementar algún sensor que ayude a diferenciar la etapa 2 en específico o reforzar con alguna característica que resalte los atributos de la etapa 2.

8.3.2. Bosques aleatorios (*Random Forests*)

Random Forests es un algoritmo que combina la predicción de varios árboles de decisión. El valor que propone cada uno de estos árboles depende de el vector de entrada que fue entregado de forma aleatoria a cada uno de los elementos del algoritmo. En un trabajo de regresión el valor final propuesto por el bosque es producto de la media de los valores propuestos por cada árbol de decisión y en un trabajo de clasificación el valor es el que escogieron la mayor cantidad de árboles. El error de generalización aparece cuando la cantidad de árboles se vuelve grande [29].

Etapa	Precisión	Exhaustividad	Valor-F
Despierto	0.795549	0.887471	0.839000
Etapa 1	0.485126	0.150675	0.229935
Etapa 2	0.788538	0.859976	0.822709
Etapas 3/4	0.796156	0.688427	0.738383
REM	0.643478	0.666917	0.654988

Cuadro 11: Métricas generales de precisión por etapa en algoritmo Random Forest.

Resultados

Métrica	Precisión	Exhaustividad	Valor-F
Precisión	0.759345	0.759345	0.759345
Promedio	0.701770	0.650693	0.657003
Promedio ponderado	0.742480	0.759345	0.741820

Cuadro 10: Métricas de precisión en algoritmo Random Forest.

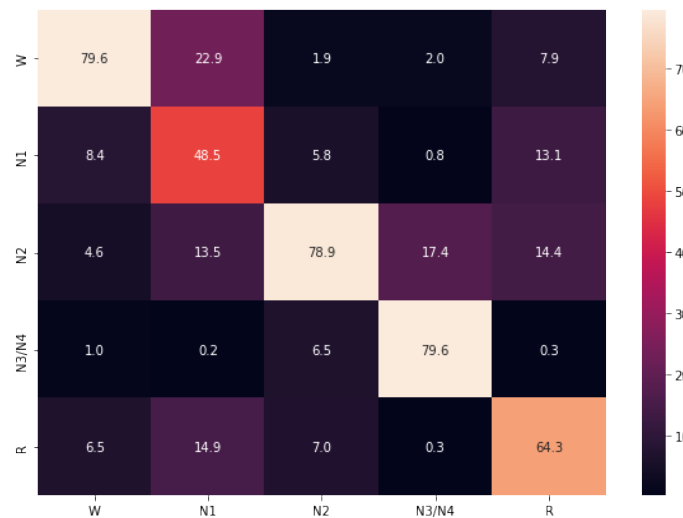


Figura 22: Matriz de confusión para algoritmo Random Forest.

Análisis

Para este algoritmo es observable que hay una mejora notable, en los Cuadros 10 y 11 podemos observar que existen mejores valores de precisión para la clase 2 y REM que fueron las que en el algoritmo *XGBoost* eran las que presentaban peores métricas de precisión. En la Figura 22 también podemos notar la mejoría. A raíz de esto se seleccionó este clasificador para implementarlo en la integración final.

9.1. Cambios en versiones anteriores y consideraciones

El módulo de pulsos binaurales ha experimentado cambios desde la primera versión planteada en [6]. Desde un principio se definió la generación en software de los pulsos, por eficiencia, optimización de calidad y evitar factores contraproducentes como ruido e interferencia y control de ganancia manual. En [8] se retoma el diseño del módulo pero esta vez con el deseo de migrar de dispositivo hacia una computadora Raspberry Pi. El proceso de migración fue motivado por la movilidad que ofrece un dispositivo de este tipo y la opción de tener una pantalla portátil en la cual se puedan escoger los valores que definen estos pulsos tales como la frecuencia central y el volumen de reproducción. En la Fig. 23 se puede observar el flujo de trabajo planteado en [8] para dicha migración y un ejemplo ilustrado de la interfaz gráfica propuesta en la Fig. 24.

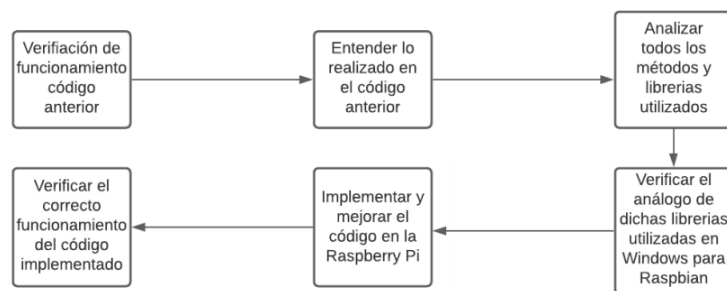


Figura 23: Flujo de migración a generador de pulsos binaurales en Raspberry Pi [8].



Figura 24: Interfaz gráfica de generador de pulsos binaurales [8].

En este trabajo se abordó una propuesta distinta en el módulo generador de pulsos binaurales puesto que se detectaron algunos inconvenientes en las propuestas pasadas. Uno de los principales problemas detectados en las investigaciones pasadas es la falta de una función que encapsula todas las características necesarias en la generación de pulsos binaurales, una función que reciba como parámetros las frecuencias centrales, tiempo de reproducción, tiempo de *crossfade* entre pulsos y volumen. Dicha función además debe de ser compatible con programación multihilos para poder recibir como parámetro la etapa de sueño y definir las frecuencias del pulso. En implementaciones que utilizan paralelismo es necesario ser eficientes en el tiempo de ejecución. Una interfaz gráfica podría representar un problema de funcionalidad. La interfaz gráfica es útil para diseñar pulsos y escucharlos en un paciente que empieza a probar los pulsos binaurales, en una implementación donde se desea evaluar su impacto no genera beneficios más que atracción visual sacrificando desempeño y funcionalidad.

En términos de compatibilidad, la Raspberry Pi al contar con una arquitectura de computador ARM, distinta de las computadoras comunes, presenta algunos problemas como la reproducción de audios generados por el módulo diseñado en [8]. Aunado a esto, existe una limitación en términos de funcionalidad, la Raspberry Pi no posee una buena interconexión inalámbrica con audífonos, limitando así a la utilización de audífonos alámbricos disminuyendo la comodidad y arriesgando a posibles pérdidas de conexión física durante la etapa de sueño.

Es por las razones antes mencionadas que en esta investigación se propuso una sola función generadora de pulsos binaurales implementada en software en el lenguaje de programación Python utilizando una computadora portátil. Función que recibe los parámetros necesarios para el diseño y producción continua de pulsos binaurales, funcional en desarrollo de programas multihilos y que además debido al tipo de entorno de reproducción es compatible con audífonos inalámbricos otorgando comodidad al paciente.

9.2. Módulo generador de pulsos binaurales

Desarrollar este módulo fue un proceso de continua investigación para encontrar la forma más rápida y eficiente de generar pulsos con características adecuadas a cada paciente. Para el pulso binaural se planteó utilizar las funciones antes desarrolladas pero carecían de la característica *crossfade*, esta aporta el sentido del desvanecimiento del audio al principio y al final aportando una transición más suave hacia el siguiente pulso manteniendo una sensación de continuidad.

Para la generación del pulso binaural se utilizó la librería de Python *AccelBrainBeat* [30]. Utilizando objetos de tipo *BinauralBeat* se crea el pulso genérico y luego implementando la función nativa de la librería *Save Beat* se pasan los parámetros deseados como las frecuencias centrales, el tiempo de reproducción y el volumen para el pulso específico.

Al pulso original se añade una pista portadora, esto debido a que durante un período de sueño es de suma importancia considerar la comodidad de una persona para optimizar el descanso. Es por ello que las frecuencias centrales y el volumen del pulso juegan un papel importante pero la reproducción sola de los pulsos puede parecer un poco tediosa por su patrón repetitivo. En esta investigación la solución propuesta fue añadir una pista portadora al pulso binaural original que como se puede observar en la Fig. 25 y 26 poseen un contenido frecuencia definido y que al añadir esta pista portadora para modular, como se puede observar en la Fig. 27 y 28 no añade contenido frecuencial importante al audio aportando solamente la comodidad en el proceso de escucha.

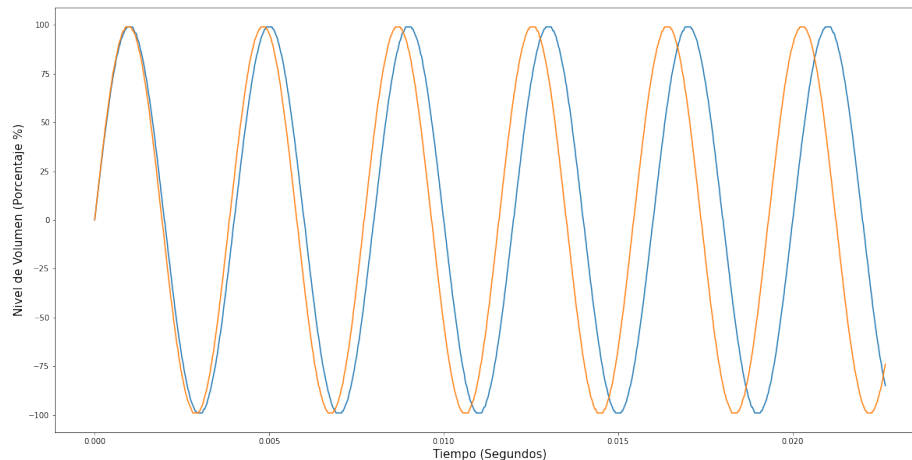


Figura 25: Pulso binaural de 12 Hz.

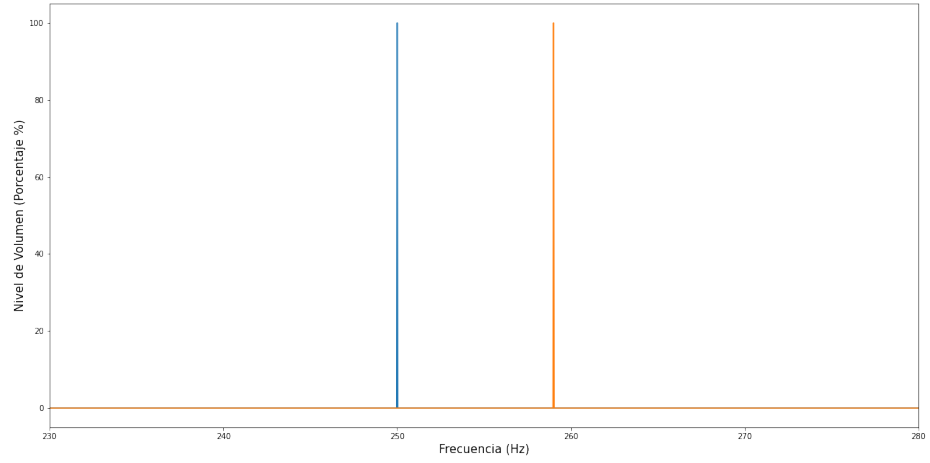


Figura 26: Transformada de Fourier de pulso binaural de 12 Hz.

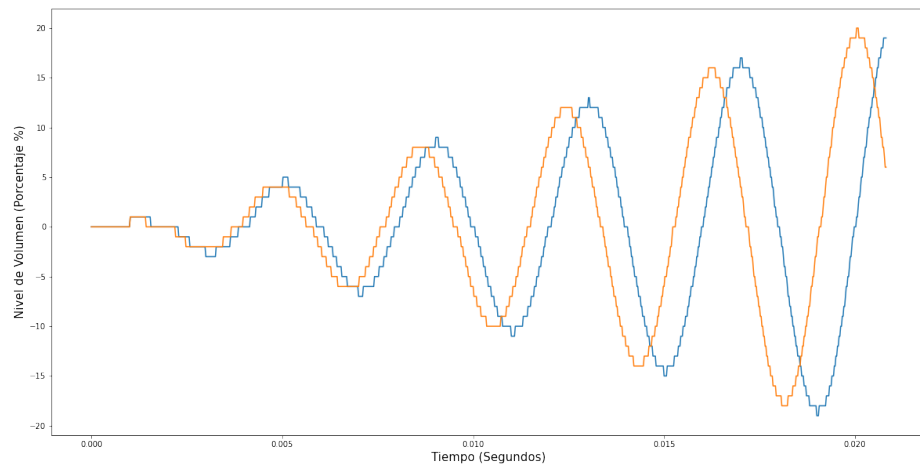


Figura 27: Pulso binaural modulado con pista portadora.

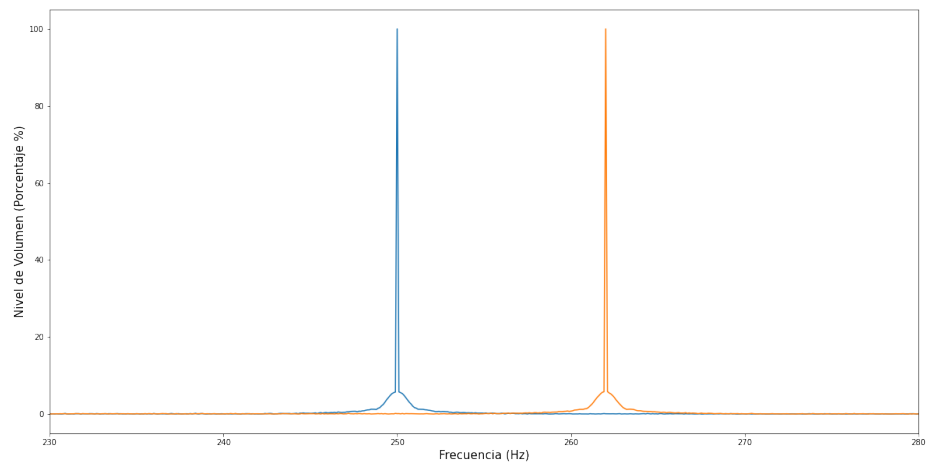


Figura 28: Transformada de Fourier de pulso binaural modulado con pista portadora.

El proceso de modulación y *crossfading* para añadir transiciones suaves fue realizado gracias a la librería *PyDub* [31] la cual permite realizar modificaciones de audio como sumar, recortar y añadir contenido externo, el flujo de creación de esta función integradora se puede apreciar en la Fig. 29.

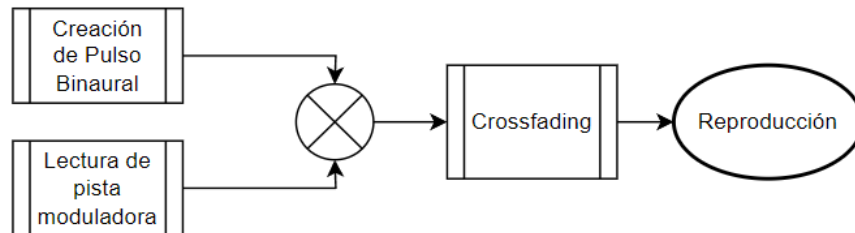


Figura 29: Proceso de creación de pulso binaural.

10.1. Cambios en versiones anteriores y consideraciones

La línea de investigación original plantea un sistema integrado de generación de pulsos binaurales retroalimentado según la etapa de sueño en la que se encuentra el paciente en cuestión. Como se observa en la Fig. 30, en un principio se esperaba que el sistema recolectara datos electroencefalográficos del *Electro-Cap* a través de la *Cython Board*, se procesaran en un lenguaje de programación para su clasificación con un modelo de inteligencia artificial para luego producir un pulso binaural que fuese escuchado por el paciente.

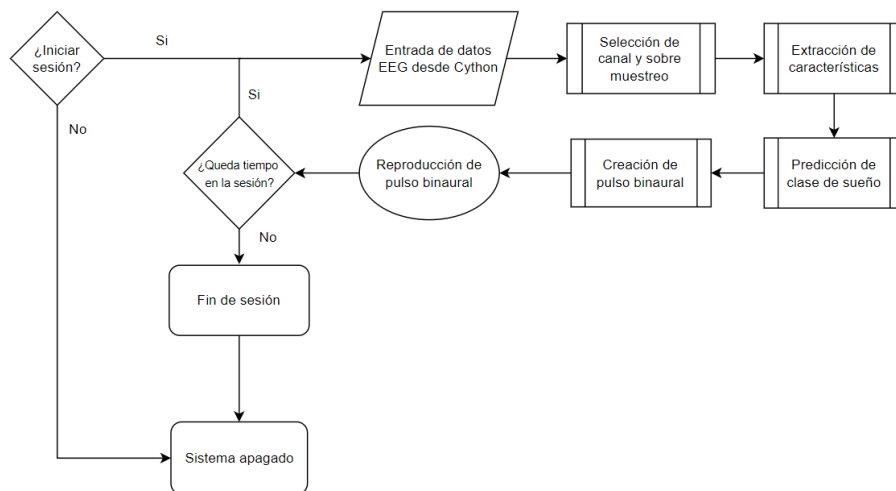


Figura 30: Sistema integrado de neuro-retroalimentación.

Esta arquitectura del sistema integrado fue desarrollado en esta investigación pero se encontraron errores de diseño a nivel del flujo de procedimientos. Una arquitectura de este tipo permite retroalimentar con pulsos binaurales la frecuencia de la misma etapa en la que se encuentra el paciente pero no seguir un flujo de sueño, aquí el problema fundamental de un proceso cíclico sin fin como se puede observar en la Fig. 31 si el impacto de los pulsos binaurales es válido como se espera. La solución al sistema se explora en la sección 10.4 utilizando un plan de sueño.

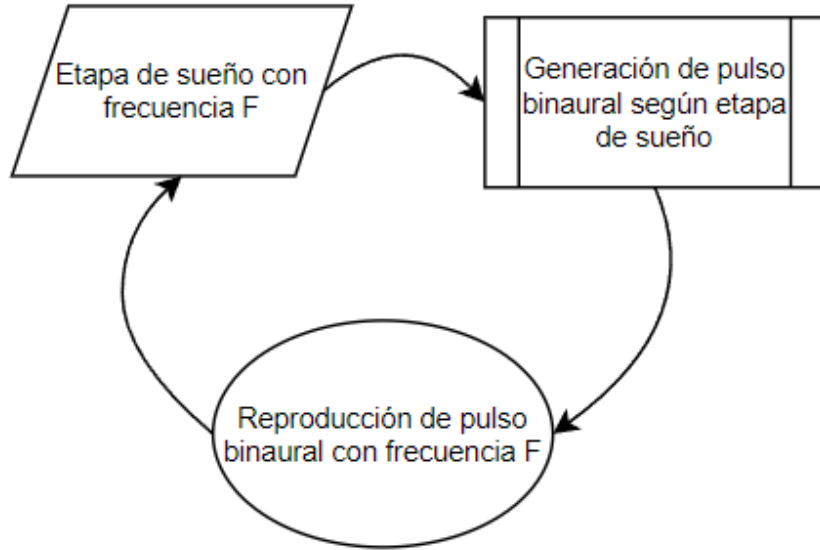


Figura 31: Retroalimentación en ciclo infinito.

10.2. Obtención de datos electroencefalográficos

10.2.1. Consideraciones y configuración de ambiente

La obtención de los datos bioeléctricos sigue un proceso como se observa la Fig. 30 en donde se recogen las señales eléctricas con los electrodos del *Electro-Cap* utilizando como *hardware* de procesamiento de señales la *Cython Board*. El lenguaje de programación y ambiente donde se procesan los datos en *software* ha variado puesto que en [6] se utilizó Python con la librería *PyLSL* en la cuál se construye un flujo de información a nivel de la capa de transporte. En la investigación siguiente [7] se utiliza también *PyLSL*, pero esta vez terminando el análisis en MATLAB por lo que se utilizaron otro tipo de conexiones en *software*.

Para este trabajo se utilizó el lenguaje de programación Python desarrollado en un entorno de Anaconda utilizando *Jupyter Notebook* con la librería *BrainFlow* [32],

la cuál nos permite la interconexión con una variedad de dispositivos de adquisición de señales bioeléctricas.

10.2.2. Librería BrainFlow

La interconexión de dispositivos, las dimensionales y el filtrado son algunos de los principales problemas en la obtención de datos bioeléctricos. Para esto, la librería *BrainFlow* (Desarrollada en colaboración con *OpenBCI*) posee varios módulos para la conexión con dispositivos de adquisición de datos, filtrado y clasificación como se puede observar en la Fig. 32. En este trabajo se utiliza la librería para la conexión con la *Cython Board*, esto nos ofrece ventajas sobre la conexión con *PyLSL* puesto que entrega los datos en micro voltios por segundo para luego ser utilizados en el preprocesamiento y posteriormente en el proceso de clasificación.

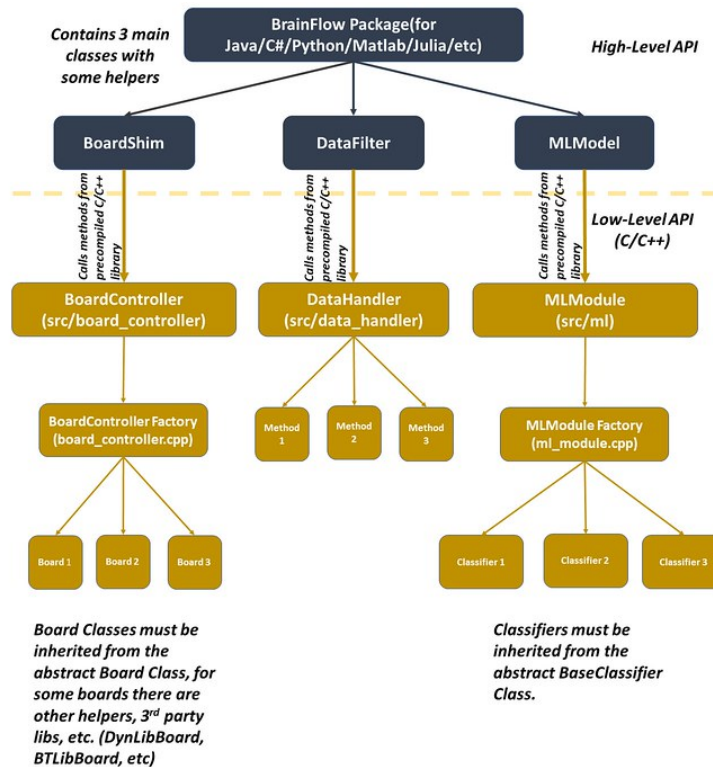


Figura 32: Módulos de *BrainFlow*.

10.3. Extracción de características y clasificación

Los datos recolectados por la etapa de adquisición de señales con *BrainFlow* son procesados, se les calcula las características y se le pasan al clasificador. Como fue descrito en la subsección 8.3.2, el clasificador de Bosques Aleatorios presentó el mejor desempeño en la tarea de clasificación frente a los otros modelos por lo que se utilizó la versión del clasificador que entregó los mejores valores de precisión entrenado con las siguientes características:

- Densidad Espectral de Potencia
- Dimension Fractal de Higuchi
- Análisis de Fluctuación sin tendencias

Las características de potencia nos ayudan a clasificar los estados macro por ejemplo, REM, NREM y Despierto, las fluctuaciones para los patrones de las ondas y el coeficiente de Dimensión Fractal de Higuchi está correlacionado con las etapas de sueño como podemos observar en la Fig. 33.

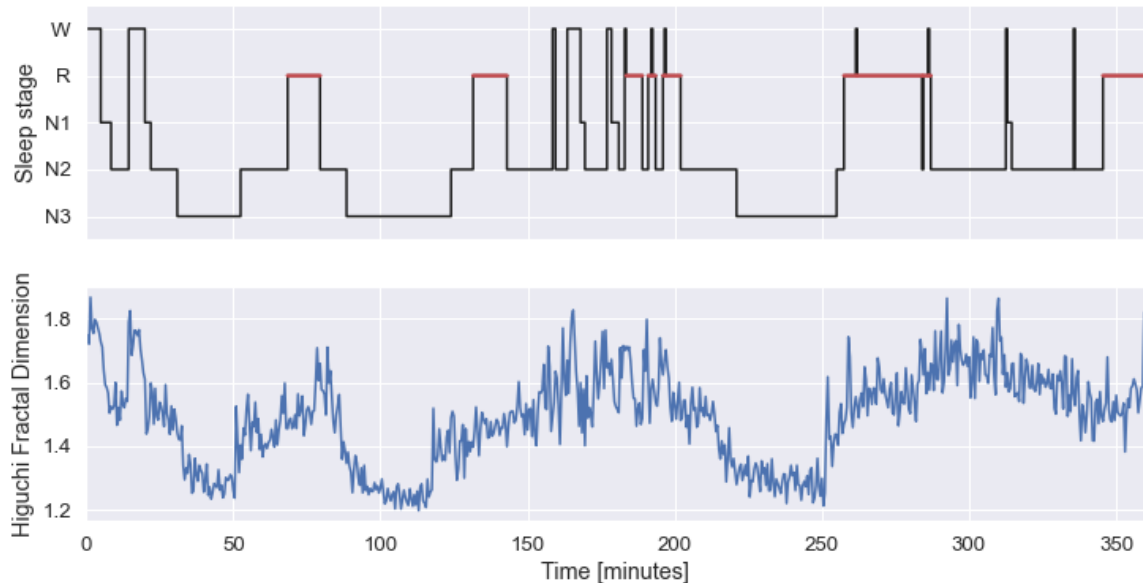


Figura 33: Relación de coeficiente de Higuchi con etapas de sueño [33].

10.4. Arquitectura de sistema con plan de sueño

Debido a que existe un problema en la arquitectura de la investigación original en el flujo de la información, parte de este trabajo fue replantear la dirección en la que fluye la información en el sistema. Debido a que la finalidad de los pulsos

Etapa	Rango de duración (Minutos)	Duración en plan de sueño (Minutos)
Despierto	Indefinido	5
N1	1-5	5
N2	25-90	30
N3/N4	30-60	50
REM	10-60	15

Cuadro 12: Distribución del sueño [34]

binaurales es inducir un sueño reparador, se realizó una investigación respecto a la distribución ideal de sueño. En un artículo relacionado a la fisiología del sueño [34], el autor describe algunos rangos de tiempo para cada etapa de sueño y como estos se van expandiendo conforme avanza el tiempo, los valores seleccionados para el plan de sueño se puede observar en la Tabla 12 siguiendo la distribución observable en la Fig. 36.

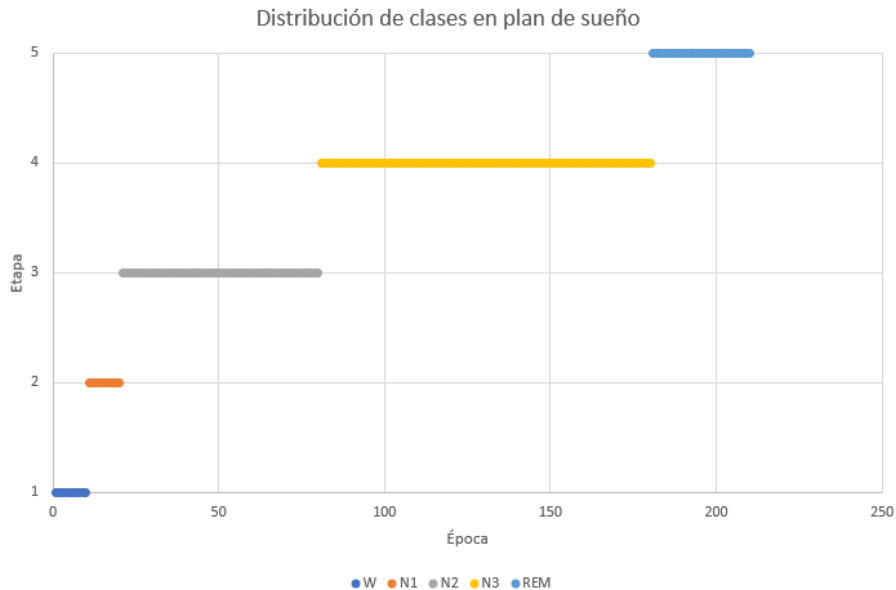


Figura 34: Distribución de etapas de sueño en plan de sueño.

10.5. Almacenamiento de datos

Como parte final del flujo del sistema integrado y pensando en el proceso continuo de mejora para el sistema, cada vez que se realiza una sesión de ejecución para el plan de sueño, se define un sistema de almacenamiento de datos utilizando la librería de Python *Pandas* la cual nos ofrece la utilidad de almacenar los datos sin procesar en formato CSV para su posterior análisis.

Diseño de estudio estadístico y documentos protocolarios

La eficacia de un nuevo tratamiento o procedimiento rara vez puede ser evaluado sin recurrir a un ensayo controlado. La mayoría de estos experimentos en los últimos 25 años se ha realizado a través de la asignación aleatoria de unidades experimentales a distintos grupos que cumplen con el perfil que el estudio requiere. La evaluación de la respuesta al tratamiento que recibe una unidad experimental o grupo y la comparación final de las respuestas promedio entre los otros grupos incluido en este el grupo control. [35]

A esta unidad experimental se le asigna un paciente, es aquí donde existe una divergencia entre la técnica de investigación a utilizar puesto que existen investigaciones intra-sujeto y entre-sujetos. La diferencia radica en que si es una investigación intra-sujeto la unidad experimental va a cambiar en cada iteración del tratamiento mientras que en una investigación entre-sujetos cada unidad experimental es siempre asignada a un paciente o sujeto.

En colaboración con la dirección de maestría en neuropsicología de la Universidad del Valle se inició un proceso de diseño de experimento tomando en cuenta sus debidas consideraciones éticas. En los documentos realizados para esta investigación se definen los perfiles que no son incluidos en esta investigación tales como mujeres embarazadas, niños y privados de libertad.

11.1. Consentimiento informado

La aceptación o el rechazo de un tratamiento es una manifestación particular del ejercicio de la autonomía en el paciente en cuestión. El consentimiento a la indicación

del procedimiento a realizar se hace sobre la base de la información que posee el paciente en relación con su enfermedad, pronóstico y opciones de tratamiento. En este punto de diseño se plantea que cantidad de información debe de saber un paciente respecto al procedimiento. La respuesta dependerá del marco ético en que se mueva el equipo que conduce la investigación. Si la conducta se rige por el principio de proporcionar al paciente el máximo beneficio, retendrá información si considera que esta provocará angustia, depresión o cualquier reacción adversa. Por el contrario, si las acciones del equipo investigador giran alrededor del respeto por la autonomía del paciente, le proporcionará toda la información necesaria antes de tomar una decisión. [36]

Para este trabajo, se realizó una investigación sobre las mejores prácticas éticas utilizadas en el diseño de un experimento que involucra pacientes humanos. Para ello se realizó un proceso de certificación en la universidad de Macquaire en Sydney, Australia. Dentro de esta certificación llamada *Human Research Ethics for the Social Sciences and Humanities* se abordaron los temas:

- Introducción
- Historia de las violaciones y regulaciones de la ética de la investigación
- Principios básicos de la práctica de la investigación ética
- 4 estudios de casos de ética en la investigación de las ciencias sociales
- Consideraciones éticas de la investigación para la investigación etnográfica
- Investigando en comunidades aborígenes australianas y de las islas del Estrecho de Torres
- Examen

Por último, se abordaron ejemplos enviados por la dirección de neuropsicología y este documento final se adjunta en los anexos de este trabajo.

11.2. Protocolo de investigación

Parte de los trabajos realizados en el fortalecimiento de las buenas prácticas éticas en esta investigación se realizó un protocolo de investigación para poder explicar de una forma menos técnica y hace más referencia a los procedimientos a realizar. Para el documento se trabajaron requerimientos planteados por el comité ético de la Universidad del Valle que toca los siguientes puntos.

Objetivos del estudio

- ¿Ha sido revisado por un IRB extranjero? (Adjuntar documento)

- Describir los procedimientos (naturaleza, secuencia y frecuencia) que se realizarán
- Describir los métodos de reclutamiento, criterios de inclusión y manejo de la información obtenida de los y las participantes
- Discutir riesgos y beneficios para los y las participantes
- Describir de qué forma se realizará el monitoreo y supervisión de la investigación por parte del investigador o investigadora principal

Cronograma y procedimientos

- Plan de análisis de los datos
- Fuentes de financiamiento y presupuesto si lo hubiera
- Declaración de conflicto de intereses
- Autorías
- Referencias
- Adjuntar los instrumentos que se utilizaran para el estudio

Documentos protocolarios

- Cartas de permiso de las instituciones, organismos o autoridades representantes de la comunidad donde se reclutará a los y las participantes
- Consentimiento informado y/o asentimiento informado (en caso sean menores de edad)
- Llenar el formulario en línea
- CV y certificados de cursos de ética en investigación de los investigadores e investigadoras principales y de los y las asistentes que tengan contacto con los y las participantes del estudio.

Estos documentos aún están en revisión por el equipo puesto que se debe de determinar y fijar un procedimiento específico para poder delimitar el tratamiento final para realizar la investigación.

El comité de ética establece la siguiente información una vez se envían los documentos:

“Después de enviar los comentarios y sugerencias el investigador y la investigadora debe hacer los cambios al protocolo y presentarlos de nuevo a la secretaria del comité.

Al ser revisados por la secretaria, esta dará su consentimiento de aprobación y se lo notificará a la presidenta, quien firmará una carta de aprobación, redactada por la secretaria.

Una vez sea autorizado un protocolo de investigación el investigador o la investigadora deberá apegarse a lo establecido en el protocolo y enviar notificación al Comité una vez sea finalizado el proyecto de investigación.

El Comité de Ética en Investigación de la Facultad de Ciencias Sociales debe revisar y aprobar todos protocolos de investigación que involucren seres humanos por lo menos una vez al año. Si el proyecto dura más de un año, es responsabilidad de el investigador o la investigadora enviar al Comité el protocolo de investigación para revisión y solicitud de aprobación de continuación del estudio.”

11.3. Evaluación inicial de impacto

Uno de los objetivos de la investigación es el diseño de un estudio que pueda medir el impacto del tratamiento con neuroretroalimentación de pulsos binaurales, para ello se utilizó la estructura original de refuerzo del estado neuronal de la etapa de sueño, para ello se utilizó la herramienta de medición utilizada en inteligencia artificial la matriz de confusión y se lograron a través de la recolección de datos de dos noches de sueño con dos personas con las siguientes características:

Sujeto 1: ´

- Género: Masculino
- Edad: 22
- Profesión: Estudiante

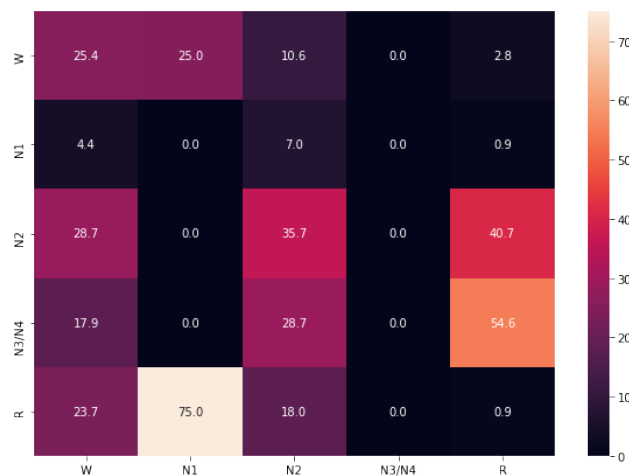


Figura 35: Impacto en sueño de terapia en sujeto 1.

Sujeto 2: ´

- Género: Masculino
- Edad: 25
- Profesión: Estudiante

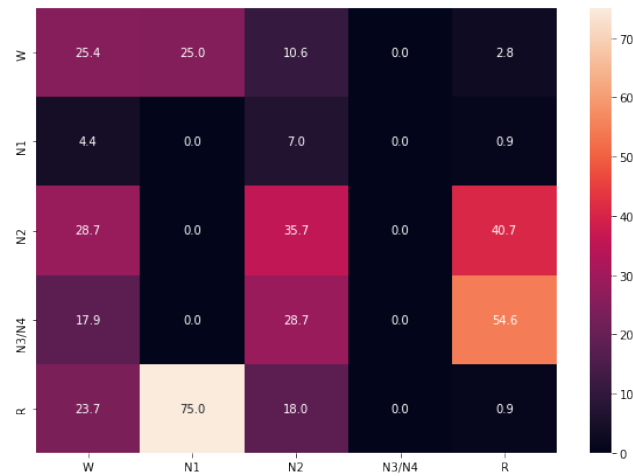


Figura 36: Impacto de sueño de terapia en sujeto 2.

Análisis

Uno de los principales descubrimientos al analizar los datos obtenidos a través del procedimiento es que existe una confusión y poca inducción en las etapas de sueño. El mejor caso fue la inducción del estado NREM 2, luego de ello fue el estado de despierto al inicio de cada uno de los períodos. Uno de los factores principales que puede afectar es el escalamiento de los datos, utilizar un preprocesamiento escalando los datos a mínimos y máximos o a través de la distribución como media y desviación estándar. Otro de los factores que pudo haber influido es la utilización de un único canal en la investigación o la falta de herramientas físicas como electroculogramas o electromiogramas de mentón. Este impacto aún no es definitivo, esto es la primera etapa de un proceso de validación con un grupo control utilizando un modelo final pero el diseño del estudio es lo importante para poder replicar.

- Se realizó una mejora considerable en el clasificador no solo en la precisión para diferenciar las etapas de sueño sino también en el cambio hacia una metodología entre-sujetos que nos añade variabilidad y robustez al variar la edad y el género del paciente.
- Se desarrolló un conjunto de herramientas para la descarga, procesamiento, extracción de características y visualización de datos electroencefalográficos. Se puede acceder a la herramienta en un navegador web sin procesos de instalación en computadora local.
- Se renovó el sistema generador de pulsos binaurales para la reproducción desde la computadora con audífonos inalámbricos y una pista moduladora de los pulsos binaurales para mayor estímulo del sueño, añadiendo también comodidad al paciente.
- Se dio inicio a un proceso de validación con el comité de ética de la Universidad del Valle de Guatemala, estableciendo un protocolo de investigación y un consentimiento informado para la ejecución de experimentos con pacientes.
- Se integró de forma exitosa el módulo clasificador y el módulo generador de pulsos binaurales utilizando programación multihilos obteniendo un buen desempeño en el sistema integrado.
- El estudio del impacto de los pulsos binaurales debe ser robusto tanto en el procedimiento de clasificación automática, el preprocesamiento de los datos y como interpretarlo puesto que los resultados se prestan a una conclusión confusa.

- Se recomienda para una futura fase la integración de un módulo de EMG para barbilla, esto con el motivo de mejorar los procesos de clasificación especialmente para etapa 2 del sueño NREM que fue el que presentó una clasificación más complicada.
- En una futura implementación sería interesante poder determinar si es posible inducir Complejos-K y husillos de sueño para inducir la etapa 2 del sueño NREM y ondas diente de sierra a través de pulsos binaurales para inducir la etapa REM.
- La investigación con potenciales evocados es clave en el análisis de neuro-retroalimentación por lo que sería bueno poder reproducir dos noches de sueño teniendo una como grupo control sin los pulsos y otra con la escucha de los pulsos binaurales.
- La reclutación de sujetos para el estudio debe de ser un proceso acompañado con la dirección de la maestría en neuropsicología. Esto con motivo de seleccionar no solo en base a variables sociodemográficos sino que también en la parte conductual de la psicología.
- Para llegar a una conclusión robusta el sistema debe de incluir sensores de EOG y EMG, sumado a un sistema de clasificación de árbol utilizando múltiples canales. [37] [38] [39]

-
- [1] S. Kuwada, T. C. Yin y R. E. Wickesberg, "Response of cat inferior colliculus neurons to binaural beat stimuli: possible mechanisms for sound localization," *Science*, vol. 206, n.º 4418, págs. 586-588, 1979.
 - [2] S. A. Reedijk, A. Bolders y B. Hommel, "The impact of binaural beats on creativity," *Frontiers in human neuroscience*, vol. 7, pág. 786, 2013.
 - [3] J. D. Lane, S. J. Kasian, J. E. Owens y G. R. Marsh, "Binaural auditory beats affect vigilance performance and mood," *Physiology & behavior*, vol. 63, n.º 2, págs. 249-252, 1998.
 - [4] J. Aguirre, "Diseño, Análisis, y Desarrollo de un Sistema de Entrenamiento para Mejorar el Desempeño de los Atletas del Comité Olímpico Guatemalteco," Tesis de Licenciatura, Universidad del Valle de Guatemala, 2018.
 - [5] M. Godoy, "Sistema de Neurofeedback para mejorar el rendimiento de los Atletas del Comité Olímpico Guatemalteco. Diseño e implementación de módulo de recopilación de señales y módulo de retroalimentación," Tesis de Licenciatura, Universidad del Valle de Guatemala, 2018.
 - [6] J. Muñoz, "Diseño de un sistema inteligente de monitoreo de ondas EEG y generador de pulsos binaurales para combatir desórdenes de sueño en los atletas," Tesis de Licenciatura, Universidad del Valle de Guatemala, 2019.
 - [7] O. López, "Análisis y Reconocimiento de Patrones de Señales Biomédicas Relacionadas a las Etapas de Sueño," Tesis de Licenciatura, Universidad del Valle de Guatemala, 2020.
 - [8] L. Guerrero, "Diseño e Implementación de un Sistema Generador de Pulsos Binaurales para el Estudio del Impacto de los Pulsos en la Calidad de Sueño," Tesis de Licenciatura, Universidad del Valle de Guatemala, 2020.
 - [9] F. S. Luyster, P. J. Strollo, P. C. Zee y J. K. Walsh, "Sleep: a health imperative," *Sleep*, vol. 35, n.º 6, págs. 727-734, 2012.

- [10] S. A. Imtiaz y E. Rodriguez-Villegas, "An open-source toolbox for standardized use of PhysioNet Sleep EDF Expanded Database," en *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2015, págs. 6014-6017.
- [11] M. M. Rahman, M. I. H. Bhuiyan y A. R. Hassan, "Sleep stage classification using single-channel EOG," *Computers in biology and medicine*, vol. 102, págs. 211-220, 2018.
- [12] C. A. Vinueza Bonilla y col., "Sistema para monitoreo de ondas cerebrales en estudios de pulsos binaurales con ritmo theta sobre los procesos cognitivos y emocionales," B.S. thesis, Quito: USFQ, 2016, 2016.
- [13] M. G. Melissa Vollrath, *The McGill Physiology Virtual Laboratory*, https://www.medicine.mcgill.ca/physio/vlab/biomed_signals/eeg_n.htm#:~:text=The%20recorded%20waveforms%20reflect%20the,amplitude%20and%20the%20slowest%20waves., [Online; accessed 30-March-2020], 2005.
- [14] J. LaRocco, M. D. Le y D.-G. Paeng, "A systemic review of available low-cost EEG headsets used for drowsiness detection," *Frontiers in neuroinformatics*, vol. 14, 2020.
- [15] B. M. Altevogt, H. R. Colten y col., "Sleep disorders and sleep deprivation: an unmet public health problem," 2006.
- [16] A. Dey, "Machine learning algorithms: a review," *International Journal of Computer Science and Information Technologies*, vol. 7, n.º 3, págs. 1174-1179, 2016.
- [17] T. O. Ayodele, "Types of machine learning algorithms," *New advances in machine learning*, vol. 3, págs. 19-48, 2010.
- [18] J. L. Grabmeier y L. A. Lambe, "Decision trees for binary classification variables grow equally with the Gini impurity measure and Pearson's chi-square test," *International journal of business intelligence and data mining*, vol. 2, n.º 2, págs. 213-226, 2007.
- [19] P. C. Jennings, "Engineering features of the San Fernando earthquake of February 9, 1971," 1971.
- [20] F. S. Bao, X. Liu y C. Zhang, "PyEEG: an open source python module for EEG/MEG feature extraction," *Computational intelligence and neuroscience*, vol. 2011, 2011.
- [21] T. W. David Anderson Dennis Sweeney, *Estadística para negocios y economía*. Av. Santa Fe, núm. 505, piso 12 Col. Cruz Manca, Santa Fe C.P. 05349, México, D.F.: South-Western Cengage Learning, 2011.
- [22] B. Kemp, A. Värri, A. C. Rosa, K. D. Nielsen y J. Gade, "A simple format for exchange of digitized polygraphic recordings," *Electroencephalography and clinical neurophysiology*, vol. 82, n.º 5, págs. 391-393, 1992.
- [23] S. R. Erfan Khadem y col., "Sleep Stage Classification," B.S. thesis, Colorado: UCD, 2017, 2017.
- [24] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen y M. S. Hämäläinen, "MEG and EEG Data Analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, n.º 267, págs. 1-13, 2013. DOI: 10.3389/fnins.2013.00267.
- [25] J. Möcks y T. Gasser, "How to select epochs of the EEG at rest for quantitative analysis," *Electroencephalography and clinical neurophysiology*, vol. 58, n.º 1, págs. 89-92, 1984.

- [26] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot y E. Duchesnay, “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, págs. 2825-2830, 2011.
- [27] G. Lemaître, F. Nogueira y C. K. Aridas, “Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning,” *Journal of Machine Learning Research*, vol. 18, n.º 17, págs. 1-5, 2017. dirección: <http://jmlr.org/papers/v18/16-365.html>.
- [28] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho y col., “Xgboost: extreme gradient boosting,” *R package version 0.4-2*, vol. 1, n.º 4, págs. 1-4, 2015.
- [29] A. Cutler, D. R. Cutler y J. R. Stevens, “Random forests,” en *Ensemble machine learning*, Springer, 2012, págs. 157-175.
- [30] L. Accel Brain Co., *AccelBrainBeat*, 2017. dirección: <https://accel-brain.co.jp/>.
- [31] J. Robert, M. Webbie y col., *Pydub*, 2018. dirección: <http://pydub.com/>.
- [32] A. Parfenov, *BrainFlow*, 2019. dirección: <https://github.com/brainflow-dev/brainflow>.
- [33] R. Vallat y M. P. Walker, “A universal, open-source, high-performance tool for automated sleep staging,” *bioRxiv*, 2021.
- [34] A. K. Patel, V. Reddy y J. F. Araujo, “Physiology, sleep stages,” *StatPearls [Internet]*, 2021.
- [35] P. Armitage y col., “Sequential medical trials.,” *Sequential Medical Trials.*, 1960.
- [36] M. d. C. Lara y col., “Sobre el consentimiento informado,” *Boletín de la Oficina Sanitaria Panamericana (OSP)*; 108 (5-6), mayo.-jun. 1990, 1990.
- [37] O. Tsinalis, P. M. Matthews, Y. Guo y S. Zafeiriou, “Automatic sleep stage scoring with single-channel EEG using convolutional neural networks,” *arXiv preprint arXiv:1610.01683*, 2016.
- [38] A. Supratak, H. Dong, C. Wu e Y. Guo, “DeepSleepNet: A model for automatic sleep stage scoring based on raw single-channel EEG,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, n.º 11, págs. 1998-2008, 2017.
- [39] B. Koley y D. Dey, “An ensemble system for automatic sleep stage classification using single channel EEG signal,” *Computers in biology and medicine*, vol. 42, n.º 12, págs. 1186-1195, 2012.

CAPÍTULO 15

Anexos

A continuación se presenta el consentimiento informado, un esfuerzo realizado en conjunto con la dirección de la maestría en neuropsicología de la Universidad del Valle para el proceso de reclutamiento y participación en el proceso de experimentación.



FORMATO DE CONSENTIMIENTO INFORMADO PARA PARTICIPACIÓN EN INVESTIGACIÓN

Este documento tiene el propósito de explicar todas las cuestiones relativas a la utilización y manejo de la información que se recabe durante las sesiones de evaluación y recolección de datos. Además, el de informar y solicitar su autorización para participar en el presente estudio de investigación. Léalo atentamente y consulte con el encargado de la investigación todas las dudas que tenga al respecto.

Información sobre la investigación

Nombre del estudio: Estudio del Impacto de un Sistema Integrado de Detección de Patrones de Señales Biomédicas y de Generación de Pulsos Binaurales en la Calidad del Sueño de las Personas

Lugar y fecha: Guatemala, agosto de 2021.

Número de participantes: 25.

Encargado de la investigación: Diego Alejandro Alegría Ochoa, estudiante del programa de Ingeniería Electrónica en la Universidad del Valle de Guatemala.

Propósito y objetivo del estudio

El presente estudio de investigación tiene como objetivo analizar el impacto en la calidad de sueño de un sistema de generación de pulsos binaurales integrado a un clasificador automático de etapas de sueño. Los resultados de este estudio permitirán obtener información preliminar sobre terapias de neuro-retroalimentación y cómo estas pueden ser útiles para mitigar la privación de sueño. Esto es de mucha importancia ya que la privación de sueño contribuye en una gran parte al desarrollo de trastornos a niveles moleculares, inmunes y neurológicos.

Procedimiento

Las personas que decidan participar en la investigación recibirán un trato digno, respetuoso y atento durante todo el procedimiento, recibiendo asesoría de personal capacitado. Si usted acepta participar en este estudio, se le solicitará la utilización de un casco con electrodos de superficie no invasivos, los cuales permiten detectar actividad eléctrica del cerebro mientras usted duerme.

A continuación, se detalla la estructura del protocolo de evaluación y recolección de datos:

1. El estudio consta de una polisomnografía básica a través de un electroencefalograma el cuál se realiza de preferencia durante la noche.
2. Antes de realizar el procedimiento de recolección de datos se realizará una evaluación conductual sobre la concepción del sueño.
3. El procedimiento consiste en la medición de la actividad eléctrica del cerebro a través de electrodos de superficie no invasivos, la colocación de audífonos inalámbricos para la escucha de los pulsos binaurales con un volumen moderado y ajustable según la necesidad tal que esto no represente incomodidades al participante.
4. Durante la recepción de la actividad eléctrica un clasificador automático predice la etapa del sueño y genera pulsos binaurales según la etapa que se detecte.

5. Al finalizar se realizará una encuesta y evaluación posterior sobre los efectos percibidos durante el estudio.

Uso y confidencialidad de los datos

Si usted decide participar en el estudio, se le garantiza que toda la información proporcionada será manejada con absoluta confidencialidad. Sus datos personales no serán publicados ni revelados a nadie. El encargado de la investigación se hace responsable de la custodia y privacidad de estos, y del riguroso cumplimiento del secreto profesional en el uso y manejo de la información obtenida. Los datos que se obtengan de su participación serán utilizados únicamente con fines científicos y académicos por el encargado de la investigación dentro del programa de Ingeniería Electrónica de la Universidad del Valle de Guatemala.

Riesgos poco frecuentes asociados a su participación en el estudio

A pesar de la adecuada selección de la técnica y el proceso de investigación detrás de este estudio, se pueden presentar efectos no deseados. Entre estos efectos no deseados asociados al procedimiento se puede presentar:

- Irritación en la piel por el contacto con el gorro que contiene los electrodos la cual desaparecería en horas o máximo un día.
- Dificultad para conciliar el sueño.
- Leve dolor de cabeza.

En el caso de presentarse alguna complicación asociada o no asociada al procedimiento el estudio será interrumpido de forma inmediata.

Beneficios asociados a su participación en el estudio

La deficiencia crónica de sueño, definida como el estado inadecuado o mínimo de sueño, es un factor determinante e infravalorado en lo que concierne al estado de salud, afectando de forma directa la longitud y calidad de vida de las personas. Su participación en este estudio es muy valiosa ya que los resultados servirán para explorar el tratamiento de trastornos de sueño a través de terapias de neuro-retroalimentación basada en pulsos binaurales. Con esta terapia se pretende ayudar a personas que tienen dificultad para conciliar el sueño, y personas cuyo sueño no sea reparador.

Voluntariedad

Su participación es voluntaria. Si usted decide no participar o desea retirarse del estudio en cualquier momento, aun cuando haya iniciado la evaluación, puede hacerlo ningún problema.

Declaración de consentimiento de participación

Yo, _____,
he leído el presente documento de consentimiento informado y he comprendido las explicaciones en él facilitadas acerca de la participación en la investigación. También he podido resolver todas las dudas y preguntas que he planteado al respecto. Además, comprendo que, en cualquier momento y sin necesidad de dar ninguna explicación, puedo revocar el consentimiento que ahora presento. Finalmente, he sido informado/a de que mis datos personales serán protegidos y serán utilizados únicamente con fines científicos y académicos.

Tomando todo lo anterior en consideración y en tales condiciones, CONSIENTO participar en la investigación y que los datos que se deriven de mi participación sean utilizados para cubrir los objetivos especificados en el documento.

Declaración del encargado de la investigación

Yo, _____,
declaro no tener ningún conflicto de intereses y que he explicado la naturaleza y el objetivo de la investigación,
y que el/la participante entiende en qué consiste su colaboración. He leído y explicado adecuadamente las
partes del consentimiento informado. Así mismo, todas las preguntas que el/la participante ha hecho han sido
contestadas de forma adecuada.

Fecha: ____/____/____

Firma de la Participante

DPI _____

Firma del Testigo

DPI _____

Firma del Investigador

DPI _____

Firma del Asesor

DPI _____

Contacto

Si tiene dudas puede comunicarse con:

Investigador

Diego Alejandro Alegría Ochoa
Teléfono: +502 55289474
Correo electrónico: ale15171@uvg.edu.gt

Comité de Ética en Ciencias Sociales

Universidad del Valle de Guatemala (UVG)
Teléfono: (502) 2507-1500
Correo electrónico: eticaccss@uvg.edu.gt

Nombre

Dr. Luis Alberto Rivera
Departamento de Ingeniería Electrónica,
Mecatrónica y Biomédica
Universidad del Valle de Guatemala
Teléfono: +502 5385-3237
Correo electrónico: larivera@uvg.edu.gt

