



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

IDENTIFICACIÓN DE BIOMARCADORES DE POTENCIAL EVOCADO Y  
CONECTIVIDAD CEREBRAL PARA LA CLASIFICACIÓN DE ESQUIZOFRENIA  
CON REDES PROFUNDAS

TESIS PARA OPTAR AL GRADO DE  
MAGÍSTER EN CIENCIAS, MENCIÓN COMPUTACIÓN

MEMORIA PARA OPTAR AL TÍTULO DE  
INGENIERA CIVIL EN COMPUTACIÓN

ALEXANDRA SOFÍA IBARRA CUESTA

PROFESORA GUÍA:  
JOCELYN SIMMONDS WAGEMANN  
PROFESOR GUÍA 2:  
ÁNGEL JIMÉNEZ MOLINA

PROFESOR CO-GUÍA:  
PABLO GASPAR RAMOS

MIEMBROS DE LA COMISIÓN:  
SERGIO OCHOA DELORENZI  
FELIPE BRAVO MÁRQUEZ  
VIOLETA CHANG CAMACHO

SANTIAGO DE CHILE  
2023

# Resumen

La esquizofrenia es uno de los trastornos psiquiátricos más graves que afecta a la población mundial debido a su diagnóstico tardío y las consecuencias de esta conlleva sobre quién la sufre. Más aún, en Chile se encuentra dentro del listado de las 100 enfermedades más importantes, las cuales están cubiertas por el sistema de salud (GES y Auge).

Este trabajo investigativo gira en torno a responder diversas preguntas de investigación que fueron estipuladas en la etapa preliminar de esta tesis, razón por la cual se elaboraron distintos objetivos que permitiesen concretar dicho fin y una hipótesis que englobase las interrogantes establecidas.

La hipótesis se centra en determinar la posibilidad de diseñar un modelo predictivo en base a redes neuronales que permita la detección de esquizofrenia en personas y, a su vez, que este nuevo modelo presente un rendimiento equiparable a otros clasificadores generados y estudiados por diferentes investigadores del mundo.

Para esto se utilizaron señales neuronales de personas con y sin esquizofrenia, lo cual permitió desarrollar un clasificador de dicha enfermedad mediante la combinación de un modelo de red neuronal recurrente con un modelo de red neuronal feedforward. Este clasificador permite utilizar de manera simultánea el potencial evocado y la conectividad cerebral obtenidas a partir de las señales cerebrales de una persona que se encuentra expuesta a distintos estímulos. Cabe destacar que estas dos características han sido utilizadas en diversos trabajos para la detección de esquizofrenia.

Para lograr extraer dichas características se dió paso a un proceso completo de limpieza y filtrado de las señales, así como también la aplicación de distintas técnicas computacionales que permitían extraer las características de interés.

Una vez obtenidas las características y diseñado el modelo predictivo se dió paso a la etapa de entrenamiento y prueba del modelo, logrando como resultado un clasificador que cumplía tanto con las expectativas establecidas en los objetivos iniciales de esta tesis, como también con un rendimiento equiparable a los existentes. Además, se pudo determinar cuales de las características utilizadas (potenciales evocados y conectividad cerebral) y estímulos a los cuales fue expuesta una persona, eran los primordiales para generar un modelo predictivo de esquizofrenia, permitiendo cumplir con la hipótesis, preguntas y objetivos establecidos.

*Cada vez que miro las estrellas, sé que están allí para guiar mis pasos.  
Por y para los que ya no están.*

# Agradecimientos

Partiendo por agradecer a mi familia, que gracias a ellos pude lograr mi sueño de estudiar la carrera que quería en la universidad que quería. Gracias a su esfuerzo incansable que me permitieron centrarme a la universidad sin preocuparme de cómo pagar la mensualidad o de tener un plato de comida para llevar.

También quiero agradecer a los que ya no están, mi abuela, mi abuelo, que sin ellos no hubiese podido muchas veces estudiar tranquila; ellos que se dedicaron a estar por y para uno, que siempre tuviese el almuerzo del día siguiente y la cena llegando de la universidad; ellos que hubiesen querido estar presentes en cada uno de mis logros, pero más no están ausentes.

Un agradecimiento especial a mis amigos, los que hice en plan común y especialidad, puesto que sin ellos el camino por la universidad no hubiese sido tan gratificante como lo fue. También por siempre darme ánimo para continuar pese a todas las adversidad que la vida colocó en el camino.

También quiero dar las gracias al laboratorio psiquiátrico de la Universidad de Chile, sobretodo a Rocío Mayol que fue de gran ayuda para entender conceptos médicos y psiquiátricos, así como también por darse el tiempo y espacio para ayudarme en la etapa de preprocesamiento y limpieza de las señales. Sin su ayuda y paciencia hubiese sido muy complejo comprender los pasos a seguir.

Y a todos los que estuvieron y seguirán estando; profesores de la U que recordaré con mucho cariño; secretarias docentes maravillosas con las cuales siempre fue grato conversar; alumnos que estuvieron presentes en las clases que dicté y que me hacían sentir feliz con el poder ejercer como profesora auxiliar. Y también a todos aquellos profesores que me formaron en mis años de colegio y liceo, sin ellos no estaría donde estoy.

Simplemente, gracias a todos.

# Tabla de Contenido

<b>1. Introducción</b>	<b>1</b>
1.1. Definición de los datos . . . . .	3
1.2. Definición del problema . . . . .	4
1.3. Hipótesis y objetivos . . . . .	6
1.3.1. Preguntas de investigación . . . . .	6
1.3.2. Hipótesis . . . . .	6
1.3.3. Objetivos . . . . .	6
1.4. Solución propuesta . . . . .	7
1.5. Metodología . . . . .	8
1.6. Estructura del documento . . . . .	9
<b>2. Antecedentes y trabajos relacionados</b>	<b>10</b>
2.1. Marco teórico . . . . .	10
2.1.1. Conceptos básicos . . . . .	10
2.1.2. Software y lenguajes utilizados . . . . .	11
2.1.3. Profundización pruebas realizadas . . . . .	13
2.2. Estado del arte . . . . .	21
<b>3. Procesamiento de datos</b>	<b>25</b>
3.1. Preprocesamiento . . . . .	25
3.2. Filtrado de señal . . . . .	29
3.2.1. Inspección y limpieza visual . . . . .	29

3.2.2. ICA y rechazo de componentes . . . . .	33
3.3. Extracción de características . . . . .	36
3.3.1. Potencial Evocado (ERP) . . . . .	36
3.3.2. Conectividad Cerebral (CC) . . . . .	38
<b>4. Clasificación</b>	<b>40</b>
4.1. Conformación del Dataset . . . . .	40
4.1.1. Concatenación de datos: conformación de un dato único . . . . .	43
4.1.2. ERP como dato independiente . . . . .	45
4.2. Modelo . . . . .	45
4.3. Selección de características . . . . .	52
4.4. Entrenamiento y prueba . . . . .	54
4.4.1. Cross-Validation . . . . .	55
4.4.2. Evaluar el sobreajuste del modelo . . . . .	59
4.4.3. Métricas y métodos de evaluación . . . . .	59
<b>5. Resultados y discusión</b>	<b>62</b>
5.1. Resultados generales . . . . .	63
5.1.1. Significancia estadística . . . . .	71
5.2. Elección del mejor modelo . . . . .	76
5.3. Obtención de Biomarcadores . . . . .	78
<b>6. Conclusión</b>	<b>80</b>
<b>Bibliografía</b>	<b>86</b>
<b>Anexo</b>	<b>100</b>

# Índice de Tablas

1.1. Cantidad de sujetos por grupo con registro de cada tarea. . . . .	4
1.2. Cantidad de sujetos por grupo con registro de cada tarea. . . . .	4
2.1. Marcas o <i>event types</i> de la prueba STB . . . . .	15
2.2. Marcas o <i>event types</i> de la prueba DBF . . . . .	16
2.3. Escenarios prueba MMN . . . . .	17
2.4. Marcas o <i>event types</i> de la prueba MMN Emotional . . . . .	18
2.5. Marcas o <i>event types</i> de la prueba MMN Emotional . . . . .	19
3.1. Sujetos por prueba y grupo . . . . .	31
3.2. Sujetos por prueba y grupo . . . . .	31
3.3. Detalle de ERP . . . . .	37
5.1. Pruebas realizadas para obtención de resultados . . . . .	62
5.2. Resultados DBF para modelo Feedforward. . . . .	63
5.3. Pruebas realizadas para obtención de resultados . . . . .	64
5.4. Resultados DBF para modelo RNN. . . . .	65
5.5. Resultados STB para modelo RNN (parte 1). . . . .	66
5.6. Resultados STB para modelo RNN (parte 2). . . . .	67
5.7. Resultados STB para modelo RNN (data concatenada). . . . .	68
5.8. Mejores resultados . . . . .	70
5.9. ANOVA STB no concatenada: Accuracy . . . . .	72
5.10. ANOVA STB no concatenada: Recall . . . . .	72

5.11. ANOVA STB concatenada: Accuracy . . . . .	73
5.12. ANOVA STB concatenada: Recall . . . . .	73
5.13. ANOVA DBF feed forward: Accuracy . . . . .	74
5.14. ANOVA DBF feed forward: Recall . . . . .	74
5.15. ANOVA DBF red recurrente: Accuracy . . . . .	74
5.16. ANOVA DBF red recurrente: Recall . . . . .	75
5.17. ANOVA Red Recurrente: Rendimiento . . . . .	75
5.18. ANOVA Red Recurrente: Recall . . . . .	76
5.19. Mejor clasificador . . . . .	76
5.20. ANOVA STB no concatenada: Accuracy . . . . .	77
5.21. ANOVA STB no concatenada: Recall . . . . .	77
5.22. ANOVA STB no concatenada: Accuracy . . . . .	78
5.23. ANOVA STB no concatenada: Recall . . . . .	78
6.1. Análisis visual para pacientes sanos en prueba STB . . . . .	100
6.2. Análisis visual para pacientes con esquizofrenia en prueba STB . . . . .	101
6.3. Análisis visual para pacientes sanos en prueba DBF . . . . .	101
6.4. Análisis visual para pacientes con esquizofrenia en prueba RS . . . . .	102
6.5. Análisis visual para pacientes con esquizofrenia en prueba DBF . . . . .	102
6.6. Análisis visual para pacientes sanos en prueba RS . . . . .	103
6.7. Análisis visual para el rechazo de componentes (ICA) en pacientes sanos para la prueba STB Parte 1 . . . . .	103
6.8. Análisis visual para el rechazo de componentes (ICA) en pacientes sanos para la prueba STB Parte 2 . . . . .	104
6.9. Análisis visual para el rechazo de componentes (ICA) en pacientes con esquizofrenia para la prueba STB . . . . .	104
6.10. Análisis visual para el rechazo de componentes (ICA) en pacientes con esquizofrenia para la prueba DBF . . . . .	105
6.11. Análisis visual para el rechazo de componentes (ICA) en pacientes sanos para la prueba DBF . . . . .	105

6.12. Análisis visual para el rechazo de componentes (ICA) en pacientes sanos para la prueba RS Parte 1 . . . . .	106
6.13. Análisis visual para el rechazo de componentes (ICA) en pacientes sanos para la prueba RS Parte 2 . . . . .	107
6.14. Análisis visual para el rechazo de componentes (ICA) en pacientes con esquizofrenia para la prueba RS Parte 1 . . . . .	107
6.15. Análisis visual para el rechazo de componentes (ICA) en pacientes con esquizofrenia para la prueba RS Parte 2 . . . . .	108
6.16. Resultados con $acc \geq 75\%$ pero con overfitting (parte 1). . . . .	108
6.17. Resultados con $acc \geq 75\%$ pero con overfitting (parte 2). . . . .	109
6.18. Resultados con $acc \geq 75\%$ pero con underfitting. . . . .	110
6.19. Resultados con $acc \geq 75\%$ pero con overfitting (parte 3). . . . .	111
6.20. Métodos estadísticos utilizados para calcular la conectividad cerebral . . . .	111

# Índice de Ilustraciones

1.1. Aproximación de la distribución espacial de los electrodos (círculos rojos). . .	3
1.2. Obtención de Potenciales evocado. EEG corresponde a la señal capturada, ERP corresponde a potenciales evocados obtenidos tras transformaciones de la señal, representados por picos y valles. . . . .	5
2.1. Ejemplo prueba STB . . . . .	14
2.2. Ejemplo trial STB . . . . .	14
2.3. Ejemplo trial DBF . . . . .	17
2.4. Ejemplo trial MMN . . . . .	18
2.5. Ejemplo trial MVD . . . . .	20
3.1. Ejemplo de problemas en la señal . . . . .	30
3.2. Ejemplo de ICA . . . . .	34
3.3. Artefactos EOG . . . . .	34
4.1. Proporción de clases por prueba . . . . .	41
4.2. Ejemplo de uso de Zero Padding. . . . .	44
4.3. Red FeedForward . . . . .	46
4.4. Red Neuronal Recurrent . . . . .	46
4.5. Funciones de activación . . . . .	49
4.6. Diseño del Modelo concatenado elaborado. . . . .	50
4.7. Diseño de la Red Neuronal Recurrente. Modelo usado para procesar datos ERP. . . . .	51
4.8. Diseño de la Red Feedforward. Modelo usado para procesar datos de Conectividad cerebral. . . . .	52

4.9. Ejemplo selección de características . . . . .	53
4.10. Proceso de Cross-validation con $k = 3$ . . . . .	55
4.11. Proceso de Cross-validation con $k = 3$ . Donde $erp_i$ corresponde al $i$ -ésimo ERP del sujeto y $S_i$ corresponde al $i$ -ésimo sujeto del cual se extrajeron los ERP. . . . .	57
4.12. Proceso de Cross-validation con $k = 3$ para el caso de uso de 2 pruebas. Donde $erp_i$ corresponde al $i$ -ésimo ERP del sujeto; $S_i$ corresponde al $i$ -ésimo sujeto del cual se extrajeron los ERP y $P_j S_i$ corresponde al $i$ -ésimo sujeto de la $j$ -ésima prueba. . . . .	58
4.13. Overfitting y Underfitting . . . . .	59

# Capítulo 1

## Introducción

La esquizofrenia es una enfermedad mental grave que afecta algunas funciones cerebrales, generando una alteración de la percepción de la realidad lo que provoca que las personas que la sufren consideren reales ciertas cosas que no son verdad [49, 5].

Esta enfermedad afecta al 1 % de la población mundial [169], mientras que en Chile tiene una prevalencia de vida del 1,02 % en la población general [52], con una incidencia de 12 nuevos casos por cada cien mil habitantes por año [50].

A su vez, se considera que la esquizofrenia tiene un importante peso en la mortalidad y discapacidad de una población, siendo la responsable de un 1,87 % del total de Años de Vida Saludable [53] (AVISA)<sup>1</sup> perdidos por todas las enfermedades en el país, lo cual se traduce también en un impacto económico, estimándose una pérdida de 85 mil millones de pesos, ocupando la posición 14 en el ranking de las 100 enfermedades más importantes de Chile [53].

Por otra parte, la esquizofrenia se considera la tercera patología psiquiátrica causante del suicidio, por debajo de los trastornos depresivos y la dependencia al alcohol y drogas [103, 119]. Lo anterior, sumado a factores relacionados con el menoscabo de la calidad de vida de la persona, provoca que se triplique el riesgo de muerte [82], disminuyendo la esperanza de vida de las personas que la sufren a casi un 10 % menos que la población general [135, 159].

La esquizofrenia es una condición crónica que tiene efectos devastadores en muchos aspectos de la vida del paciente, y conlleva un alto riesgo de suicidio y el deterioro de la calidad de vida de la persona [52, 176]. Por esto, se necesita de un tratamiento que permita reducir los síntomas que se presentan y así mantener controlada la enfermedad, permitiendo que el paciente pueda realizar una vida lo más normal posible.

La intervención temprana juega un rol fundamental, puesto que cuanto antes reciba tratamiento una persona con esquizofrenia, mayor será el grado de recuperación y la probabilidad de reinserción de esta [167, 140]. Por esta razón es de suma importancia la detección precoz de pacientes con riesgo a este tipo de enfermedades.

---

<sup>1</sup>Factor para dimensionar la magnitud de daño

Actualmente, no existe ninguna prueba específica que permita asegurar que un paciente padece de esquizofrenia [166, 176]. Debido a esto, el proceso de diagnóstico contempla un largo período de evaluación, el cual es bastante complejo debido a la cantidad de observaciones y pruebas que se le deben realizar al paciente [168, 18] para excluir el resto de los trastornos psicóticos dado que estos tienden a presentar los mismos síntomas [165]. Por esta razón, se necesita de un período ininterrumpido de al menos 6 meses de síntomas [170] y haberse producido varios episodios con ciertas características específicas [164] para determinar si realmente se está en presencia de un trastorno esquizofrénico o no.

Debido a lo anterior, y al gran impacto que genera esta enfermedad, hoy en día el mundo psiquiátrico está indagando nuevas alternativas entre las cuales se encuentra la búsqueda y reconocimiento de biomarcadores<sup>2</sup> cerebrales [175, 121, 113, 39], que permitan realizar la detección de un cuadro psicótico en etapas tempranas (previo al desarrollo como tal de la enfermedad) y lo más rápido posible. Esto con la finalidad de detener la propagación de los efectos de la esquizofrenia en los pacientes que la padezcan.

Algunos de los estudios enfocados en la búsqueda de estos biomarcadores cerebrales, utilizan como base factores que presentan diferencias significativas entre sanos y esquizofrénicos. Entre estos, se encuentra el uso de los potenciales evocados o ERP, que son cambios en el potencial eléctrico cerebral ocasionados por algún estímulo [51], el cual se registra mediante un electroencefalograma (EEG). Así como también, la conectividad cerebral, la cual muestra la interacción entre distintas regiones cerebrales [66].

Del párrafo anterior se desprende que, desde un punto de vista médico, se está en búsqueda de métodos que permitan predecir la esquizofrenia. Esto mismo, desde el punto de vista de la ciencia de los datos, se trata de resolver un problema de clasificación.

Para contextualizar, un problema de clasificación busca encontrar un sistema capaz de identificar automáticamente la clase a la cual pertenece un objeto, por ejemplo un sistema que pueda discernir entre diferentes animales, o que puede diferenciar entre un paciente con diabetes y uno sano [17]. Para abordar estos problemas existen los métodos de aprendizaje automático, en particular las técnicas de Machine Learning y Deep Learning, las cuales son usadas para enseñar a los computadores a aprender en base a ejemplos.

Machine Learning se caracteriza por identificar patrones de comportamiento, los cuales utiliza para predecir comportamientos futuros, siendo capaces de identificar objetos [156]. Por otra parte, Deep Learning es una técnica que permite reconocer y distinguir entre distintos objetos en base a la idea de que dados unos parámetros hay una forma de combinarlos para predecir un cierto resultado [129, 156].

En base a la información previa, es notorio ver que resulta de gran importancia el estudio de diversos factores que permitan detectar rápidamente si un paciente sufre de esquizofrenia, con el objetivo de suministrar un tratamiento temprano para evitar el avance y empeoramiento de la enfermedad. Por esta razón, en esta tesis se abordará el problema de clasificación de esquizofrenia para evaluar la importancia del potencial evocado y conectividad cerebral en el

---

<sup>2</sup>Marcador biológico que puede proporcionar información sobre el estado de salud de una persona. Se puede usar para detectar una enfermedad, un cambio fisiológico, una respuesta al tratamiento o un trastorno psíquico.

proceso de detección de esta enfermedad.

Para esto, se utilizarán métodos de Deep Learning, que como se mencionó en el párrafo anterior, sirven para afrontar problemas de clasificación buscando la combinación de patrones que permita dar con el resultado correcto. Todo lo anterior, bajo el objetivo de determinar biomarcadores asociados al potencial evocado y conectividad cerebral para detectar esquizofrenia.

## 1.1. Definición de los datos

En total se tienen datos asociados a 118 sujetos, los cuales fueron recolectados por el laboratorio psiquiátrico de la Universidad de Chile, Psiquislab. Estos datos cuentan con todas las medidas legales y de confidencialidad necesarias para hacer uso de estos.

Los sujetos evaluados presentan distintos grados de evolución de la enfermedad diferenciados según la severidad de sus síntomas. Por esta razón se distinguen 5 grupos, APS (pacientes adultos con pródromo<sup>3</sup> de esquizofrenia), FEP (pacientes adultos con menos de 2 años desde su primer episodio psicótico), EQZ (pacientes adultos con esquizofrenia crónica), CNTA (controles adolescentes sanos (14-18 años)) y CNTF (controles adultos sanos (19-30 años)).

Cada sujeto utilizó una gorra de 64 electrodos para registrar las señales EEG, así como también 8 electrodos extras (6 en el contorno de los ojos y 2 en los mastoides<sup>4</sup>), los cuales se usan para eliminar el ruido en las señales producto del parpadeo. En la Figura 1.1 se puede observar una aproximación a la distribución global de los electrodos.

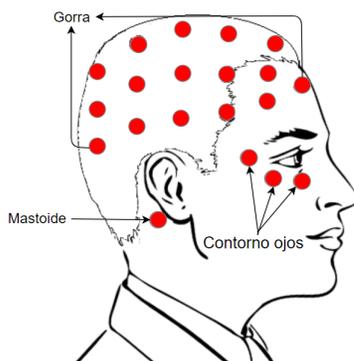


Figura 1.1: Aproximación de la distribución espacial de los electrodos (círculos rojos).

Los sujetos fueron sometidos a 5 pruebas en total, con y sin presencia de estímulos. Cabe destacar que los datos a utilizar fueron recopilados de diferentes proyectos realizados durante un período de 5 años, transcurso en el cual se fueron agregando nuevas tareas. Por esta razón, no todos los sujetos fueron sometidos a todas las pruebas realizadas.

En la tabla 1.1 se muestra un detalle de cada una de las pruebas realizadas con su breve descripción. Por otra parte, en la tabla 1.2 se observa la distribución de sujetos por grupo

<sup>3</sup>Síntomas iniciales que preceden al desarrollo de una enfermedad.

<sup>4</sup>Apófisis del hueso temporal situada detrás y debajo de la oreja.

con registro de cada tarea.

Prueba	Definición
RS	Prueba <b>sin</b> estímulos. 2 minutos con ojos abiertos, luego 2 minutos con ojos cerrados.
STB	Prueba con estímulo visual. Secuencia de imágenes a recordar con el objetivo de responder, al finalizar la serie, si una tarjeta apareció o no en dicha secuencia.
DBF	Prueba con estímulo visual y auditivo. Mezcla entre flashes blancos y un beep; el sujeto debe responder cuántos flashes logró ver.
MMN	Prueba con estímulo auditivo. Secuencia de un sonido frecuente con aparición de un sonido disruptivo que perturba la serie.
MVD	Prueba con estímulo visual. Se observan imágenes de puntos blancos aglomerados los cuales pueden estar estáticos o en movimiento.

Tabla 1.1: Cantidad de sujetos por grupo con registro de cada tarea.

Grupos	N Sujetos	RS	STB	DBF	MMN	MVD
APS	30	25	19	16	10	16
FEP	18	16	18	11	0	1
EQZ	10	8	0	2	0	0
CNTA	11	10	9	1	2	8
CNTF	49	36	19	14	15	13

Tabla 1.2: Cantidad de sujetos por grupo con registro de cada tarea.

## 1.2. Definición del problema

La esquizofrenia es una de las enfermedades más conocidas dentro de los trastornos psíquicos, la cual tiende a presentarse en la edad formativa de las personas (entre 18 y 30 años) [54], y cuyos síntomas o señales varían según cada individuo.

Las personas que sufren esta enfermedad se ven afectados negativamente en el ámbito laboral y social, así como también su calidad de vida, provocando problemas más serios asociados a complicaciones de salud, depresión e intentos de suicidios, lo que incrementa el riesgo de muerte de quien la padece.

Por estas razones, es de suma importancia detectar de manera rápida y oportuna la esquizofrenia en pacientes, puesto que mientras más tiempo pase desde el inicio de síntomas

psicóticos hasta la pérdida total del juicio de realidad, más agudos se harán los síntomas, siendo más compleja su recuperación y menor las posibilidades de reinserirse socialmente.

Los pacientes con esquizofrenia sufren de la alteración de ciertos factores cerebrales los cuales, a su vez, han servido para detectar la presencia de esta enfermedad en las personas. Entre estos factores se encuentran los potenciales evocados y la conectividad cerebral.

Los potenciales evocados (ERP) son cambios de voltajes contenidos en un segmento de las señales EEG, los cuales adoptan formas de picos o valles (ver cuadro inferior (ERP) de la Figura 1.2). El estudio de los potenciales evocados permite observar cómo reacciona el sistema nervioso frente a estímulos visuales, táctiles o auditivos, lo cual se traduce en impulsos eléctricos que son recogidos y amplificados por un electroencefalograma. Debido a que la señal es muy pequeña, es necesario realizar varias repeticiones de un mismo estímulo para luego promediar las respuestas, obteniendo así los potenciales [187] (ver proceso representado en la Figura 1.2).

Por otra parte, la conectividad cerebral (CC) se refiere a la interacción entre diferentes regiones del cerebro que anatómicamente están separadas, la cual es fundamental en los procesos cognitivos.

A partir de lo anterior, se plantea abordar el problema de clasificación de esquizofrenia mediante modelos de Deep Learning, usando los potenciales evocados y la conectividad cerebral de las personas. Para esto se utilizarán las señales EEG de pacientes con distintos niveles o estados de esquizofrenia, de las cuales se filtrarán los potenciales evocados presentes y se calculará la conectividad cerebral del sujeto.

Lo anterior es con la finalidad de determinar el nivel de significancia de cada uno de estos componentes en la clasificación, así como también características interpretables en el contexto clínico. De esta manera se pretende establecer cuáles factores asociados al potencial evocado y conectividad cerebral, son relevantes a la hora de detectar esquizofrenia en pacientes.

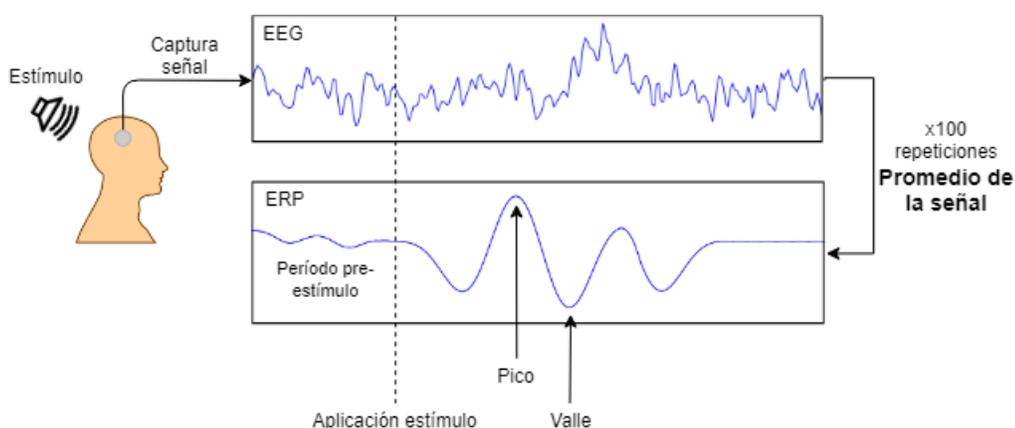


Figura 1.2: Obtención de Potenciales evocado. EEG corresponde a la señal capturada, ERP corresponde a potenciales evocados obtenidos tras transformaciones de la señal, representados por picos y valles.

## 1.3. Hipótesis y objetivos

### 1.3.1. Preguntas de investigación

En base del problema que se desea resolver, se plantearon las siguientes preguntas de investigación:

- ¿Es posible diseñar un modelo de clasificación utilizando el potencial evocado y conectividad cerebral, que permita detectar pacientes con esquizofrenia con un porcentaje de acierto equiparable a los clasificadores existentes para este fin?
- ¿Qué componentes del potencial evocado y la conectividad cerebral permiten clasificar la esquizofrenia? y ¿Cuál es su nivel de significancia dentro de la clasificación?
- ¿Cómo impacta en el proceso de clasificación el uso del potencial evocado y conectividad cerebral, discriminando entre estado de reposo y la memoria de trabajo frente a estímulos?

### 1.3.2. Hipótesis

En base a las preguntas de investigación, se formula la siguiente hipótesis:

*“Es posible obtener biomarcadores para la esquizofrenia asociados a componentes del potencial evocado y la conectividad cerebral, que permitan desarrollar un modelo de clasificación con una tasa de acierto superior al 75 %”*

A modo de acotación, es importante señalar que se define un umbral de 75% dado que, en la literatura, ese es el orden de desempeño de los modelos de clasificación que se han logrado [181, 22, 87, 61].

### 1.3.3. Objetivos

Para este trabajo se cuenta con una serie de datos que fueron obtenidos por el equipo de investigadores del laboratorio PsiquisLab<sup>5</sup>, de la Clínica Psiquiátrica Universitaria de la Universidad de Chile, quienes cuentan con la autorización del comité de ética de la Facultad de Medicina de la Universidad de Chile.

Con esto en cuenta, y en base a la hipótesis señalada en la sección anterior, se establecen los siguientes objetivos para esta tesis.

---

<sup>5</sup><https://www.psiquislab.cl/>

## Objetivo general

Determinar la significancia del potencial evocado y la conectividad cerebral en el proceso de clasificación de esquizofrenia, con la finalidad de establecer biomarcadores asociados a estos.

## Objetivos específicos

1. Calcular las características a estudiar para la detección de esquizofrenia (potencial evocado y conectividad cerebral).
2. Analizar y categorizar los datos en base a los distintos pacientes.
3. Establecer conjunto de características en base al potencial evocado y la conectividad cerebral de cada paciente.
4. Establecer arquitectura y técnicas a utilizar para definir un modelo de Deep Learning para la clasificación de esquizofrenia.
5. Establecer métricas para evaluar el desempeño del clasificador.
6. Examinar e interpretar resultados.
7. Identificar biomarcadores.

## 1.4. Solución propuesta

Para poder resolver el problema planteado se dividió la interrogante en dos partes principales. Primero, en cómo lograr evaluar el “peso” de cada una de las características del potencial evocado y la conectividad cerebral, para así obtener los biomarcadores asociados. Y segundo, qué modelo o modelos de Deep Learning serían pertinentes para procesar cada uno de los datos.

Para resolver la primera interrogante se planteó evaluar todas las combinaciones posibles de características frente a las distintas configuraciones y modelos implementados, lo que representa una forma básica y rudimentaria de colocar “peso” a cada característica. Por ejemplo, en caso de tener 2 características ( $C_1$  y  $C_2$ ) entonces las combinaciones a probar serían ( $C_1$ ), ( $C_2$ ), ( $C_1, C_2$ ); donde la primera permite medir el impacto únicamente de la característica 1 (diciendo que el “peso” de la segunda característica es nulo en la determinación de la esquizofrenia); la segunda combinación permite ver el caso contrario (impacto únicamente de la característica 2) y la tercera combinación permite ver el impacto combinado de ambas características.

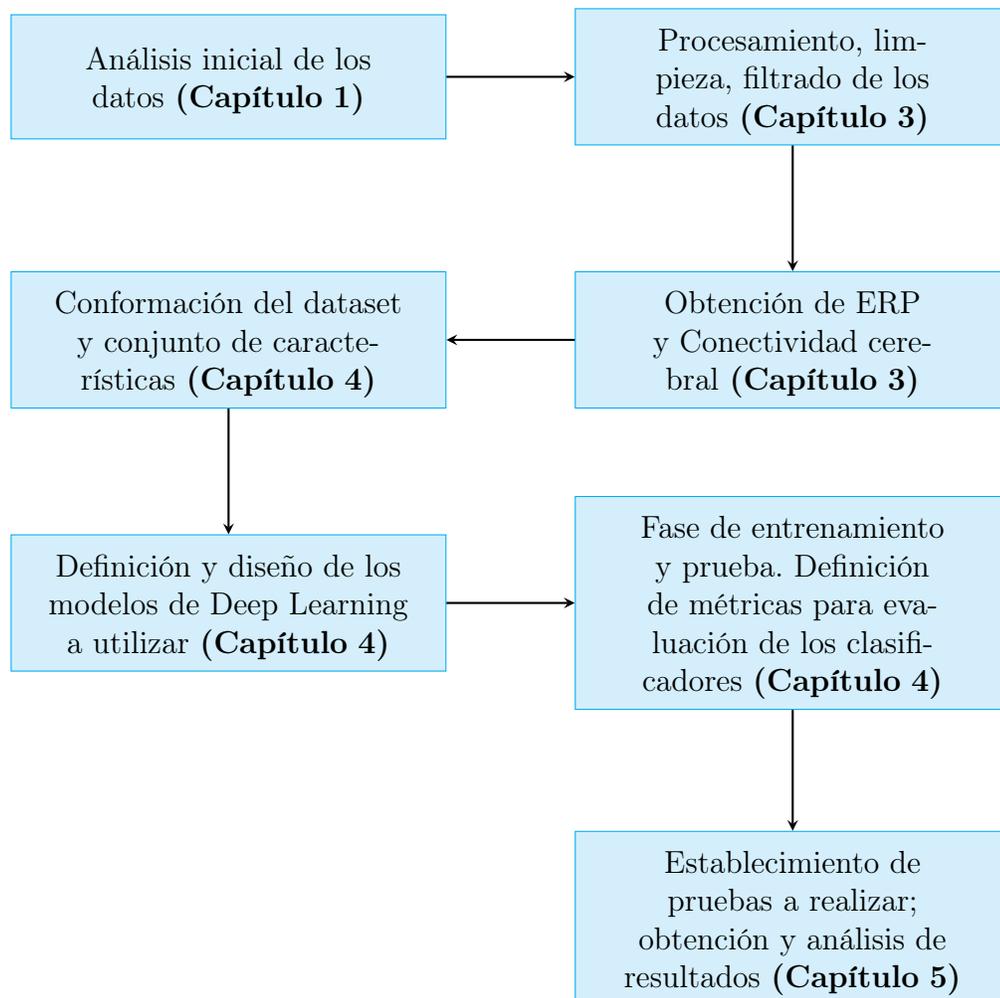
Debido a lo anterior, el clasificador se verá enfrentando a todas las posibilidades y se podrá determinar que combinación genera los mejores resultados. De esta forma se podrán

determinar los biomarcadores buscados para la detección de esquizofrenia, que serían las características que conforman la combinación perteneciente al clasificador con mejor resultado.

Para determinar cuál clasificador presenta un mejor rendimiento se utilizarán las métricas de accuracy, dado que en base a este dato se podrá validar la hipótesis inicial. Además se utilizará la métrica recall dado que esta representa una mayor importancia a nivel médico en cuanto a la detección de enfermedades de gran importancia.

Por otra parte, para resolver la interrogante respecto a que modelo o modelos utilizar se determinó utilizar 3 clasificadores diferentes que permitiesen medir el impacto de todas las combinaciones posibles. Se utilizará una red neuronal recurrente para poder procesar los datos asociados al potencial evocado; una red feed forward para entrenar los datos de la conectividad cerebral y, para el caso en que se deban utilizar ambos datos en simultáneo, se utilizará una red concatenada que permite unir la red recurrente con la red feed forward, las cuales se unen mediante una capa densa.

## 1.5. Metodología



En resumen, se debe realizar un análisis inicial de los datos para poder ver si hay problemas o errores en ellos. Luego se realiza el pre-procesamiento de los datos con el fin de eliminar el ruido de las señales y cualquier perturbación que se encuentre en estas que conlleve a resultados erróneos de cálculos.

Con los datos preparados se da paso a la obtención de características, es decir, extraer el potencial evocado y conectividad cerebral a partir de las señales cerebrales de cada paciente.

Con esto se da paso a la conformación del dataset, para lo cual es necesario dividir el set según clase (esquizofrénico/no esquizofrénico) y evaluar la necesidad de un balanceo de clases.

Después se da paso a la elección de un modelo de Deep Learning apto para las características a utilizar. Para esto se optó por una red feed forward, una red recurrente y una combinación de ambas.

Finalmente se da paso a la obtención y análisis de resultados, para esto se utilizó como métrica de evaluación el accuracy y el recall, por su significado a nivel médico y para permitir la validación de la hipótesis.

## 1.6. Estructura del documento

En el presente documento partiremos definiendo una serie de conceptos que serán necesarios aprender para poder entender cada parte de esta tesis, así como también ilustrar brevemente investigaciones que están relacionadas con el problema a abordar.

Posteriormente se dará paso a la presentación de los datos con los cuales se trabajó en esta tesis y, por sobretodo, como estos fueron obtenidos.

Luego se empezará de lleno en la metodología general, partiendo con la explicación del preprocesamiento realizado a los datos y la extracción de las características asociadas al potencial evocado y la conectividad cerebral.

Siguiendo con la metodología, se dará paso a la explicación exhaustiva de la etapa de clasificación, respondiendo preguntas tales como ¿Cuales modelos se utilizaron y porqué?, ¿Cómo se conforma el grupo de características?, ¿Cómo se realizará el entrenamiento de los clasificadores?, ¿Cómo se evaluará que clasificador es mejor?, entre otras.

Finalmente logrará encontrar los resultados obtenidos y cuales fueron las pruebas realizadas para poder determinar si se cumple o no el objetivo principal de esta tesis, además de una serie de discusiones y conclusiones obtenidas a partir del análisis de los resultados.

# Capítulo 2

## Antecedentes y trabajos relacionados

### 2.1. Marco teórico

#### 2.1.1. Conceptos básicos

El electroencefalograma o EEG, es un examen que sirve para medir la actividad eléctrica del cerebro [132]. Para este examen se utiliza un casco con un número determinado de electrodos, que corresponden a dispositivos como una placa metálica o una aguja pequeña que llevan señales eléctricas de los músculos, cerebro, corazón, piel u otras partes del cuerpo, hasta los aparatos de registro para ayudar a diagnosticar ciertas afecciones [137].

El examen de EEG se realiza frente a estímulos, que corresponden a señales sensitivas, motoras y/o cognitivas que se envían hacia una parte en específico del organismo con el fin de esperar una respuesta por parte del cerebro u órgano de destino [78], así como también en estado de reposo (sin estímulos).

A partir de las señales cerebrales es posible extraer los potenciales evocado o ERP, que son cambios de voltajes contenidos en un segmento de la señal del electroencefalograma; y la conectividad cerebral, que es la interacción entre diferentes regiones cerebrales anatómicamente separadas. Los cuales han sido utilizados como biomarcadores para la esquizofrenia, es decir, marcadores biológicos que puede proporcionar información sobre el estado de salud de una persona, que se utilizan para la detección de una enfermedad [45].

Para poder llevar a cabo estos procesos es necesario el filtrado de las señales, que es un procedimiento para reducir o eliminar el ruido de la señal [86]; junto con la aplicación del método ICA, que se encarga de extraer factores o componentes independientes de una señal [197].

La clasificación es un proceso que permite agrupar conjuntos de objetos en categorías o clases [47]. Este proceso es posible automatizarlo mediante la creación de un clasificador, algoritmo utilizado para asignar un elemento entrante no etiquetado en una categoría concreta conocida.

Hay dos técnicas para el diseño del clasificador: Machine Learning o aprendizaje automático, que es un tipo de inteligencia artificial que permite a un sistema aprender de los datos en lugar de aprender mediante programación explícita [101]; y Deep Learning o aprendizaje profundo, que es un tipo de aprendizaje automático que imita la forma en que los humanos obtienen ciertos tipos de conocimiento y el cual está inspirado en la arquitectura neuronal del cerebro humano [34].

La clasificación cuenta con una fase de entrenamiento, procedimiento utilizado para llevar a cabo el proceso de aprendizaje en una red neuronal [38], y una fase de prueba; esto con la finalidad de calibrar el clasificador y modificar sus hiperparámetros tales como las épocas, que corresponden a cada iteración de la red por el lote de entradas en la que haya ajuste de variables (fase de entrenamiento) [185]; la utilización de dropout, que es un procedimiento que permite desactivar aleatoriamente un porcentaje determinado de neuronas de las capas ocultas [65]; el número de capas ocultas, etc.

La modificación de estos hiperparámetros se realizan con el objetivo de mejorar la clasificación y disminuir el overfitting o sobreajuste de los datos, que se produce cuando un modelo se aprende los datos de entrenamiento perfectamente, por lo que no es capaz de generalizar y comprender nuevos datos de entrada [4]; y underfitting o subajuste de los datos, se produce cuando un modelo no es capaz de identificar patrones [117]

Un clasificador por red neuronal cuenta con una capa de entrada, que es la primera capa de una red neuronal la cual está compuesta por neuronas que reciben datos o señales procedentes del entorno, es decir, son las encargadas de recoger la entrada bruta [147]; capas ocultas, que son las capas intermedias de una red neuronal, no tiene una conexión directa con el entorno (no recibe señales externas ni entrega el resultado final); y una capa de salida, que es la última capa de una red neuronal la cual está compuesta por neuronas que proporcionan la respuesta de la red neuronal [147].

La capa oculta proporciona grados de libertad a la red neuronal gracias a los cuales es capaz de representar más fehacientemente determinadas características del entorno que trata de modelar [147]. Además, cada capa cuenta con su propia función de activación, que son funciones utilizadas para transformar el valor de salida de una capa, permitiendo acotar los valores de salida de una neurona y mantenerlos en un cierto rango [59].

Además, cuando el set de datos para la etapa de entrenamiento y prueba es pequeño, se opta por la utilización de cross-validation, que es una técnica para evaluar modelos de clasificación mediante el entrenamiento de varios modelos en subconjuntos de los datos de entrada disponibles y evaluarlos con el subconjunto complementario de los datos [13].

### 2.1.2. Software y lenguajes utilizados

Para el desarrollo de esta tesis se optó por el uso de dos lenguajes de programación diferentes. Para realizar toda lo correspondiente al procesamiento de los datos, se utilizó `Matlab`, puesto que este lenguaje es el que utilizan directamente en el laboratorio (*PsiquisLab*) y además cuenta con diferentes frameworks (conjunto de herramientas y módulos que pueden ser reutilizados para varios proyectos [10]) que permiten llevar a cabo el procesamiento de

los datos de forma más sencilla, debido a que cuentan con una interfaz gráfica que facilita el proceso de inspección, limpieza visual y rechazo de componentes.

Por otra parte, para todo el proceso de clasificación, el cual abarca desde la división del set de datos hasta obtener los resultados finales del clasificador, se utilizó `Python` dado que también es conocido y utilizado por el laboratorio, además que a opinión personal `Python` posee un mejor manejo y funcionalidades para trabajar con deep learning.

Por otra parte, se probaron inicialmente dos framework de `Matlab` para el procesamiento de los datos. Por una parte está `EEGLAB` cuya ventaja es que la interfaz gráfica que posee permite realizar todo los pasos de filtrado sin tener que escribir un *script* de código, además este framework fue uno de los aconsejados por *PsiquisLab* y de los más utilizados por ellos. Por otro lado se probó también con `FieldTrip`, que es un framework que se centra mucho más en la escritura de código para realizar cada paso del procesamiento.

En un inicio se optó por realizar todo el filtrado y la obtención de características de la señal mediante `EEGLAB` debido a su facilidad gráfica para realizar cada paso, lamentablemente cuando ya se había realizado todo el preprocesamiento de los datos más el filtrado de los mismos dando paso a la extracción de características se tuvo un problema.

`EEGLAB` no permitía obtener la conectividad cerebral de las personas pues el programa fallaba al momento de intentar realizarlo y era un problema de código interno que no se pudo solucionar. Por esta razón, en una primera instancia, se optó por pasar los datos ya filtrados a `FieldTrip` (pues posee una función de compatibilidad en el traspaso de información entre `EEGLAB` y `FieldTrip`) y calcular con este framework solamente la conectividad cerebral de los usuarios, para así evitar realizar todo el procesamiento nuevamente pues era lo que más tiempo conllevaba.

Lamentablemente, al realizar dicha conversión de los datos desde `EEGLAB` a `FieldTrip` había una variable que no se encontraba o bien no era generada al momento de realizar el traspaso de información de un framework a otro, lo cual generaba un error al intentar utilizar cualquier función propia de `FieldTrip`. En base a esto se tenían tres opciones:

- Arreglar el código fuente de `EEGLAB` para poder solucionar el error que se generaba al momento de determinar la conectividad funcional.

Esta opción se descartó por dos motivos principalmente, uno porqué se consideró que conllevaría más tiempo intentar encontrar el porqué se producía el error y así arreglarlo, de hecho se probaron con varias versiones de `EEGLAB` y cada una generaba algún error diferente y además había que instalar librerías extras para el cálculo tanto de la conectividad como de los potenciales evocados.

La otra razón es que el proceso para obtener los potenciales evocados no era muy claro, habían varias opciones diferentes en la interfaz para poder calcularlos y no había un buen manual de uso, por lo que para evitar errores se optó por descartar completamente `EEGLAB`.

- La otra opción que había era encontrar el porqué esa variable se “perdía” en el momento de traspasar la data desde `EEGLAB` a `Fieldtrip`, y en base a eso arreglar el error. O

simplemente encontrar un modo de hacer “aparecer” ese dato extra con la información correcta.

Se intentó junto con Rocío Mayol (una de las encargadas del equipo de investigación de *PsiquisLab* y encargada de la realización de las pruebas que se tomaron en los pacientes) hacer que ese parámetro volviese a estar en la información de cada dato, pero no dió frutos pues no había información concreta al respecto sino que un par de preguntas realizadas basadas en dicho error, pero sin solución alguna.

- La última opción correspondía a hacer todo el procesamiento de los datos nuevamente pero en `FieldTrip`. Finalmente esta fue la opción que se llevó a cabo, por lo que se realizó todo el preprocesamiento, filtrado y obtención de características de cada uno de los pacientes nuevamente.

### 2.1.3. Profundización pruebas realizadas

En total de realizaron cinco pruebas las cuales se detallarán a continuación.

#### Prueba 1: Resting State (RS)

La primera prueba que se le realizó a los pacientes fue la de estado basal, para medir como se comportan las señales cerebrales de cada uno cuando no se encuentran frente a estímulos.

En esta prueba, el paciente debe estar por 2 minutos con los ojos cerrados y sin realizar absolutamente nada. Pasado dicho tiempo, debe estar otros 2 minutos más con los ojos abiertos.

#### Prueba 2: Stemberg (STB)

Esta prueba consiste en mostrar al sujeto una secuencia de imágenes (1, 3 o 6 tarjetas) las cuales debe ser capaz de recordar, con la finalidad de que al finalizar cada serie responda si una cierta tarjeta apareció o no en la secuencia vista previamente.

Por ejemplo, en la figura 2.1 se muestra un ejemplo de esta prueba. En la parte superior se puede observar la secuencia que se le muestra al paciente, cada imagen aparece por un par de segundos. Mientras que en la parte inferior se observa la pregunta que se le realizó tras finalizar toda la secuencia. En este ejemplo la respuesta correcta es sí, dado que esa carta pertenece a la secuencia que está en la parte superior.

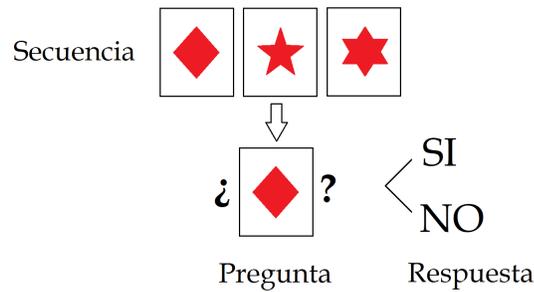


Figura 2.1: Ejemplo prueba STB

Las respuestas dadas por el paciente pueden pertenecer a alguna de las siguientes 4 categorías:

- Correcta (la tarjeta estaba)
- Correcta (la tarjeta no estaba)
- Incorrecta (no responde correctamente)
- Incorrecta pues responde fuera de tiempo

En la Figura 2.2 se ilustra parte de la señal EEG de un sujeto obtenida para esta prueba. Ahí se pueden observar varias marcas, señaladas con el nombre de *status*, que se utilizan dentro de la señal para indicar diferentes eventos durante el transcurso de la prueba.

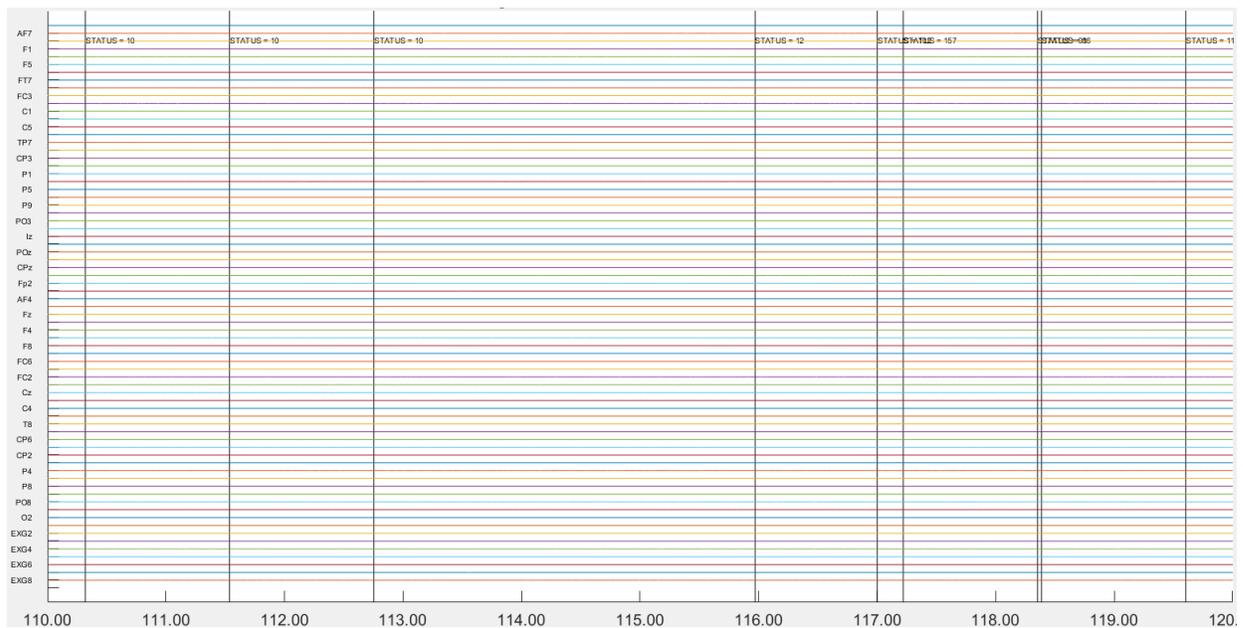


Figura 2.2: Ejemplo trial STB

Estas marcas, llamadas en el ámbito científico *event types*, se puede observar en la tabla 2.1.

Marca	Significado
157	Término de cada ronda o tarea de la prueba
102	Respuesta correcta (la tarjeta estaba)
101	Respuesta correcta (la tarjeta no estaba)
92	Respuesta incorrecta (no responde correctamente)
91	Respuesta incorrecta (fuera de tiempo)
86	Se muestran 6 tarjetas en la tarea
83	Se muestran 3 tarjetas en la tarea
80	Se muestra 1 tarjeta en la tarea
12	Se realiza la pregunta

Tabla 2.1: Marcas o *event types* de la prueba STB

Hay otras dos marcas (10 y 11) que no fueron consideradas en el estudio, por instrucción de la investigadora principal del laboratorio psiquiátrico, dada su poca influencia en este.

### Prueba 3: Double Flash (DBF)

En esta prueba, el paciente se encuentra sentado frente a un computador con su pantalla completamente en negro. Cada cierto instante de tiempo, aparece en pantalla uno o dos flashes blancos por una fracción de segundo.

El objetivo es que la persona identifique cuantos flashes observó. La peculiaridad de esta prueba radica en que además de un estímulo visual (flash) existe un estímulo auditivo (un beep) que viene a alterar la percepción del paciente.

Por ejemplo, cuando se genera un sonido beep seguido de un flash, el paciente podría percibir ese sonido como un flash extra, provocando que la cantidad de flashes visto por éste sean dos en vez de uno, lo que se considera como la *ilusión del doble flash*.

Las respuestas dadas por el pacientes se dividen en 6 categorías:

- Respuesta correcta, reporta ver 0 flashes.
- Respuesta correcta, reporta ver 1 flash.
- Respuesta correcta, reporta ver 2 flashes.
- Respuesta incorrecta, reporta ver 0 flashes.
- Respuesta incorrecta, reporta ver 1 flash.
- Respuesta incorrecta, reporta ver 2 flashes (ilusión del doble flash).

En la tabla 2.2, se pueden observar los *event types* para esta prueba

Marca	Significado
121	Época de ilusión de doble flash
100	Época sin estímulo
101	Época donde se muestra 1 flash
102	Época donde se muestran 2 flashes
111	Época donde hay un sonido seguido de un flash
120	Época donde hay dos sonidos seguidos
1	Reporta ver 0 flashes (respuesta correcta)
2	Reporta ver 1 flash (respuesta correcta)
3	Reporta ver 2 flashes (respuesta correcta)
10	Reporta ver 0 flashes (respuesta incorrecta)
11	Reporta ver 1 flash (respuesta incorrecta)
12	Reporta ver 2 flashes (respuesta incorrecta)

Tabla 2.2: Marcas o *event types* de la prueba DBF

Cabe destacar que para 6 datos (pacientes) de esta prueba se consideraron al revés las marcas que indican que la respuesta está correcta, es decir:

- 3: Reporta ver 0 flashes.
- 1: Reporta ver 1 flash.
- 2: Reporta ver 2 flashes.

Lo cual afectó a los pacientes del grupo APS (APS\_001 APS\_009, APS\_012, APS\_016, APS\_019, APS\_023).

En la Figura 2.3 se puede observar un ejemplo de las marcas que aparecen en la prueba.

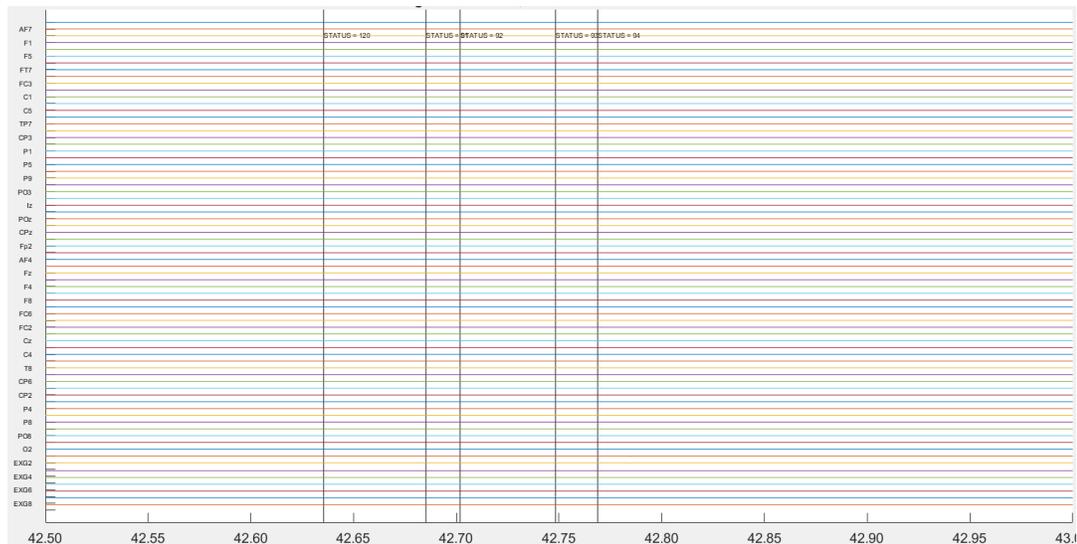


Figura 2.3: Ejemplo trial DBF

#### Prueba 4: MMN Emocional (MMN)

En esta prueba, el paciente se encuentra sentado escuchando una serie de sonidos. Estas series auditivas se componen siempre de sólo dos estímulos, uno que es frecuente (que es el que se escuchará constantemente) y otro llamado infrecuente, que es el que viene a romper la continuidad del estímulo anterior en la secuencia.

Al aparecer el estímulo infrecuente y el paciente sentir este cambio, se genera también un cambio en las señales cerebrales captadas por el electroencefalograma, el cual se verá reflejado en la generación de un potencial evocado específico llamado MMN.

En esta prueba, el paciente no debe responder ningún tipo de pregunta sino que se trata solamente de escuchar los sonidos que van apareciendo en cada serie auditiva.

En la tabla 2.3 se señalan los 5 escenarios existentes los cuales están divididos en 5 archivos diferentes (R1, R2, R3, R4, R5), donde EF corresponde al estímulo frecuente y EI al infrecuente

EF	EI
Violín alegre	Violín triste
Violín triste	Violín alegre
Violín	Flauta
Flauta	Violín
Tono simple 1kHz, 100 ms	Tono simple 1kHz, 50 ms

Tabla 2.3: Escenarios prueba MMN

Es importante señalar que las secuencias no siempre se mostraban en el mismo orden, sino

que se iba variando según paciente pues así los pacientes no inferían cuál sería la secuencia siguiente a la que ya estaban escuchando.

En la tabla 2.4 se observan los *event types* de la prueba MMN Emotional

Marca	Significado
180	Estímulo frecuente inicial
157	Inicio de la secuencia o pausa
80	Estímulo frecuente
70	Estímulo infrecuente

Tabla 2.4: Marcas o *event types* de la prueba MMN Emotional

Hay otros dos *event type* que aparecen en la señal pero que no están relacionados con la tesis por lo que no se consideran (158 y 1). Cabe destacar además que hay dos marcas que consideran el estímulo frecuente, cuya diferencia radica en que el 180 marca el primer estímulo frecuente que aparece en la secuencia, todos los que le siguen a ese tendrán la marca 80.

En la Figura 2.4 se puede observar un ejemplo de la señal neuronal con las marcas correspondientes.

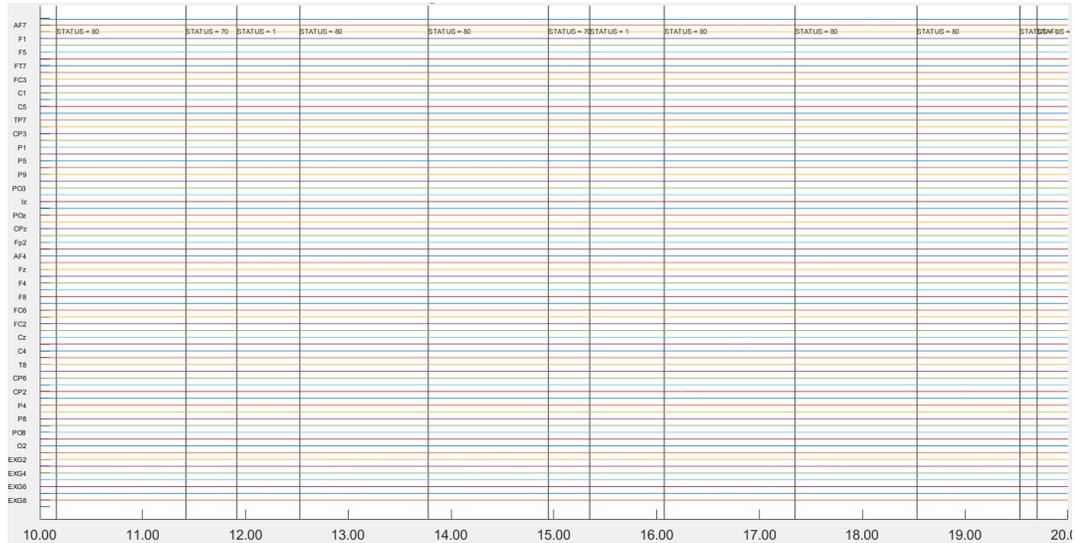


Figura 2.4: Ejemplo trial MMN

## Prueba 5: Moving Dots (MVD)

Por último está la prueba *moving dots*, en la cual el paciente se encuentra sentado frente a una pantalla de computador, en una cabina aislada y sin otro estímulo. Aquí está constantemente mirando una pantalla completamente en negro en la cual van apareciendo diferentes

escenarios, los que se componen de puntos blancos aglomerados que se encontrarán estáticos o en movimiento.

En base a lo anterior, existen 4 tipos de imágenes o escenarios posibles:

- Estática (los puntos no se mueven).
- Movimiento 100 % coherente (los puntos se mueven todos en el mismo sentido).
- Movimiento 50 % coherente (la mitad de los puntos se mueven en un sentido y la otra mitad en otro).
- Movimiento 0 % coherente (todos los puntos se mueven en sentidos diferentes).

En la tabla 2.5 se observan los *event types* para esta prueba.

Marca	Significado
55	Imagen con movimiento de 100 % de coherencia
35	Imagen con movimiento de 50 % de coherencia
25	Imagen con movimiento de 0 % de coherencia
40	Imagen estática

Tabla 2.5: Marcas o *event types* de la prueba MMN Emotional

Los marcadores de coherencia aparecen al inicio de cada segmento de señal. En la Figura 2.5 se muestra un ejemplo de trial, donde se marca al inicio con el evento 25 lo que quiere decir que en dicho trial (en la zona marcada con verde) se muestra una imagen con 0 % de coherencia. Esta imagen dura hasta que aparece la siguiente marca, en este caso 40.

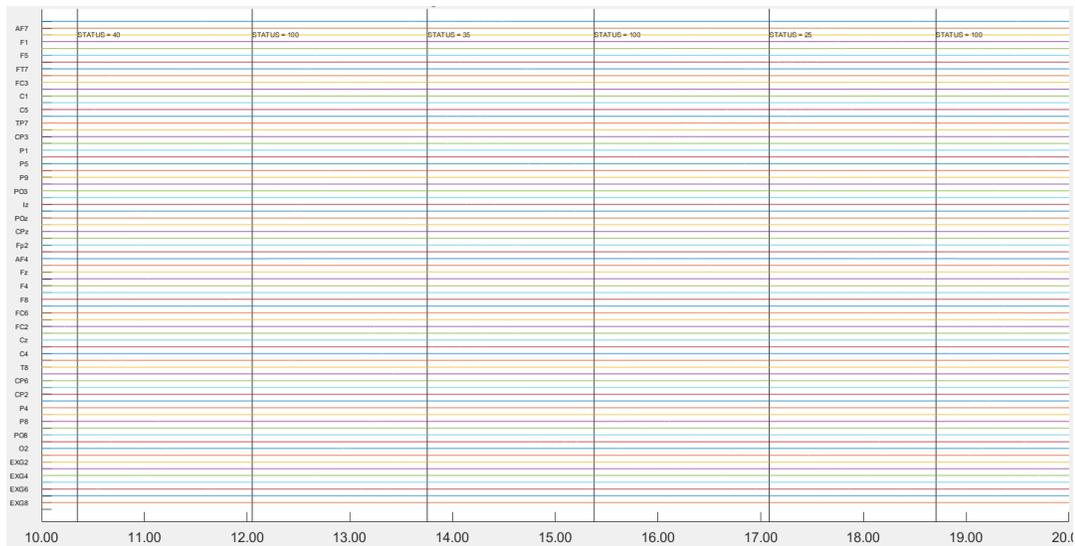


Figura 2.5: Ejemplo trial MVD

En este caso hay una marca que tampoco se considera para la tesis que es la 100.

### Selección del set de datos

Tal como se mencionó anteriormente, los datos los podemos dividir en 5 grupos de personas:

- **FEP**: Pacientes esquizofrénicos que tienen menos de 2 años desde que se generó su primer episodio psicótico.
- **APS**: Pacientes con síntomas iniciales de esquizofrenia. Estos síntomas también se pueden confundir con otro tipo de trastorno psicótico por la similaridad de los mismos, por lo que son pacientes en observación para determinar si su cuadro psicótico deriva en una esquizofrenia finalmente o no.
- **EQZ**: Pacientes con esquizofrenia crónica.
- **CNTF**: Pacientes sanos, cuya edad fluctúa entre los 19 y 30 años (grupo adulto-sano).
- **CNTA**: Pacientes sanos con una edad entre los 14 y 18 años (grupo joven-sano).

Para el desarrollo de esta tesis, no se utilizaron todos los grupo de pacientes puesto que entre el grupo **APS** eran personas que recién empezaban a presentar algún síntoma de trastorno mental pero que no necesariamente desarrollarían esquizofrenia como tal en un futuro, además presentan un período muy corto de tiempo de desarrollo de algún par de síntomas por lo que no se pueden considerar esquizofrénicos como tal.

A su vez, hay estudios que demuestran que la edad de inicio de síntomas de esquizofrenia y enfermedades mentales suele rondar alrededor de 20.5 años promedio [186]. Esto sumado

a que la edad de inicio de síntomas incide en la gravedad de los mismos, mostrando que individuos jóvenes (<18 años) tienen déficit cognitivos severos mientras que aquellos con inicio de síntomas más tardíos presentan funciones cognitivas más preservadas; lo cual comprueba que la gravedad del proceso de la enfermedad se asocia con diferentes edades de inicio, por ende hay diferencias significativas entre pacientes mayores y menores de edad [155].

Por estas razones, se determinó eliminar el grupo **CNTA** por el rango etario debido a que los grupos de pacientes esquizofrénicos (**FEP** y **EQZ**) son sobre 18 años, por lo que para hacerlos comparables se decidió descartar ese grupo y quedarse solo con **CNTF** que son adultos sanos.

## 2.2. Estado del arte

La esquizofrenia es una enfermedad mental de interés científico debido a las consecuencias que ocasiona en la persona que la sufre, siendo necesario someterse a un tratamiento para controlar sus síntomas. De mano con lo anterior, se hace imprescindible el aplicar un tratamiento lo más temprano posible dado que mientras más tardía sea la intervención, los pacientes corren mayor riesgo de sufrir complicaciones adversas [131]. Por esta razón, se considera necesario la detección precoz de pacientes con esquizofrenia [98].

Para poder determinar factores importantes que permitan diferenciar a personas con esquizofrenia de personas sanas, se han realizado varios estudios focalizados en buscar distintos indicadores que presenten diferencias significativas entre ambos grupos.

Dentro de estas investigaciones se encuentran aquellas que estudian las anomalías presentes en la conectividad cerebral de pacientes con esquizofrenia, encontrándose alteraciones en la conectividad funcional [183], así como también una disfuncionalidad frente a estímulos [208], presentando además una estructura cerebral anormal [144]. En base a lo anterior, se ha determinado la relación entre la actividad neuronal de una persona con la presencia de desórdenes mentales [160], de ahí su importancia para la detección de esquizofrenia.

Para poder hacer uso de la conectividad cerebral, esta ha sido calculada en base de imágenes de resonancia magnética [126] o señales EEG [89, 118], a partir de las cuales se permite obtener diferentes métricas relacionadas con la conectividad cerebral cuyos análisis han permitido la detección de esquizofrenia en personas [81, 58].

A su vez, otras investigaciones han determinado la contribución de los potenciales evocados en la detección de esquizofrenia [2]. Estos estudios se han encargado de determinar las alteraciones presentes en las componentes de los potenciales evocados frente a estímulos auditivos, señalando que estos presentan diferencias significativas que permiten diferenciar pacientes sanos de pacientes con esquizofrenia [85, 28, 148]; por lo cual se han utilizado fuertemente para este objetivo [151, 201, 79, 14, 210].

También se ha investigado que los potenciales evocados sufren variaciones frente a procesos de lectura [138, 63]. Así como también se ha estudiado el uso de características lingüísticas; los tópicos de los posts en redes sociales [15]; la data recolectada a partir del sensor de celular,

como por ejemplo, características relacionadas con el sueño, el habla, la movilidad, etc [202]; las bandas de frecuencia de los potenciales evocados [173] y la descomposición de las señales EEG [182, 109] para la detección de esquizofrenia.

A partir de ambos casos de estudio, se ha establecido que tanto el estudio de la conectividad cerebral [142] como el estudio de los potenciales evocados [139] sirven como potenciales biomarcadores para la detección de esquizofrenia.

En base a lo anterior, ambos parámetros se han utilizados como características en procesos de clasificación. En este ámbito, trabajos posteriores utilizan la conectividad cerebral, más específicamente la conectividad funcional, calculada tanto en estado de reposo [194, 7, 181] como durante la memoria de trabajo [22] para clasificar pacientes con esquizofrenia.

Por otro lado, también se han utilizado los componentes del potencial evocado para la predicción de esquizofrenia [200, 87], observando además, cómo afectaba en la clasificación el uso de algunos de estos potenciales (P300 y MMN) si es que se usaban de manera independiente o conjunta [115].

Cabe destacar que no se ha encontrado ninguna investigación que utilice todos los potenciales evocados en el proceso de clasificación, ni tampoco alguna donde se use tanto la conectividad cerebral como el potencial evocado como características para la clasificación. A su vez, muy pocos estudios han hecho uso de la actividad neuronal para la predicción de esquizofrenia en base a EEG [90].

Desde el punto de vista de la clasificación, también se ha estudiado el uso de señales EEG como indicador para predecir la esquizofrenia, para lo cual se han extraído diferentes parámetros que han sido usados en la confección de los modelos [27, 199]. En base a lo anterior, se ha establecido que el uso de señales EEG sirve como potencial biomarcador en la detección de esquizofrenia [61], razón por la cual se hará uso de estas señales para el desarrollo del presente trabajo.

En relación a los clasificadores, cabe resaltar que varios estudios han utilizado técnicas relacionadas con Machine Learning para discriminar entre pacientes sanos y esquizofrénicos [181, 11, 88], pero hoy en día se ha estado apuntando el uso de Deep Learning como herramienta potencial para clasificación [180], debido a los mejores resultados que estas técnicas presentan frente a las de Machine Learning [188], surgiendo la implementación de nuevas arquitecturas de clasificación basadas en métodos de Deep Learning para la detección de esquizofrenia [41, 145].

Las técnicas de Machine Learning han sido fuertemente explotadas para el desarrollo de clasificadores de detección de esquizofrenia, existiendo una gran cantidad de investigaciones con relación a su uso [71, 179, 72, 196, 26, 96, 20, 42]. Gran parte de estas investigaciones utilizan como característica base el uso de señales EEG, a partir de las cuales extraen los potenciales evocados para su análisis [71, 179, 72, 196]. Por otro lado, también hay estudios que utilizan la conectividad cerebral para generar un clasificador para esquizofrenia, la que es calculada generalmente a partir del estudio de imágenes de resonancia magnética [190, 37].

Por otra parte, hace un par de años, Deep Learning apareció como una herramienta con buen potencial para desarrollar clasificadores para detección de esquizofrenia, lo que ha

aumentado poco a poco el interés de la comunidad científica para su uso en reemplazo de Machine Learning [43].

Por dicha razón, actualmente han habido una serie de investigaciones donde se han utilizado imágenes de resonancia magnética como características de entrada a la red neuronal [84, 209, 97], donde se usa la conectividad cerebral como biomarcador para la detección de esquizofrenia. También se han empleado las señales EGG, desde las cuales se obtienen diferentes características de los potenciales evocados (tiempo, frecuencia, bandas de frecuencia, etc) para utilizarlos como características de entradas tanto de una red neuronal artificial básica como de redes neuronales LSTM [153, 174].

Por otro lado, hoy en día han aparecido un par de estudios donde se utiliza una combinación de redes neuronales para la generación de clasificadores para detectar esquizofrenia. Uno de estos estudios utiliza únicamente la conectividad neuronal, la cual es procesada con una red convolucional y una red recurrente LSTM [16, 211]. También, salió recientemente un estudio que utiliza biomarcadores funcionales y estructurales de imágenes de resonancia magnética, donde se usa un clasificador que une una red neuronal feed-forward y una red recurrente LSTM [154]. A parte de estos estudios, todos los demás utilizan únicamente un solo tipo de red neuronales para la construcción del clasificador.

Machine Learning, en parte, tiende a ser una herramienta utilizada cuando la información a usar es de baja complejidad y las tareas o automatización es menos complicada [108]. Los datos obtenidos mediante EEG son señales complejas y robustas, por lo que requieren varias épocas de entrenamiento para su correcta interpretación [161]. Por dicha razón, se establece que Deep Learning puede ser una herramienta útil, por sobre Machine Learning, para el análisis de este tipo de datos, dado que estos métodos son perfectos cuando se trata de problemas complejos como clasificación de imágenes, reconocimiento de voz y procesamiento de señales [127].

Si bien las técnicas de Deep Learning han mostrado superioridad frente Machine Learning cuando la cantidad de datos es mayor [112], también se ha señalado que funcionan mejor en problemas de este tipo, pese a tener un set de datos pequeño [198].

En base a un estudio actual, del año 2021, se logra comprender que las técnicas de Machine Learning no logran obtener todo el potencial entregado por los potenciales evocados para la detección de esquizofrenia; así como también, indica que varios estudios muestran el potencial de Deep Learning para este mismo objetivo, señalando que estos nuevos métodos pueden aprender diferentes patrones de las señales EEG que permitirían una mejor clasificación para sujetos con esquizofrenia [19].

Por otra parte, También hay estudios que demuestran que el uso de Deep Learning genera modelos con un mejor rendimiento para este tipo de problemas por sobre métodos como SVM<sup>1</sup>, Random Forest y Logistic Regression [188, 198, 189, 6], obteniendo en promedio un accuracy 5.4% superior a los de modelos tradicionales con Machine Learning [161].

Tampoco se espera que los métodos de Machine Learning reemplacen a los actuales sistemas de diagnóstico para esquizofrenia y otros trastornos psiquiátricos. En cambio, los al-

---

<sup>1</sup>Support Vector Machine

goritmos de Deep Learning pueden complementar y ayudar a discriminar esquizofrénicos de no esquizofrénicos, así como también a mejorar la identificación de biomarcadores y establecer relaciones entre las medidas del EEG y los síntomas. Estas técnicas, especialmente los algoritmos de Deep Learning, tienen el potencial de mejorar nuestra comprensión de los mecanismos neuronales subyacentes a la esquizofrenia [20].

En base a las razones anteriores, se justifica el uso de métodos de Deep Learning por sobre los tradicionales para atacar el problema de predicción de esquizofrenia.

# Capítulo 3

## Procesamiento de datos

Uno de los grandes problemas de los datos consistió en el procesamiento de los mismos, debido a que no existe un consenso ni regla general respecto a cómo deben procesarse y limpiarse, provocando que este paso dependa completamente de la decisión propia del investigador a cargo.

Esto se puede distinguir al momento de revisar el “manual” de Fieldtrip y EEGLAB respecto a sus recomendaciones y/o indicaciones para la limpieza de la data, por ejemplo, Fieldtrip realiza primero una división por trials (fragmento de señal) para luego aplicar todo el proceso de la aplicación de filtros, re-referenciado de la señal y su posterior eliminación de trials con problemas de ruido, para así aplicar el rechazo de componentes EOG; en cambio EEGLAB opta por realizar el proceso de filtrado a la señal completa y después prefiere realizar el rechazo de componentes EOG previo a la eliminación de trials [77, 69].

Por esta razón, el procesamiento que se detallará a continuación estará basado en las recomendaciones de Fieldtrip, dado que es el framework utilizado, y la orientación dada por Rocío Mayol, encargada del *PsiquisLab* y de los datos con los cuales se trabajó en esta tesis.

El código asociado a cada una de las siguientes secciones lo podrá encontrar en el siguiente [repositorio](#).

### 3.1. Preprocesamiento

El preprocesamiento de los datos consistió en la extracción de segmentos de la señal original utilizando como punto de partida el estímulo de interés, este segmento usualmente suele llamarse *trial*.

Para conformar cada trial, se utilizó como referencia lo mencionado en el libro **An introduction about event-related potential technique**, en el cual se señala que el tiempo pre-estímulo puede ir desde -100ms hacia atrás y el post-estímulo generalmente se encuentra entre 500ms hasta 1500ms [122].

En base a lo anterior, y sumado a las recomendaciones del laboratorio psiquiátrico, se fijó un tiempo pre-estímulo de 0.3 segundos y un post-estímulo de 1 segundo; conformando un trial de 1.3 segundos de longitud. Cabe destacar, que producto que los estímulos no se encuentran siempre a una distancia constante, los trial conformados pueden solaparse por una fracción de segundos.

Posterior a la obtención del trial, se aplican diferentes métodos para remover ciertos rangos de Hertz de la señal dado que estos provocan ruido en la misma, lo cual implicaría la obtención de data no prolija. Para esto se aplican dos filtros, uno pasa alto que se encarga de eliminar frecuencias muy grandes ( $>15-100\text{Hz}$ ); y un filtro pasa bajo, que permite eliminar frecuencias muy pequeñas ( $<0.01-0.1\text{Hz}$ ) las cuales suelen ser ruido de fuentes no neurales, por ejemplo, las frecuencias muy pequeñas son consideradas como ruidos producidos por el aparato electrónico que tomó las señales [123].

Por tanto, en primer lugar se aplica el filtro para eliminar todas las frecuencias inferiores a 0.1Hz, para así excluir el ruido propio del aparato electrónico. Y luego se eliminan las frecuencias superiores a 50Hz.

La razón de este último filtro se relaciona con el teorema de Nyquist para determinar el máximo de la señal. Este teorema indica que para evitar el aliasing es necesario utilizar como frecuencia tope la *frecuencia de Nyquist*, la cual se obtiene en base al doble de la frecuencia máxima de la señal EEG.

Para el caso de las señales de electroencefalograma, la cual se divide en distintas bandas de frecuencias, la máxima banda de frecuencia es de 13-30Hz lo que provoca que la frecuencia de Nyquist sea entre 26-60Hz. Esto, sumado al hecho de que en señales EEG existe ruido a partir de los 60Hz y recomendaciones del *Psiquislab*, se optó por filtrar sobre los 50 Hz para así evitar el posible aliasing generado y la mayor cantidad de ruido posible [123].

De esta forma el trial final quedaba de un rango de 1.3 segundos y una amplitud entre los 0.1 Hz hasta los 50 Hz.

Este proceso de pre-procesamiento de la data se realizaba para cada prueba. Una vez realizada esta limpieza y filtrado inicial, se da paso al último paso que permite conformar los trials, que es la identificación de los estímulos.

Para esto es necesario determinar para cada prueba los *event type* que marcan los estímulos que finalmente permitirán obtener los potenciales evocados y la conectividad cerebral de cada paciente. En base a lo anterior, los *event type* que se establecieron en cada prueba fueron los siguientes:

- **STB:** Se decidió que los event type (detallados en el Capítulo 2.1.3) que se utilizarían para determinar las características de este set de datos serían las que marcan que indican cuándo el paciente responde la pregunta realizada, que corresponden a los eventos **91 (respuesta fuera de tiempo)**, **92 (respuesta incorrecta)**, **101 (respuesta correcta)** y **102 (respuesta correcta)**
- **DBF:** Para esta prueba se consideraron como event type (detallados en el Capítulo 2.1.3) aquellos que marcan el momento en que el paciente responde la cantidad de flashes

que observó, es decir, los eventos **1 (respuesta correcta)**, **2 (respuesta correcta)**, **3 (respuesta correcta)**, **10 (respuesta incorrecta)**, **11 (respuesta incorrecta)** y **12 (respuesta incorrecta)**.

- **MMN Emocional:** En este caso, en base a los tipos de eventos detallados en el Capítulo 2.1.3, se consideró que el evento a utilizar sería el 70 (aparición estímulo infrecuente), que es el que marca el momento en que aparece el estímulo infrecuente pues esta acción viene a romper la secuencia.
- **MVD:** Se decidió que los event type (detallados en el Capítulo 2.1.3) que se utilizarían serán aquellos que marcan el momento en que se muestra cada imagen, es decir, los eventos **25 (0% coherencia)**, **35 (50% coherencia)**, **40 (imagen estática)** y **55 (100% coherencia)**. Cómo cada tipo de imagen proyectada es importante pues puede gatillar distintos eventos cerebrales en el usuario, todos estos eventos fueron considerados como principales para la conformación de los trials.

Una vez extraído el trial, se realiza una re-referenciación de los canales de los electrodos del electroencefalograma para eliminar el ruido innecesario propio de los canales. Eligiendo como método de re-referenciación como el promedio sobre todos los canales [157, 68], el cual suele ser el más utilizado en las investigaciones.

Por otra parte es importante señalar que cada señal y prueba poseen diferente cantidad de trials, dado que depende de la cantidad de eventos o estímulos elegidos para obtenerlos y la longitud de la señal misma, la cual se relaciona con el tiempo de duración de la prueba de un determinado sujeto.

## Resting state

Para el caso de *resting state* el proceso fue diferente, dado que en estado basal el individuo no se encuentra bajo ningún estímulo por lo cual no es posible extraer trials de la señal y, por tanto, tampoco es posible obtener ERP en base a esta prueba.

Es importante recordar que esta prueba consistía en que el sujeto estuviese 2 minutos con los ojos abiertos, sin hacer nada, y luego otros 2 minutos con los ojos cerrados. El proceso de toma de esta prueba fue continuo, sin ningún indicador que permitiese identificar el momento en que el sujeto estaba con los ojos abiertos o cerrados. Además, haciendo una inspección de la señal se encontró que la longitud de onda no era la misma para todos los sujetos y tampoco correspondía a los 4 minutos que debiese haber durado la prueba en totalidad, sino que algunos presentaban más segundos de duración.

Para esta prueba los datos se guardaron de dos formas diferentes, algunos sujetos tenían sus señales divididas en dos archivos,  $X\_R1$  que almacena la información para el momento en que el sujeto  $X$  estuvo con los ojos abiertos; y  $X\_R2$  que contiene la señal obtenida para el momento en que el sujeto  $X$  estuvo con los ojos cerrados. Pero por otro lado, la data del paciente podía encontrarse toda junta en un único archivo.

Para abordar este último caso, y lograr separar la data, se tuvo que dividir la señal en dos

partes iguales, donde la primera mitad de la señal correspondería al caso “ojos abiertos” y la segunda mitad sería el caso “ojos cerrados”. Para hacer esta división simplemente se tomó el instante de tiempo donde iniciaba la señal (*tiempo mínimo*, que era siempre 0) y el instante de tiempo donde terminaba la señal (*tiempo máximo*, el cual era variable).

Con esta información se calculaba el tiempo medio y así se podían extraer los dos intervalos de señal, siendo el primero desde el tiempo mínimo hasta el tiempo medio mientras que el segundo correspondía al intervalo de señal que se encontraba desde el tiempo medio hasta el tiempo máximo de la señal original.

Tras la primera parte, se tendría por separado la información obtenida para ojos abiertos y ojos cerrados, pero había un problema dado que la longitud de onda para cada caso eran diferentes, lo que sumado a que entre los individuos también existía discrepancia con el largo de onda, generaba que los datos que se tenían no fuesen 100 % comparables entre sí puesto que una longitud de onda más larga es equivalente a más data y, por ende, más información se puede extraer de la señal obteniendo sujetos con más datos que otros.

Por otra parte, existía un problema más relevante producto de la división de una misma señal en dos partes iguales para extraer la información del sujeto con los ojos cerrados y abiertos, el cual correspondía a que el método utilizado no aseguraba que en todo el primer segmento extraído de la señal hubiese solo información del sujeto con los ojos cerrados y, por tanto, tampoco aseguraba que en la segunda mitad solo se tuviese información para el caso de ojos abiertos.

Lo anterior se daba dado que el largo de la onda no era siempre el esperado, sino que en varias ocasiones superaba los 4 minutos que debiese haber durado, lo que podía implicar que el sujeto estuvo más tiempo con los ojos cerrados o con los ojos abiertos, o que esto podía ser producto a que el investigador que realizó la prueba empezó a grabar segundos antes de empezar la prueba como tal o simplemente cortó “la grabación” segundos después de haber terminado la prueba.

Independiente del porqué, sí se podía asegurar completamente que en el centro o mitad de cada segmento se iba a encontrar solamente la información correspondiente al instante en que el sujeto estuvo con los ojos abiertos y cerrados. Por esta razón, se optó por sacar los extremos de cada segmento de la señal para así extraer solo la información central.

El tamaño de la ventana fue completamente arbitrario, pensando que la señal debía tener un máximo de 2 minutos por segmento (120 segundos) dado que es lo que debían durar las pruebas, se optó por extraer los 90 segundos centrales de cada segmento de señal utilizado provocando la eliminación de los extremos del segmento, generando finalmente que la data total tuviese una longitud de 180 segundos (90 segundos para ojos abiertos y 90 segundos para ojos cerrados).

Esto se replicó para ambos casos de almacenamiento de la data mencionados anteriormente para así tener data comparable y evitar los problemas ya señalados.

Una vez extraída la data con la cual se iba a trabajar, se procedió de forma similar que para las otras pruebas, exceptuando el hecho de la identificación de eventos para la generación de trials puesto que estos no existían.

Luego se aplicaron los mismos filtros pasa alto y pasa bajo que para el resto de las pruebas, y finalmente se realiza la re-referenciación de los canales del electroencefalograma para la nueva data. Todo esto se aplicó de igual forma para cada segmento de la señal (caso ojos abiertos y caso ojos cerrados).

## 3.2. Filtrado de señal

Una vez finalizado el pre-procesamiento de la señal, se dió paso a 2 nuevas instancias de filtrado de señal. La primera consistía en una inspección visual de la señal resultante, la cual permite encontrar inconsistencias que debían ser removidas. Mientras que la segunda es la obtención de componentes a partir de la técnica ICA (*Independent Component Analysis*), para posteriormente eliminar los componentes que presentaban problemas.

### 3.2.1. Inspección y limpieza visual

Tras el pre-procesamiento se dió paso a la inspección visual de la señal para eliminar aquellos trials que presentaban problemas o ruido notorio, puesto que de otra forma se generaría data inconsistente.

Para realizar esto se tuvo que ir trial por trial viendo si presentaban algún problema o no. Para poder hacer este análisis se consultó con Rocío Mayol Troncoso (investigadora del PsiquisLab a cargo de los datos y psicóloga clínica). Utilizando como guía su experiencia y lo señalado en la página oficial de Fieldtrip (framework utilizado para todo el procesamiento y limpieza de datos), se determinó que lo fundamental de esta limpieza visual era eliminar el ruido generado a partir de posibles movimientos del aparatos o mal toma de la señal.

Para esta fase es necesario señalar que cada canal genera un señal la cual debe ser una línea prácticamente recta, sin movimientos diagonales o verticales, por lo que cualquiera de estos movimientos o la generación de choques entre señales de distintos canales son comportamientos anormales tales como:

- Saltos repentinos de la señal (Figura 3.1 (c)), más conocidos por su nombre en inglés *jump*.
- Canales cruzados en la señal (Figura 3.1 (a) y (b)).
- Movimiento notorio producto de movimientos musculares, el cual es el que comúnmente se le conoce como ruido (Figura 3.1 (d)).

Cabe destacar que si el ruido era muy pequeño en comparación con la longitud total del trial o muy poco notorio, no se realizaba la eliminación de este. Así mismo, si el problema se encontraba en un único canal o en los canales EXG (canales no pertenecientes al electroencefalograma como tal puesto que estos son electrodos colocados en el contorno de los ojos y en la zona posterior de las orejas) o en el canal Status (canal generado automáticamente

tras el pre-procesamiento de la señal y que tampoco se utilizaría para obtener la información final para el proceso de clasificación), entonces el trial tampoco era eliminado.

De esta forma se evitaba la eliminación innecesaria de información, puesto que mientras más trials se eliminaban, más información se perdía. Por dicha razón también se estableció no eliminar más del 15% del total de trials de la señal, decisión basada en la teoría señalada en el libro *An Introduction about event-related potential technique* donde se indica que para la eliminación de artefactos, el rechazo de trials debe estar entre el 10% y 25% del total de trials originales [124].

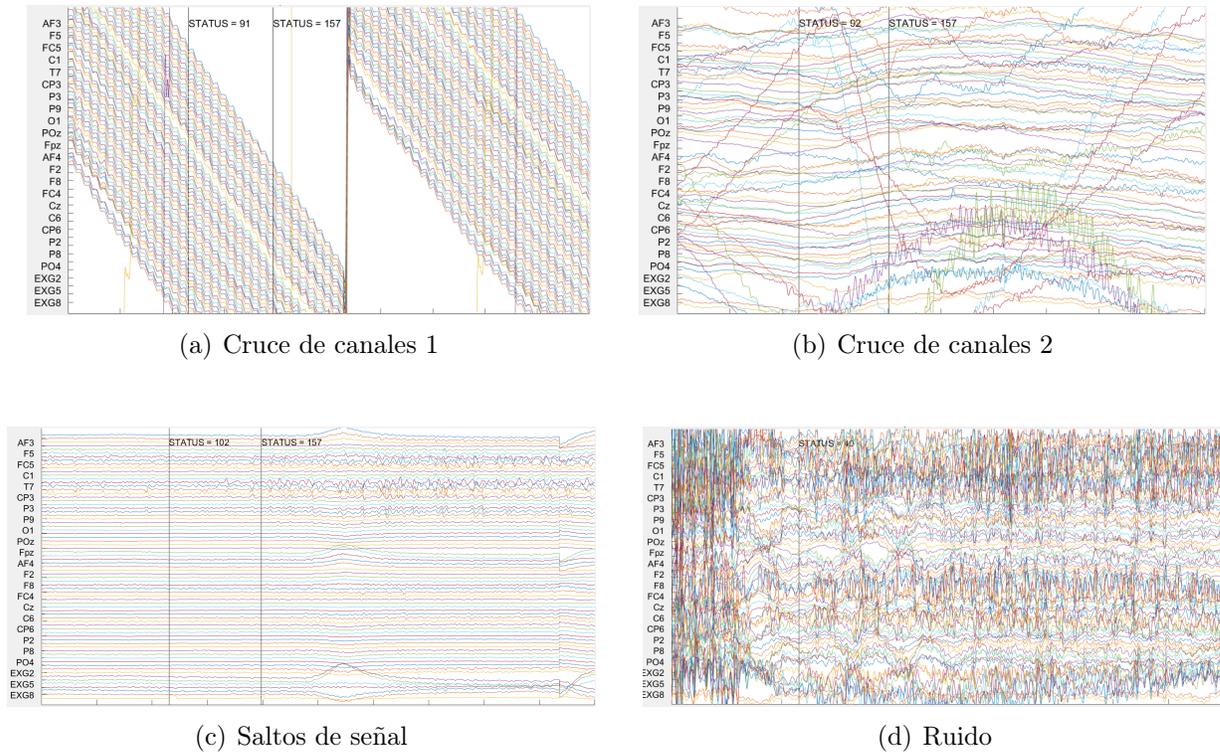


Figura 3.1: Ejemplo de problemas en la señal

Antes de entrar a detallar los trials eliminados y las razones del porqué, es necesario poner en contexto el universo de personas con las cuales se trabajará por cada prueba y cantidad de trials obtenidos por cada uno de los pacientes, el cual se puede apreciar en la tabla 3.1.

Dado el objetivo de la tesis, de obtener biomarcadores para la detección de la esquizofrenia utilizando para esto un modelo de clasificación, hubieron grupos que fueron excluidos de la data total como lo son el grupo APS y CNTA tal y como se detalló en el capítulo 2.1.3.

Por lo anterior, se puede reducir el grupo con el cual se trabajó el análisis visual y todo el procesamiento posterior al señalado en la tabla 3.2.

Prueba	APS	FEP	EQZ	CNTA	CNTF
Resting State	25	16	12	10	36
STB	19	18	0	9	18
DBF	16	11	2	1	14
MMN Emotional	10	0	0	2	15
MVD	16	1	0	8	13

Tabla 3.1: Sujetos por prueba y grupo

Prueba	Total	FEP	EQZ	CNTF
Resting State	64	16	12	36
STB	36	18	0	18
DBF	27	11	2	14
MMN Emotional	15	0	0	15
MVD	14	1	0	13

Tabla 3.2: Sujetos por prueba y grupo

Haciendo un análisis preliminar de los datos, es posible ver que las pruebas MMN Emotional y MVD presentan a lo más 1 paciente esquizofrénico, lo que conlleva a eliminar estas dos pruebas por completo.

Lo anterior se fundamenta en dos razones principales. Por una parte, la única prueba donde podría haber sido factible considerar un balanceo de clases era en la MVD, pero dado que solo cuenta con 1 dato para la clase de esquizofrenia, si se realiza un oversampling eso implicaría repetir únicamente dicho dato, lo cual no tiene sentido dado que el clasificador generaría overfitting puesto que se aprendería muy bien que solo frente a dicho dato hay presencia de esquizofrenia, clasificando la mayoría de los datos como no esquizofrénico.

Por otra parte, en caso de querer hacer undersampling, eso implicaría eliminar casi todos los datos de pacientes sanos, quedando muy pocos datos para dicha prueba lo cual nuevamente llevaría a construir un mal clasificador.

La segunda razón es que dichas pruebas tampoco sirven para ser consideradas en conjunto con los datos de las demás pruebas, dado que en caso de hacerlo (y suponiendo que el desequilibrio de clases no sea significativo) habría un problema para establecer qué tanto aportaron estas pruebas (y por ende sus features) en la determinación de si un paciente tiene o no esquizofrenia, dado que solo hay datos para pacientes sin esquizofrenia

Lo anterior conllevaría a que el clasificador tomara la tendencia de decir que para dichas pruebas todo dato ingresado pertenece a la clase “no esquizofrénico”. Esto podría generar un clasificador con un alto *accuracy* sin que este presentase overfitting, pero sería un clasi-

ficador que prácticamente indicaría que si una persona realiza la prueba MMN Emotional o MVD sería inmediatamente clasificada como “no esquizofrénico” siendo una falsa aseveración, provocando una inconsistencia en los análisis y resultados que se obtendrían con esta tesis.

A partir de lo anterior, el análisis resumido fue el siguiente:

### **Stemberg (STB)**

De los 18 pacientes sanos 17 presentaban al menos un trial con problema; esto mismo ocurrió con los 18 pacientes esquizofrénicos. Por otra parte, ningún dato fue eliminado por completo dado que ninguno presentaba problemas en más del 15 % de sus trials.

El análisis realizado se puede encontrar en los anexos, en la lista [6.1](#) y la lista [6.2](#).

### **Double Flash (DBF)**

De los 14 pacientes sanos 9 presentaban al menos un trial con problema; mientras que de los 13 pacientes esquizofrénicos solo 7 presentaron problemas en al menos un trial. Por otra parte, ningún dato fue eliminado por completo dado que ninguno presentaba problemas en más del 15 % de sus trials.

El análisis realizado se puede encontrar en los anexos, en la lista [6.3](#) y la lista [6.5](#).

### **Resting State (RS)**

Para este caso se analizó por separado las 2 pruebas que se realizaron en estado de reposo. Por esta razón un mismo sujeto tendrá dos análisis, siendo el con terminación *R1* aquel relacionado al momento en que el sujeto estuvo con ojos cerrados, mientras que la terminación *R2* será para el momento en que el sujeto estuvo con ojos abiertos.

Para esta prueba, debido al pre-procesamiento realizado sobre el set, el cual consistió en la extracción central de 90 segundos de la señal, el largo de cada señal para todos los sujetos fue exactamente el mismo y por tanto la cantidad de trials que se extrajeron fue el mismo para cada caso.

Tal como se mencionó previamente, el estado basal no cuenta con ningún estímulo por tanto los trials no se podían extraer en base a estos. En este caso, para la extracción y conformación de cada trials se optó por dividir la señal en segmentos de 1.3 segundos.

La razón de la división de la señal en trials se debe al hecho de que así era factible eliminar problemas de la señal sin que esta quedase con vacíos, puesto que en caso de no dividirla todo error encontrado en la señal debía ser seleccionado y removido, generando que dicha sección eliminada de la señal quedase sin información, formando una señal discontinua. Al conformar trials, si una parte de la señal presentaba problemas entonces este se eliminaba por completo, evitando así este corte de señal.

Por otra parte, la razón de formar trials de 1.3 segundos está fundamentada en el hecho de que el largo de los trials de las otras pruebas era exactamente de 1.3 segundos. En caso de elegir un tamaño de trial diferente provocaría que el cálculo de la conectividad cerebral de resting state tuviese diferente dimensionalidad con respecto al resto de las pruebas, debido a que esta calcula las frecuencias en base al largo temporal de los trials, provocando que a mayor cantidad de segundos del trial mayor es la dimensionalidad de la conectividad extraída y más compleja se torna su clasificación.

Tras esto, de cada archivo se obtuvieron 69 trials en total, razón por la cual no era factible eliminar más de 10 trials, correspondientes al 15 % del total obtenido.

De los 36 pacientes sanos, 14 de ellos presentaban problemas en al menos un trial mientras que de los 28 pacientes esquizofrénicos, 11 de ellos presentaban problemas en al menos un trial. Por otra parte, se eliminaron 3 datos (1 paciente sano y 2 esquizofrénicos) dado que presentaban problemas en más del 15 % de sus trials.

El análisis realizado se puede encontrar en los anexos, en la lista [6.6](#) y la lista [6.4](#).

### 3.2.2. ICA y rechazo de componentes

Una vez finalizada la limpieza visual de cada uno de los trials obtenidos a partir de las señales de cada prueba, el siguiente paso consta en la aplicación de un método llamado ICA (Independent Component Analysis), el cual consiste en descomponer un dataset complejo en múltiples subpartes independientes unas de otras [[212](#), [76](#)].

Se ha demostrado su eficacia al utilizarse para detectar, separar y eliminar de forma eficaz los artefactos oculares de incluso registros de EEG muy contaminados, siendo incluso más favorables que los que se obtienen con otros métodos existentes para realizar esta misma labor, como lo son métodos de regresión y PCA (Principal Components Analysis) [[106](#), [107](#), [191](#)].

Por este motivo, el objetivo principal del uso de este método para el procesamiento de señales EEG es eliminar ruido severo generado producto de movimientos oculares (también conocidos como sacadas) y parpadeos, dado que estos representan un grave problema para el análisis e interpretación de la data proveniente tanto del electroencefalograma como de las características a utilizar.

El método de ICA se utilizará, por tanto, para la detección de artefactos EOG (electro-oculograma), es decir, aquellos artefactos provenientes del movimiento de ojos o sacadas y el pestañeo. Para esto se realiza una inspección visual manual de los componentes independientes generados por el ICA (Figura [3.2](#)), con la finalidad de detectar cuál de estos presenta los artefactos que se buscan eliminar.

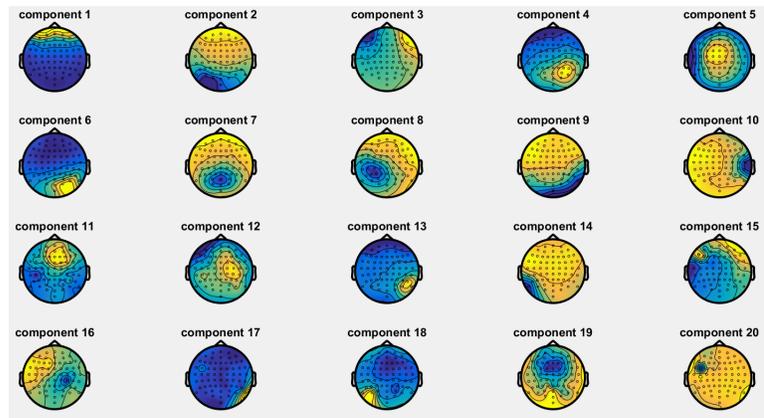
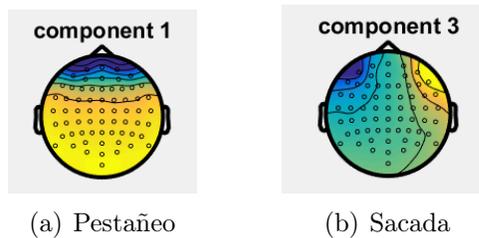


Figura 3.2: Ejemplo de ICA

Los artefactos a eliminar presentan una forma característica y reconocible una vez que se aplica el ICA. En la Figura 3.3.a se puede apreciar la imagen visual del mapa de calor que se conforma para un componente que representa un pestañeo, mientras que en la Figura 3.3.b se observa la imagen visual para un componente que representa una sacada de izquierda (movimiento ocular hacia la izquierda) y de derecha (movimiento ocular hacia la derecha). Si algún componente generado presentaba alguna de estas dos “formas”, entonces era eliminado.



(a) Pestañeo

(b) Sacada

Figura 3.3: Artefactos EOG

Al igual que para el caso anterior donde se hacía limpieza de la señal por trials, en este caso tampoco uno puede eliminar gran cantidad de componentes puesto que de igual forma se pierde información relevante de la señal. Pese a esto, la decisión de cuántos componentes rechazar no es unánime por lo que depende del investigador a cargo; en este caso, dado que se siguen los lineamientos del Psiquilab y se está bajo la tutela de Rocío Mayol para el manejo y limpieza de datos, se determinó que cómo máximo se pudiesen eliminar 3 componentes puesto que esta era la manera en que ella trabajaba la limpieza de datos.

Para llevar a cabo la aplicación del método ICA y el rechazo de componentes se siguieron los pasos indicados en el framework utilizado para el procesamiento de los datos.

En primer lugar, tras la eliminación de trials con problemas de ruido y como preparación para la aplicación de este método, se realiza un downsampling de la señal a 256 Hz, lo cual permite reducir la complejidad computacional permitiendo obtener una menor cantidad de características para los modelos de clasificación.

Posteriormente se aplica el método ICA incorporado en Fieldtrip para lo cual se realiza la eliminación de los canales no pertenecientes al electroencefalograma, es decir los canales

EXG, status y el canal de referencia que se conforma tras el preprocesamiento inicial de los datos, de tal manera que solo se consideren los canales del 1 al 64. Con esto Fieldtrip realiza el método ICA generando los componentes independientes a partir de la señal.

Como una breve explicación al funcionamiento de este método es importante señalar que el ICA trabaja bajo dos suposiciones, primero que los componentes a buscar deben ser estadísticamente independientes es decir, que la información sobre un componente  $x$  no te da información sobre un componente  $y$  y viceversa; y segundo, no deben tener una distribución gaussiana. Luego, para descubrir cuales son los componentes independientes utiliza medidas no gaussianas y ambos supuestos anteriores para maximizar la no gaussianidad, lo cual se traduce en una resolución de un problema de optimización [193, 60].

Una vez finalizado el cálculo de los componentes se da paso a la inspección visual de cada uno para determinar cuales eliminar. Para esto se realiza una visualización de los componentes obtenidos por el método, para luego ser inspeccionados y determinar cuales de estos presentan artefactos EOG (detallados en la Figura 3.3).

Es importante señalar que durante la ejecución de este método se tuvo problemas con un par de datos, puesto que el método no convergía y por tanto no se podían definir los componente independientes. Luego de analizar el problema se descubrió que esto se debía a que tras el rechazo de trials a veces existía un canal extra (el canal de referencia) provocando una inconsistencia en la cantidad de canales, por tanto para arreglar el problema netamente se tenía que señalar en la función la cantidad de canales que debían ser usados para la obtención de los componentes independientes. En esta caso, la cantidad de canales aptos para ser usados eran 64.

A continuación se realizará un resumen respecto al rechazo de componentes para cada prueba. El detalle completo se puede encontrar en la sección de anexos.

### **Stemberg (STB)**

De los 18 pacientes sanos se realizó un rechazo de componentes para 16 de ellos. Para el caso de los pacientes esquizofrénicos, se realizó un rechazo de componentes para 14. Por otra parte, no se observaron inconvenientes mayores que provocaran la eliminación de algún dato.

El análisis realizado se puede encontrar en los anexos, en las listas 6.7, 6.8 y la lista 6.9.

### **Double Flash (DBF)**

De los 14 pacientes sanos, se realizó un rechazo de componentes para 14 de ellos. Para el caso de los pacientes esquizofrénicos, se realizó un rechazo de componentes para 10 de los 13 sujetos. Por otra parte, no se observaron inconvenientes mayores que provocaran la eliminación de algún dato.

El análisis realizado se puede encontrar en los anexos, en la lista 6.11 y la lista 6.10.

## Resting State (RS)

De los 36 pacientes sanos, se realizó un rechazo de componentes para 25 de ellos. Para el caso de los pacientes esquizofrénicos, se realizó un rechazo de componentes para 15 de los 28 sujetos. Por otra parte, se encontraron problemas en el cálculo general del ICA para 3 pacientes (2 sanos y 1 esquizofrénico) producto de la presencia de canales malos en la señal; por dicha razón se eliminaron esos datos del set de datos para la clasificación.

El análisis realizado se puede encontrar en los anexos, en las listas [6.12](#), [6.13](#) y las listas [6.14](#), [6.15](#).

## 3.3. Extracción de características

Tras haber terminado con la aplicación del método ICA y el rechazo de componentes correspondientes a artefactos EOG, se procede a la extracción de las características necesarias para la conformación del set de datos para la etapa de clasificación.

### 3.3.1. Potencial Evocado (ERP)

Los potenciales evocados son señales eléctricas producidas por el sistema nervioso en respuesta a un estímulo sensorial externo, las cuales son detectadas por un electroencefalograma (EEG), electromiograma (EMG) u otro método electrofisiológico.

Por otra parte, las respuestas de potencial evocado (evoked-response potential o ERP) son fluctuaciones significativas de voltaje, resultantes de la actividad neuronal evocada. Estas señales eléctricas son usadas para el diagnóstico y monitoreo de enfermedades y se caracterizan por la amplitud, polaridad, inter-peak, latencia y la morfología de la señal [[120](#), [1](#), [70](#), [105](#)].

Para la obtención de estos ERPs se realizó lo siguiente para un sujeto [[73](#)]:

1. Primero se elije uno de los eventos de la prueba y se extraen todos los trials relacionados para dicho evento.
2. Luego se les aplica un filtro pasa bajo a 40 Hz, principalmente para reducir un poco más el ruido restante en la señal (suavizar la onda) dado que con el filtrado que presentaba provocó que los primeros ERP calculados tuviesen ruido notorio en la señal presentada.
3. Tras esto se hace el cálculo del ERP para los trials relacionados a dicho evento. El cual, como se mencionó previamente, se obtiene mediante el promedio de la señal cerebral utilizando una función propia de Fieltrip para esto.
4. Una vez obtenido el ERP para el evento asociado al trial, se realiza una corrección de línea base con la finalidad de hacer comparables los distintos datos o sujetos.

Para la elección de la línea base, se recomienda que su período sea de a lo menos 20% de la duración total del trial [[122](#)]. En este caso el trial tiene una duración de

1300ms, por tanto la línea base debe tener a lo menos una duración de 260ms (20 % de 1300); además, se recomienda que la duración de la línea base sea múltiplo de 100ms para poder completar un ciclo completo de la banda gamma asociada a las señales EEG [122]. Por esta razón, se optó por una línea base de 300ms de duración.

Para esto, la línea base se calcula en un cierto rango previo al estímulo, en este caso se tomó desde el inicio de la señal (-0.3 segundos previos al estímulo) hasta la aparición del estímulo (0 segundos previos al estímulo), con esto se calcula el promedio de esa fracción de señal de todos los trials obtenidos y se hace la corrección al resto de la señal y trials.

5. Finalmente se extrae el ERP correspondiente, que es uno por evento, y posteriormente se repite el mismo procedimiento para los demás eventos. Cabe destacar que además cada canal del EEG genera un ERP distinto.

Estos pasos se repiten por cada sujeto y prueba. El detalle de la cantidad de ERPs obtenidos por prueba se detalla en la tabla 3.3.

Prueba	Sujetos	Eventos/ERPs por sujetos	ERPs totales
STB	36	4	144
DBF	27	6	162

Tabla 3.3: Detalle de ERP

En total se tienen 306 ERPs, que corresponden a 10 tipos de ERPs diferentes (uno por cada evento).

Es importante recordar que para resting state no existen eventos, razón por la cual no se pueden extraer ERPs pues estos sí o sí necesitan de un estímulo para que se manifiesten.

Los ERPs se calculan utilizando la función *ft\_timelockanalysis* de Fieldtrip [75], la cual retorna un objeto cuyo campo más importante es el llamado *avg*. Este campo es una matriz de  $N \times M$ , con  $N$  la cantidad de canales utilizados del electroencefalograma y  $M$  corresponde al total de puntos temporales existentes. Finalmente, el valor de la casilla  $(i, j)$  corresponde a la frecuencia de la señal en el tiempo  $j$  para el canal  $i$ .

Lo anterior implica que se obtiene una señal por cada canal, cada uno de estos corresponden a un ERP del mismo tipo. Esta matriz será la utilizada para procesar los ERP en el clasificador.

Para este caso, los ERP tienen una dimensionalidad de  $64 \times 333$ . Además, es importante señalar que cada ERP recibió el nombre del estímulo en base al cual se obtuvieron, por ejemplo el *ERP N* se obtuvo bajo el evento  $N$ .

### 3.3.2. Conectividad Cerebral (CC)

La conectividad cerebral se refiere a un patrón de enlaces anatómicos, de dependencias estadísticas o interacciones casuales entre distintas unidades dentro de un sistema nervioso. Esas unidades corresponden a neuronas individuales, poblaciones de neuronas o zonas cerebrales anatómicamente separadas. La actividad neuronal está limitada por la conectividad, la cual es crucial para dilucidar como las neuronas y redes neuronales procesan la información.

Existen varios tipos de conectividad neuronal, para ser más específico, la conectividad funcional, anatómica y efectiva. La conectividad anatómica se refiere a una red de conexiones físicas o estructural (sinápticas) que unen conjuntos de neuronas o elementos neuronales, así como sus atributos biofísicos estructurales.

La conectividad funcional es fundamentalmente un concepto estadístico. En general, la conectividad funcional es definida como la dependencia temporal de regiones cerebrales anatómicamente separadas. La dependencia estadística puede ser estimada midiendo la correlación o covarianza, la coherencia espectral o phrase-locking. La conectividad funcional es usualmente calculada entre todos los elementos del sistema, independientemente de si estos elementos están conectados por enlaces estructurales directos.

Por último, la conectividad efectiva puede ser vista como la unión entre la conectividad estructural y funcional, y es descrita como la influencia que ejerce un sistema neuronal sobre otro, esto refleja la interacción causal entre la actividad de áreas cerebrales.

La usada para el desarrollo de esta tesis es la conectividad funcional, la cual puede ser calculada directamente usando FieldTrip que utiliza diferentes métodos estadísticos, por lo que hay que definir cuál de esos se desea usar para realizar el cálculo de la conectividad neuronal. El detalle de todos los métodos se puede encontrar en la sección de anexos en la tabla 6.20.

Los métodos más conocidos y utilizados son “coherence (coh)” y “correlation” (corr) siendo la coherencia (coh) la más usada, razón por la cual esta fue la empleada en Fieldtrip para ser calculada durante la obtención de la conectividad funcional. La coherencia es una herramienta utilizada para encontrar la relación o similaridad entre las regiones cerebrales y analizar las convulsiones que se producen entre dichas regiones. La correlación ( $C_{xy}$ ) es calculada por medio de la siguiente fórmula matemática [55, 203, 110]:

$$C_{xy}(f) = \frac{G_{xy}(f)^2}{G_{xx}(f)G_{yy}(f)}$$

Donde  $G_{xy}(f)$  es la densidad espectral de potencia cruzada (cross-power spectral density) entre la señal  $x$  e  $y$ , mientras que  $G_{xx}(f)$  y  $G_{yy}(f)$  son las densidades autoespectrales de  $x$  e  $y$  respectivamente.

La coherencia siempre es un valor entre 0 y 1, si el valor es 0 entonces las señales no están correlacionadas pero si el valor es 1, entonces las señales están fuertemente correlacionadas, es decir, mientras más cercano a 1 sea el coeficiente de correlación, entonces más fuertemente correlacionadas están las señales.

La conectividad cerebral fue calculada utilizando la fórmula *ft\_connectivityanalysis* de FieldTrip, la cual se encarga de computar varios tipos de conectividad entre los diferentes canales del electroencefalograma en una relación todos a todos, generando un tensor con la conectividad entre todos los canales para diferentes frecuencias. Para poder aplicar correctamente la fórmula para obtener la conectividad cerebral se realiza lo siguiente:

- Para el caso de las pruebas con presencia de estímulo, se deben diferenciar los trials según el evento o estímulo que lo caracteriza. Por tanto, lo primero es dado un evento, obtener todos los trials relacionados con este.

Para el caso de resting state, dado que no cuenta con ningún estímulo, netamente se seleccionan todos los trials generados.

- Una vez obtenidos los trials se realiza un análisis de frecuencia y tiempo-frecuencia en datos de series de tiempo en los trials obtenidos, usando para esto la función *ft\_freqanalysis* [74].

Este análisis se aplica sobre el espectro completo para toda la longitud de los datos y utiliza el método “multiple taper” o “multitaper” (multiple taper method o MTM), el cuál está basado en las secuencias de Slepian o también conocidas como DPSS, para la obtención del espectro de Fourier, pues este es necesario cuando se trabaja en el cálculo de la conectividad cerebral bajo el método estadístico “coherence” (coh) [205, 184, 204].

- Luego, se corrige por línea base los datos; para esto se sigue el mismo criterio señalado para el cálculo de los ERP, obteniendo una línea base de 300ms (de 0 ms a 300ms pre-estímulo).

Para el caso de los sujetos de la prueba RS, no se realizó corrección por línea base dado que no había un estímulo existente para calcularla.

- Finalmente se calcula la conectividad cerebral bajo el método “coherence”.

El tensor generado tiene una dimensionalidad de  $N \times N \times F$ , con  $N$  la cantidad de canales utilizados del electroencefalograma y  $F$  la cantidad de frecuencias posibles. Este tensor corresponde al campo *freq* del objeto que es retornado por la función *ft\_connectivityanalysis*. Para este caso, la dimensionalidad de la conectividad es de  $64 \times 64 \times 168$ .

Cabe destacar que para el caso de los sujetos relacionados con la prueba Resting State (RS), se obtuvieron 2 conectividades, una para el caso de ojos abiertos y la otra para el caso de ojos cerrados. Mientras que para el resto de las pruebas se obtiene 1 conectividad cerebral por cada evento asociado a la prueba.

Por otra parte, es importante señalar que para obtener la conectividad cerebral no se realizó una nueva corrección de la señal (como para el caso del cálculo de los potenciales donde se aplicó un filtro pasa bajo a los 40Hz); esto fue debido a que pueden existir conectividades a frecuencias mayores dependiendo del sujeto.

Además, es importante señalar que cada Conectividad Cerebral (CC) recibió el nombre del estímulo en base al cual se obtuvieron, por ejemplo la *CC N* se obtuvo bajo el evento  $N$ .

# Capítulo 4

## Clasificación

A continuación se explicará todo lo relacionado con el proceso de clasificación de los datos, partiendo por los criterios utilizados para la conformación del dataset, para luego dar paso al diseño del modelo de clasificación y finalizar con el detalle de las diferentes pruebas de clasificación realizadas.

El código asociado a cada una de las siguientes secciones lo podrá encontrar en el siguiente [link](#).

### 4.1. Conformación del Dataset

Para constituir el set de datos inicial ser usado en la clasificación, en primer lugar se seleccionaron todos los datos de prueba que no presentaban problemas de consideración durante la etapa de limpieza de la señal es decir, que cumplieran con el criterio impuesto de no contar con más de un 15 % de los trials con problemas de ruido o canales defectuosos [125].

Posteriormente, se realizó un análisis de los datos seleccionados con el objetivo de determinar la proporción de clases en cada una de las pruebas. Para esto hay que tener en consideración dos puntos importantes, el primero consiste en que cada prueba se tomó como un set de datos independientes de las otras pruebas. Esto dado que cada una de estas pruebas genera datos que son tomados bajo diferentes condiciones, por lo que durante el proceso de clasificación también deberán evaluarse de forma independiente para comprobar la incidencia que tiene cada prueba realizada en la detección de la esquizofrenia.

El segundo punto importante es indicar que las clases que se consideraron para el proceso de clasificación, y con las cuales se realizó el análisis de clases, son las siguientes:

- Esquizofrénico (clase 1): La cuál estaba conformada por los datos que pertenecían a los sujetos *EQZ* y *FEP*.
- No esquizofrénico (clase 0): La cuál estaba conformada por los datos pertenecientes a los sujetos *CNTF*.

En el gráfico 4.1 se puede visualizar la proporción de datos por clase para cada prueba.

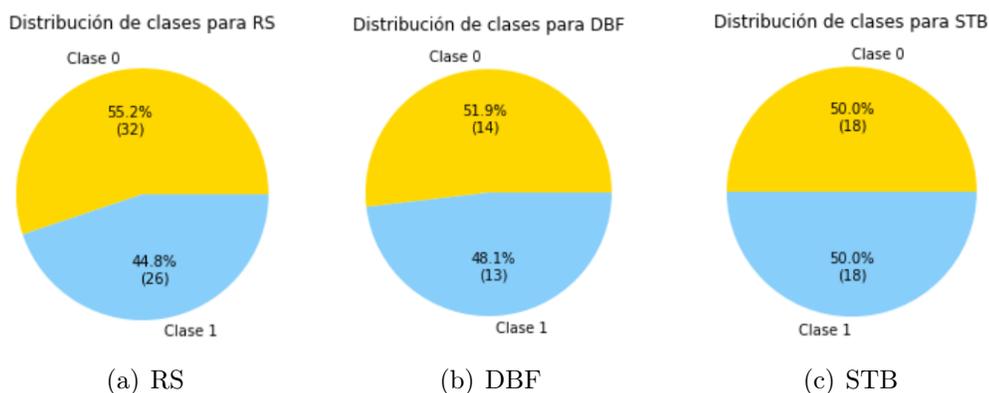


Figura 4.1: Proporción de clases por prueba

No se utilizó ninguna técnica para equilibrar las clases dado que no fue necesario. Eso sí, se consideró que podría ser un problema el pequeño desajuste existente en la prueba de resting state puesto que si bien la proporción de clases no representa un desequilibrio, sí podría generarse un problema dada la poca cantidad de datos de dicha prueba.

Por esta razón, se realizaron ensayos para analizar el impacto que tenía el asignar un “puntaje” a la clase minoritaria (clase 1); pero los resultados empeoraban bastantes dado que bajaba la detección para la clase 0 y aumentaba bastante el error de detectar esquizofrenia en pacientes sanos, pese a que el puntaje asignado era muy pequeño (0,9 para la clase 0 y 1,1 para la clase 1).

Ahora, en caso que existiera un desequilibrio de clases al momento de considerar los datos en conjunto con los de otras pruebas, entonces habría que hacer oversampling o undersampling para equilibrarlas, el problema con esto es que si se quiere replicar o eliminar datos no se podrían replicar o eliminar datos de la misma prueba (por los problemas mencionados anteriormente) sino que tendría que ser de otra prueba, lo cual no tiene sentido dado que son features completamente diferentes puesto que los estímulos de cada prueba es distinto y las features se obtienen a partir de estos, y se podrían estar perdiendo features realmente importante generando un clasificador con sesgo.

Lo que conlleva a que el set de datos final esté compuesto por:

Prueba	Total datos por clase	Total de datos
Resting State (RS)	32 (clase 0) 26 (clase 1)	58
Stemberg (STB)	18 (clase 0) 18 (clase 1)	36
Double Flash (DBF)	14 (clase 0) 13 (clase 1)	27
Conjunto total	64 (clase 0) 57 (clase 1)	121

Lo cuál, considerando la cantidad de features que genera cada prueba, se tendría lo siguiente:

Prueba	Features obtenidas por sujeto	Features totales por sujeto <i>Caso dato único</i>	Features totales por sujeto <i>Caso datos independientes</i>	Totalidad de datos a procesar <i>Caso dato único</i>	Totalidad de datos a procesar <i>Caso datos independientes</i>
Resting State (RS)	2 CC 0 ERPs	1	2	64	128
Stemberg (STB)	4 CC 4 ERPs	2	8	74	296
Double Flash (DBF)	6 CC 6 ERPs	2	12	54	324
Conjunto total	12 CC 10 ERPs	5	22	193	456

Es necesario recordar que se obtiene un ERP y conectividad cerebral (CC) por evento. Además, cabe destacar tres cosas importantes al respecto de la tabla anterior; en primer lugar, las *features totales por sujetos* están siendo consideradas como la suma total de características que aporta cada prueba para un solo individuo, es decir, la suma entre la cantidad de ERPs más conectividades cerebrales que se obtienen.

Por otra parte, la *totalidad de datos a procesar* viene dada por  $features\_totales * n\_sujetos$ .

Por último, ambos datos anteriores entregan dos resultados diferentes debido a hay dos

formas de considerar el ingreso de los datos al clasificador dada la posibilidad de poder asociar clase-sujeto (por tanto era necesario concatenar toda data del sujeto) o simplemente abstraer el sujeto y considerar ERP-clase y/o CC-clase .

Además, se optó por considerar los datos de ERP de forma matricial dada la información que estos contenían y que cada dato tenía una dependencia con los otros. Mientras que la conectividad cerebral, dada su independencia de cada valor de correlación uno de otros, se optó por traspasar el tensor generado a un vector unidimensional de forma de facilitar el manejo de la entrada a la red del clasificador.

A continuación se detallarán las dos maneras en que se consideraron los datos de entrada. Ambos casos presentan pros y contras que deben ser considerados, provocando a priori que ningún método sea mejor que el otro o dé mejores resultados que el otro, por esta razón se deberán trabajar con ambos métodos de forma separada.

#### **4.1.1. Concatenación de datos: conformación de un dato único**

Una forma es considerar los ERPs o conectividades como un solo gran dato, es decir, concatenar los ERPs y las conectividades de cada sujeto (un dato para los ERP y un dato para las conectividades), lo cual permitirá la clasificación por sujeto como una única entidad.

El problema con esto, es que los ERPs y por sobretodo la conectividad cerebral ya presentan una dimensionalidad importante a considerar como datos únicos y por tanto, al hacer la concatenación esto generaría un dato muy grande provocando que la primera capa de la red deba contener un número bastante elevado de neuronas (para algunos casos se estaría hablando de más de un millón de neuronas en la capa de entrada).

Además, la concatenación de los ERP pueden ser concatenados tanto por el eje X, provocando un aumento en la cantidad de columnas de la matriz; como por el eje Y, lo que genera un aumento en la cantidad de filas. Esto se debe a que los ERP son matrices. A priori, ninguna de las dos formas de concatenación puede ser considerada mejor que la otra, por tanto se evaluaron ambas.

En caso de utilizarse la concatenación por X (aumento de columnas), hay un problema no menor cuando se quiere trabajar con los datos de dos o más pruebas, que es que cada prueba tiene un número diferente de ERPs y conectividades y por tanto, la concatenación tiene una dimensionalidad diferente por prueba; esto conlleva a buscar una forma que permita trabajar con un número variable de neuronas. La respuesta a este problema fue utilizar la técnica de zero-padding en los datos de ERP concatenados.

Cabe destacar que este problema no se genera al concatenar por Y debido a que las redes neuronales recurrentes (RNN) pueden trabajar bien frente a un número variables de filas.

## Uso de zero-padding

Antes de introducir el uso de zero-padding en los datos, es necesario recordar que las dos pruebas de las cuales se pueden extraer ERPs (STB y DBF) se obtienen distintas cantidades de estos (4 para STB y 6 para DBF) y que cada ERP, independiente la prueba, es representado por una matriz de dimensión  $64 \times 333$ . Por dicha razón, en caso de concatenar los ERPs por prueba se obtiene una matriz de dimensión  $64 \times 1332$  para la prueba STB y una de dimensión  $64 \times 1998$  para la prueba DBF, es decir, dos matrices con igual cantidad de filas pero distinta cantidad de columnas.

Para el caso en que se utilizan ambas pruebas para realizar el proceso de clasificación, es evidente observar que la dimensionalidad de los datos de ingreso no son compatibles y por tanto era necesario realizar un arreglo de estos para evitar el ingreso de datos con tamaño variable, pues esto implicaría que la capa de entrada tuviese que trabajar con una cantidad de neuronas variables. Para resolver este problema, y así utilizar un número fijo de neuronas en la capa de entrada, se optó por emplear el método de padding, en específico, zero-padding [152, 102].

La técnica de zero-padding consiste en añadir ceros al final del tiempo de dominio de la señal, con el objetivo de incrementar la longitud de esta [95].

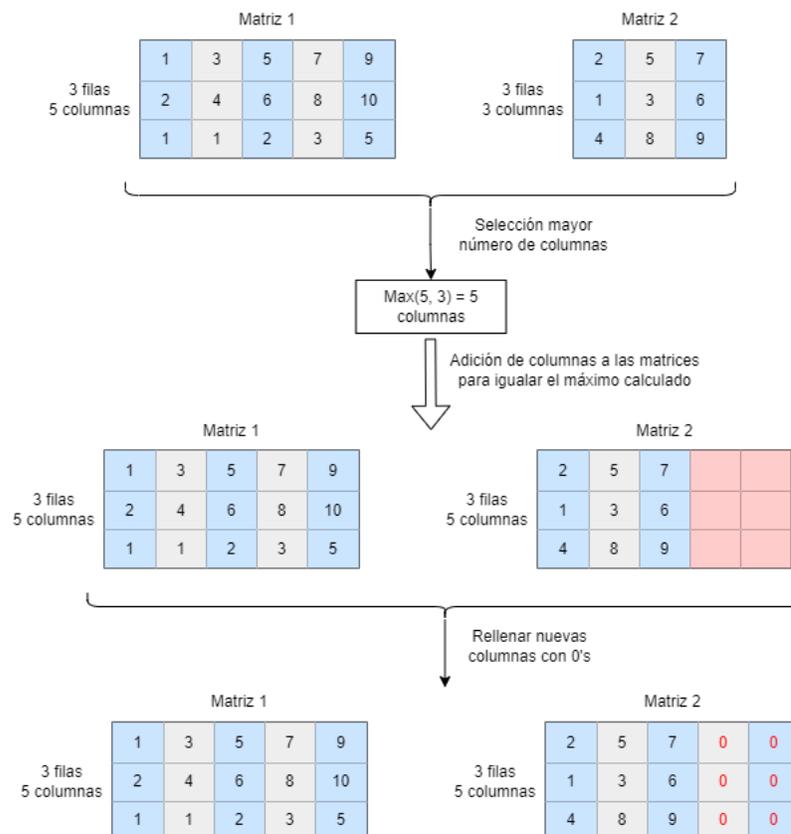


Figura 4.2: Ejemplo de uso de Zero Padding.

Para poder implementar esta técnica en los datos concatenados, primero se determina

cuál es la matriz que presenta una mayor cantidad de columnas. Luego, a cada uno de los datos se les agregan tantas columnas como se necesiten para equiparar el total de columnas de la matriz más grande. Cada una de estas nuevas columnas se completan con ceros (ver Figura 4.2).

### 4.1.2. ERP como dato independiente

La otra forma es considerar cada ERP y conectividad cerebral como datos independientes debido a que estos, si bien son del mismo sujeto, son obtenidos bajo eventos diferentes y por tanto son datos diferentes.

Si bien el que aumente la cantidad de datos por clase no genera inconveniente dado que el set de datos seguiría estando equilibrado, sí se pierde el concepto de “sujeto” dado que pasaría cada tupla ERP-CC a conformar un sujeto distinto para el clasificador, lo cual podría generar problemas.

A su vez, esto trae consigo tres puntos relevantes a favor; primero, este método tiene más sentido a nivel conceptual dado que considera cada ERP como un dato totalmente diferente a los otros y no como una única señal; segundo, genera que la dimensionalidad de los datos de entrada a la red sea considerablemente menor que en el otro caso; y por último, no hay problemas con el largo variable del input entregado a la red al momento de utilizar dos o más pruebas en conjunto dado que siempre ingresaría 1 conectividad junto a 1 ERP, que tienen la misma dimensión independiente de la prueba.

## 4.2. Modelo

Es importante recordar que la tesis se enfoca en identificar si es factible utilizar un modelo de Deep Learning para la clasificación de esquizofrenia. Por esta razón, la solución propuesta solo utiliza modelos Deep Learning y no Machine Learning.

Para el diseño del modelo de clasificación se utilizó la combinación de una Red Neuronal Recurrente (RNN) y una Red FeedForward (FNN).

La red feedforward es la red neuronal más clásica y básica, y por ende corresponde a la estructura más conocida de lo que es una red neuronal. Una red neuronal pretende simular al cerebro humano y las conexiones entre sus neuronas, las cuales son llamadas nodos o perceptrón en una red neuronal artificial.

Para el caso de una red feedforward, estos nodos están organizados en varias capas y cada nodo de una capa está conectado a todos los nodos de la capa siguiente provocando el flujo de información en una sola dirección, sin formar ciclos (ver Figura 4.3). Cabe destacar que las conexiones conformadas no son todas iguales puesto que estas cuentan con un valor propio asociado, llamado **peso**, los cuales se van ajustando durante la fase de entrenamiento para mejorar la predicción generada por la red [146, 57].

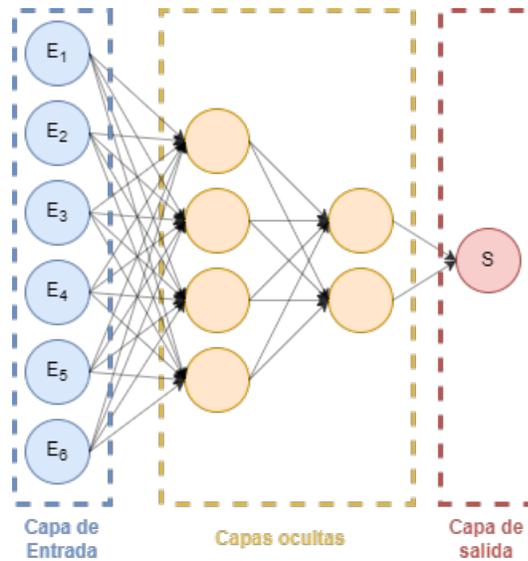


Figura 4.3: Red FeedForward

Por otra parte, la red recurrente es una red neuronal similar a la feedforward pero cuya diferencia radica en la existencia de bucles, permitiendo que la información viaje desde el nodo inicial al final pero también haciendo un ciclo de retroalimentación en cada nodo de capas intermedias, provocando que la información de salida de una neurona intermedia sea utilizada como información de entrada de la misma (ver Figura 4.4).

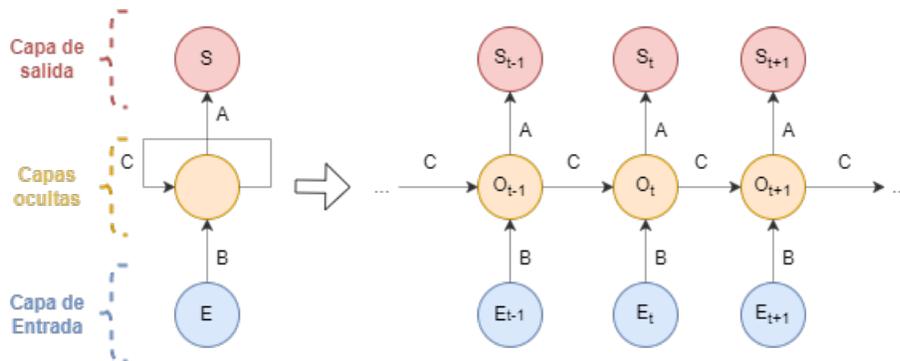


Figura 4.4: Red Neuronal Recurrent

Este tipo de red, a diferencia de la red feedforward, se utiliza para el procesamiento de datos de tipo temporal o secuencial; es decir que supone que los datos de entrada presentan alguna relación entre ellos. La red feedforward no considera una relación temporal o secuencial entre los datos es decir, los datos de entradas son tratados como datos totalmente independientes [100, 114].

En base a lo anterior, se optó por utilizar una red neuronal recurrente para procesar y clasificar los datos relacionados con los ERPs, dado que estos corresponden a secuencias temporales haciendo que este tipo de red sea el modelo más idóneo para realizar dicha labor.

Además, las redes recurrentes han mostrado tener buenos resultados para determinar esquizofrenia en base a ERPs y presentan menor complejidad en comparación con otras

utilizadas, como por ejemplo redes convolucionales, combinación entre convolucional y RNN, entre otras [12, 172, 40].

A su vez, se eligió la red feedforward para procesar los datos de la conectividad cerebral puesto que estos no tienen relación temporal ni secuencial entre ellos, sino más bien cada conectividad calculada es independiente una de otra.

Si bien la conectividad cerebral presenta una variable espacial, lo que implica que se pueda obtener una imagen visual de esta y por tanto puede ser procesada en base a una red convolucional (una de las redes más clásicas para procesamiento de imágenes), se optó de igual forma utilizar una red feedforward dado que esta red también había sido usada con anterioridad para el procesamiento de la conectividad cerebral [92, 91, 25], y debido a que las veces que se han usado redes convolucionales ha sido cuando se ha obtenido una imagen visual de la conectividad cerebral de una persona mediante una resonancia magnética (MRI o Magnetic resonance imaging), lo cual no se cumplía en esta ocasión.

Es importante realizar una aclaración respecto al uso de la red recurrente para la clasificación en base a los ERPs. Tal como se mencionó anteriormente, se consideraron dos formas diferentes para el ingreso de múltiples ERPs como datos de entrada a la red.

En la primera forma, el análisis para determinar si un paciente tiene o no esquizofrenia es por ERP (solo uno), que corresponde 100 % a una secuencia temporal provocando que el uso de una red recurrente calce coherentemente.

Por otra parte, en la segunda forma el análisis se hace sobre la unión de los ERPs como un solo dato, por lo que el orden secuencial y temporal se puede ver afectado dado que la relación temporal se quiebra en el punto de unión de los ERPs (pese a que siempre se unan en el mismo orden); pese a esto, de igual forma se utilizó una red recurrente para este caso dado que no hace sentido el uso de una red feedforward, pues se perdería por completo la “forma de señal” del ERPs por perder la secuencialidad de los datos de un ERP, y además se espera que la red recurrente aprenda que existen estos cortes en los datos por la unión de ERPs.

Lo anterior se suma a los contra ya señalados al querer utilizar la unión de ERPs en un solo dato para realizar el clasificador.

Debido al uso de dos clasificadores diferentes para procesar los dos tipos de datos de entradas del modelo, se diseñó un clasificador a partir de la unión de una red recurrente y una red feedforward.

Esta nueva red opera de la siguiente manera, por una parte se tiene la red recurrente cuyo tamaño de la primera capa es variable según el método que se utilice para ingresar los ERPs a la red y la prueba a utilizar (STB o BDF; recordar que RS no cuenta con ERP).

En caso de usarse los ERPs como datos independientes unos de otros, entonces el tamaño de la primera capa es de 333 nodos para las pruebas STB y DBF. Mientras que si se realiza la concatenación de ERPs, no existe un tamaño de primera capa preciso pues todo dependerá de la cantidad de ERPs concatenados (cada ERP es de tamaño  $64 \times 333$ ) y la forma en que estos sean concatenados (si por tiempo, representado por la variable  $x$ , o por frecuencia,

representado por la variable  $y$ ).

Esta red recurrente, correspondiente a una SimpleRNN, además cuenta con un tamaño de output variable, dado que este variará para probar distintas configuraciones, el cual se conecta con una capa densa cuyo output final sería de tamaño variable. El hecho de utilizar una capa densa al final de la red recurrente es debido a que esto genera mejores resultados para clasificaciones binarias [23]. Además, la cantidad de estados ocultos (hidden state) es de 64, que corresponde a la cantidad de canales existentes; aunque este variaría en caso de utilizar la data de ERPs concatenadas aumentando la cantidad de filas.

Por otra parte, se tiene que la red feedforward cuenta con una capa inicial, una intermedia y una capa final como configuración principal, la cual puede ir alterándose en busca de mejores resultados. La capa inicial cuenta con un tamaño de 688,128 nodos, para las pruebas DBF y STB puesto que el tamaño del tensor que representa a la conectividad cerebral es de  $64 \times 64 \times 168$ ; y un tamaño de 1,052,672 para la prueba RS dado que el tensor es de tamaño  $64 \times 64 \times 257$ . El tamaño de la capa intermedia y del output (capa final) sería variable.

Finalmente, los valores de salida de ambas capas finales de cada uno de los modelos entran como valores de entrada a una capa común de un tamaño correspondiente a la cantidad de salidas totales que hay entre la red recurrente y la feedforward (la suma de ambas), que corresponde a la capa que permite la unión entre ambos modelos, para luego pasar por una capa intermedia de tamaño variable y finalmente una capa final de tamaño 1 que corresponderá al resultado que entrega el clasificador. El tamaño de esta capa final debe ser de 1 pues el resultado del clasificador es binario, tiene esquizofrenia (1) o no tiene esquizofrenia (0).

Cabe destacar que todas las capas intermedias de la red son capas densas (Dense Layer) o también llamadas capas fully-connected puesto que cada neurona recibe información de todas las neuronas de la capa anterior; además estas no presentan una cantidad fija de neuronas debido a que no hay forma de determinar uno, por tanto hay que probar distintas configuraciones.

Las capas finales de las redes recurrentes y feedforward tienen una función de activación *softmax*; la capa final del modelo concatenado tiene una función de activación *sigmoid*; y el resto de las capas del modelo utiliza una función de activación *relu*.

La función de activación sigmoid toma cualquier valor real y entrega siempre un valor de salida entre 0 y 1. Se usa cuando un clasificador se enfrenta a clases binarias y por tanto solo existe un único nodo de salida, razón por la cual se utilizó en la capa final de cada uno de los modelos presentados (modelo concatenado, red recurrente para solo ERP, red feedforward para solo conectividad), puesto que estos debían arrojar 0 o 1 dependiendo si el dato era de un paciente con esquizofrenia o no [24].

Para hacer uso de esta función en la fase de entrenamiento y prueba, se debe establecer un umbral (threshold); este umbral determina que si un dato da un valor de salida mayor o igual al umbral, entonces el dato pertenece a la clase 1, sino pertenece a la clase 0. Para el caso de los modelos, el umbral fue de 0,5.

La función sigmoid se representa de la siguiente manera [171]:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Por otra parte, la función de activación softmax convierte un vector de números en un vector de probabilidades de tal forma que la suma de estas probabilidades siempre sea 1. Esta función se usa cuando se trata de una clasificación multiclase o bien cuando se tiene más de un nodo en la capa de salida, por eso se utilizó para la capa final de la red recurrente y feedforward utilizadas en el modelo concatenado, puesto que ambos modelos arrojaban más de un resultado los cuales luego eran ingresados a la capa densa que permitía la unión de ambos modelos anteriores [24].

La función softmax se representa de la siguiente manera [32]:

$$f(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$$

Y por último, una función de activación Relu, que es una función lineal por partes, que genera 0 en caso que la entrada sea negativa o el mismo valor de entrada en caso que esta sea positiva. Esta suele ser una de las funciones de activaciones más comunes para utilizar en capas densas intermedias dado que con esta función los modelos suelen aprender más rápido y generar mejores resultados; por dicha razón fue utilizada en las capas ocultas de las redes feedforward y del modelo concatenado.

La función Relu se representa de la siguiente manera [30]:

$$f(x) = \max(x, 0) = \begin{cases} 0, & \text{if } x \leq 0 \\ x, & \text{if } x > 0 \end{cases}$$

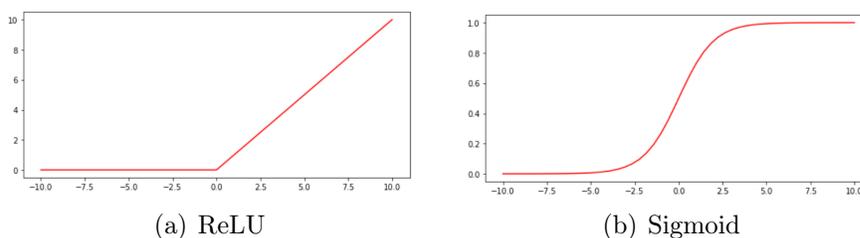


Figura 4.5: Funciones de activación

En la figura 4.5 se pueden apreciar los gráficos de los valores entre los cuales se mueven los resultados entregados por cada una de las funciones de activación; cabe destacar, que softmax no es una función como tal dado que no recibe como entrada un único elemento sino que un vector, razón por la cual no se representa gráficamente. En la figura 4.6 se puede apreciar un

diagrama del diseño completo del modelo utilizado para la clasificación de esquizofrenia en esta tesis.

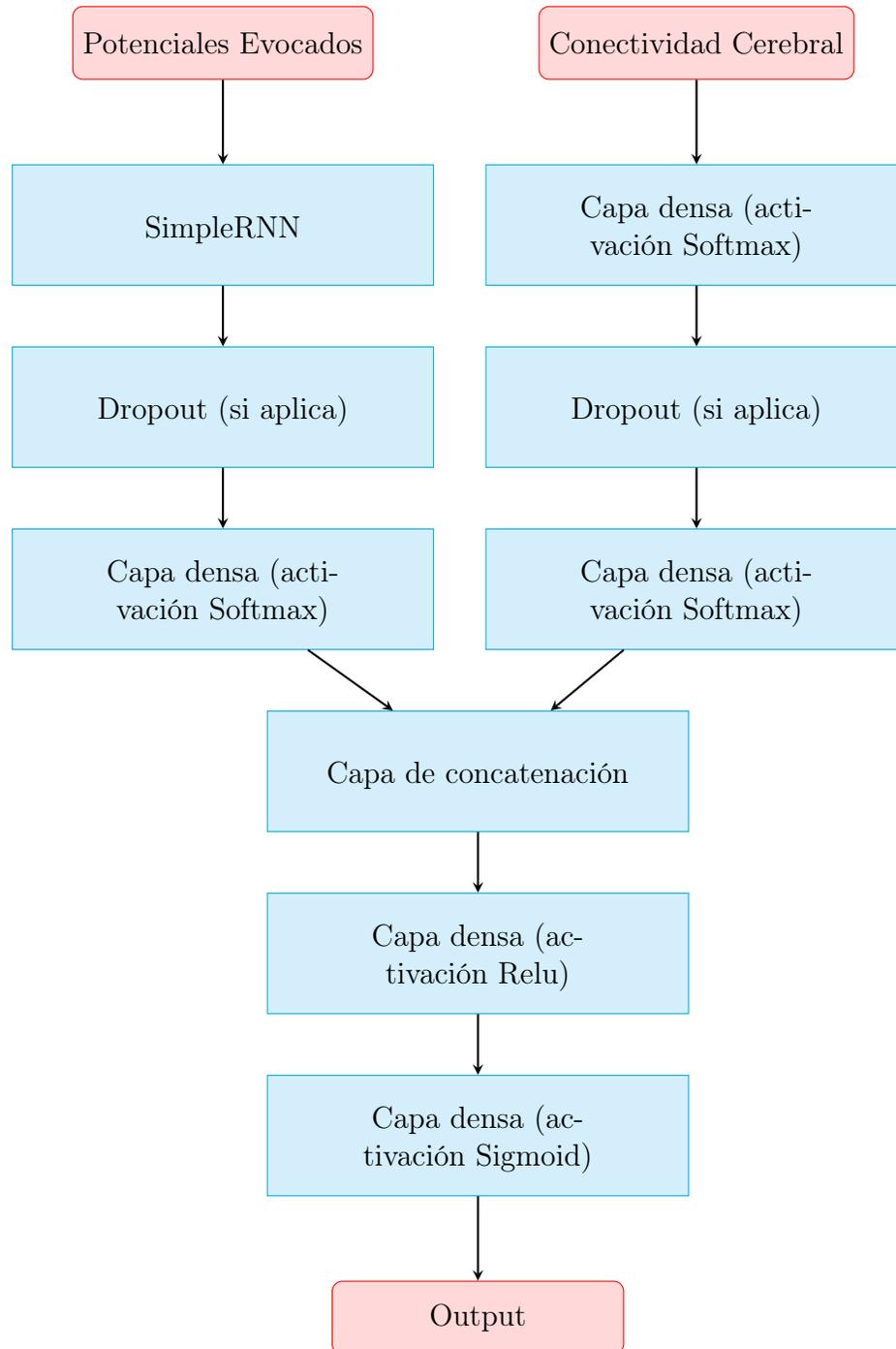


Figura 4.6: Diseño del Modelo concatenado elaborado.

Debido al objetivo de la tesis, el diseño del clasificador sufre ciertas variaciones leves para poder analizar los diferentes tipos de entradas (ERP y conectividad cerebral) de forma tanto independiente como en conjunto, para así poder determinar la influencia que tiene cada una de estas en la determinación de esquizofrenia y, así, determinar los biomarcadores que se están buscando.

En caso que se utilicen únicamente los datos de los potenciales evocados obtenidos en cada prueba, el modelo de clasificación utilizado es únicamente la red neuronal recurrente de tres capas (entrada, salida e intermedia), cuya última capa se compone de un nodo y una función de activación *sigmoid* (el resto de las capas mantienen la misma composición mencionada previamente en el modelo concatenado) tal como se aprecia en la Figura 4.7.

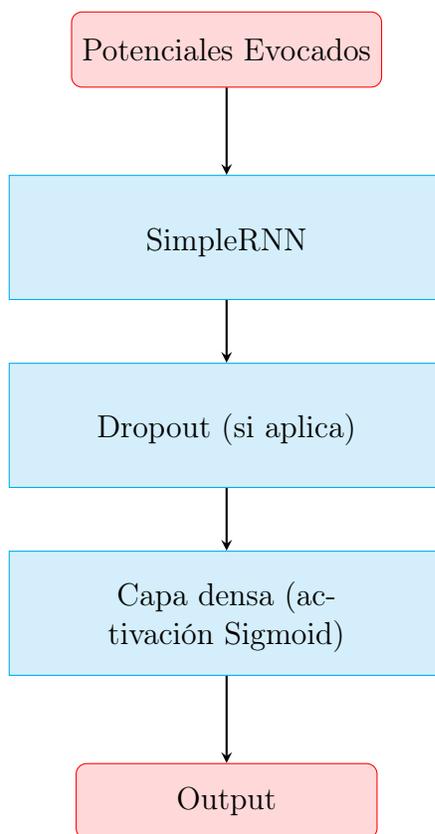


Figura 4.7: Diseño de la Red Neuronal Recurrente. Modelo usado para procesar datos ERP.

Por otra parte, en caso que se utilicen solamente los datos de conectividad cerebral se usa, como modelo de clasificación, la red neuronal feedforward cuya configuración es idéntica a la mencionada para el modelo concatenado, modificando únicamente su última capa para que esta solo tenga un nodo de salida y utilice una función de activación *sigmoid*, tal como se aprecia en la Figura 4.8.

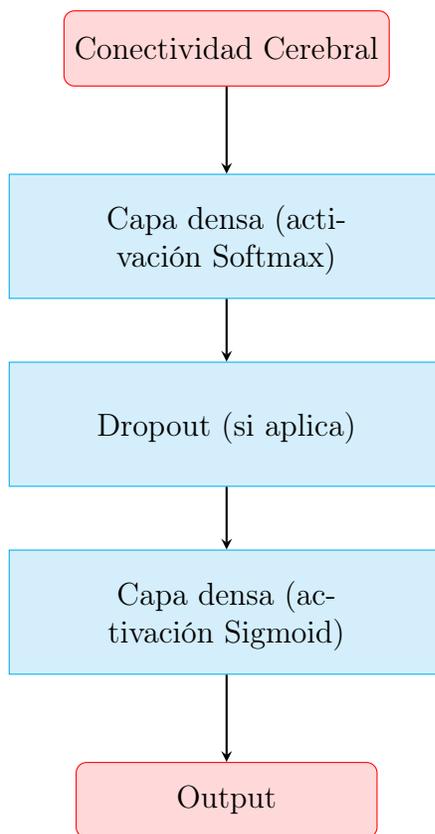


Figura 4.8: Diseño de la Red Feedforward. Modelo usado para procesar datos de Conectividad cerebral.

Finalmente, en caso de utilizar tanto la conectividad cerebral como el potencial de evocado como datos de entradas para el clasificador, se usa el modelo concatenado detallado previamente.

### 4.3. Selección de características

La selección de características, junto con la limpieza de los datos, es una de las partes más importantes al momento de realizar un buen clasificador puesto que permite eliminar información irrelevante la cual puede generar un impacto negativo en el rendimiento del clasificador [178]. Además, esta también compone uno de los puntos más fundamentales de esta tesis puesto que va a permitir determinar finalmente el conjunto de biomarcadores buscados.

La selección de características es un proceso donde se selecciona, de forma manual o automática, un subconjunto del total de características o variables que componen el set de datos, eligiendo aquellas que se consideran no redundantes y que representan de mejor manera al set completo. Existen diversos métodos de selección que se dividen en tres grandes categorías (métodos de filtrado, embedded y wrapper) y cuyo uso suele depender del tipo de variables de ingreso y salida del clasificador (variables numéricas o categóricas) [56, 93].

Para esta tesis, el método automático de selección no era factible debido al tipo de datos utilizados en los clasificadores. Tal como se mencionó previamente un dato de entrada estaba representado por un vector (conectividad cerebral) o por una matriz (ERP), provocando que cada neurona de la capa de entrada se encargara de uno de los valores de estos vectores y no de un dato completo.

En caso de utilizar un método de *feature select*, estos no seleccionarían ERP o conectividades cerebrales diferentes, sino que elegirían valores de un vector (que representa a un único ERP o conectividad cerebral) como el conjunto de características a utilizar para la clasificación, provocando la pérdida o modificación del ERP y conectividad cerebral obtenido para un cierto individuo conllevando al uso de datos incoherentes y finalmente una clasificación y biomarcadores incorrectos.

Por dicha razón se optó que para realizar el proceso de selección de características, que permitiría determinar que conjunto de ERPs y/o conectividades cerebrales eran relevantes, se realizarán múltiples clasificadores que utilizarán como datos de entradas uno de los conjuntos que se conforman del total de combinaciones posibles que existen entre todos los ERPs y conectividades cerebrales.

Para ejemplificar lo anterior suponga que en total existen 2 ERP ( $ERP_1, ERP_2$ ) y 1 conectividad cerebral ( $CC_1$ ), por tanto todas las combinaciones posibles de datos son  $ERP_1$ ;  $ERP_2$ ;  $ERP_1, ERP_2$ ;  $ERP_1, CC_1$ ;  $ERP_2, CC_1$ ;  $ERP_1, ERP_2, CC_1$ ;  $CC_1$ . Luego, se utiliza un clasificador de los mencionados en la sección anterior (red neuronal recurrente, red neuronal feedforward o red concatenada) para evaluar cada una de las combinaciones obtenidas, lo que implica el uso total de 7 clasificadores diferentes tal como se puede apreciar en la Figura 4.9.

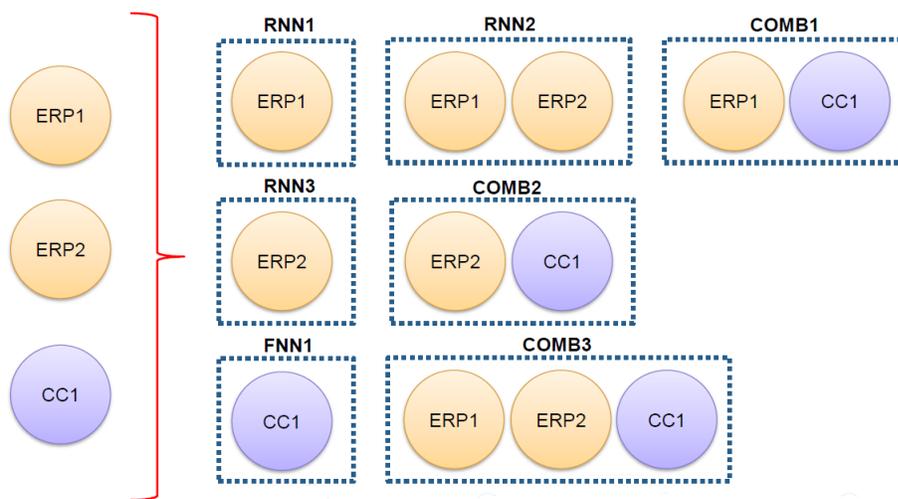


Figura 4.9: Ejemplo selección de características

Finalmente, los biomarcadores buscados corresponderán al conjunto o combinación de datos que fueron utilizados en el clasificador que obtuvo los mejores resultados.

Cabe destacar que en esta tesis hay un total de 4095 combinaciones de datos. Esto debido a que hay 15 combinaciones para la prueba STB, que tiene 4 ERPs y conectividades cerebrales; 63 combinaciones para la prueba DBF, que tiene 6 ERPs y conectividades cerebrales; 3 para

la prueba RS, que tiene únicamente 2 conectividades. Además de 945 combinaciones para el caso de la combinación de las pruebas STB-DBF; 45 combinaciones para el uso de las pruebas STB-RS en conjunto; 189 combinaciones para las pruebas DBF-RS en conjunto y 2835 combinaciones para el uso conjunto de las tres pruebas.

Por dicha razón, sumado además a la dimensionalidad de los datos (principalmente conectividad cerebral), se dificultó el entrenamiento de múltiples clasificadores puesto que el uso de memoria RAM era bastante grande, provocando que algunos necesitasen más de 25Gb de RAM para poder ejecutarse en su totalidad. Esta fue una de las limitaciones de esta tesis debido a que solo se contaba con un computador de 25Gb de RAM para realizar las pruebas y entrenar los clasificadores.

## 4.4. Entrenamiento y prueba

La primera fase de los algoritmos de clasificación es la fase de entrenamiento, que es donde el clasificador aprende a catalogar los datos en cada una de las categorías establecidas [207]. En esta fase se utiliza una porción del set de datos para realizar una comparación entre el valor real y el valor predicho por el clasificador, la cual es usada por el algoritmo para ajustar los pesos o factores internos del modelo con la finalidad de mejorar la clasificación [62]. Estos valores predichos por el clasificador serán los utilizados para determinar la clase a la cual pertenece el dato ingresado.

La segunda y última fase corresponde a la fase de prueba, que es donde se utiliza el clasificador ya entrenado frente a datos nunca antes vistos con el objetivo de evaluar el desempeño, la eficacia y el rendimiento del modelo [33]. De esta forma se determina cómo funcionaría el clasificador en la vida real [83, 44].

Para poder realizar ambas fases, usualmente se sigue un enfoque simple de dividir los datos iniciales en 2 partes o conjuntos, el de entrenamiento y el de prueba, eligiendo entre el 70 % y el 80 % de los datos para conformar el set de entrenamiento [162, 3]. A su vez, la conjunto de entrenamiento suele subdividirse en dos grupos más, conformando el set de validación y el set de entrenamiento real; esta subdivisión suele estar en proporción 80:20, donde el 80 % es utilizado como datos de entrenamientos y el 20 % como datos de validación. Esta subdivisión se realiza para evitar el sobreajuste del modelo y determinar sus hiperparámetros (por ejemplo la cantidad de capas ocultas), labor para la cual se usan los datos del conjunto de validación [83, 44, 177].

Lamentablemente esta técnica de división de la data no funciona bien para los casos en los que no se cuenta con grandes conjuntos de datos y por tanto no puede ser aplicada directamente. Para solucionar este problema se aplica una variación de la técnica de validación cruzada, cuyo método más conocido se llama *k-fold cross-validation*, puesto que se aplicará sobre el set de datos completo además de presentar ciertas variaciones para la conformación de los grupos, pero utilizando el mismo algoritmo e idea del cross-validation original.

### 4.4.1. Cross-Validation

Cross-validation o validación cruzada, es un técnica que se utiliza cuando se tiene un set de datos pequeños con el fin de asegurar que el modelo pueda funcionar bien en datos no vistos. Además, es útil para contrarrestar la naturaleza estocástica tanto al construir los conjuntos de datos, como de las variables y parámetros propios de los modelos de clasificación. Dicha característica provoca que los resultados obtenidos para un clasificador no necesariamente serán siempre los mismos y, en consecuencia, las predicciones y errores también serán distintos [29].

Para aplicar el método de validación cruzada, se debe dividir el set de datos en  $k$  conjuntos. Luego, se toma uno de esos conjuntos como el set de prueba y los  $k - 1$  conjuntos restantes se utilizan en el proceso de entrenamiento del clasificador. Este proceso se repite  $k$  veces, eligiendo cada vez un set de prueba diferente a los anteriores. Una vez terminado el proceso, se realiza un promedio de las métricas de evaluación obtenidas en cada iteración, obteniendo así las métricas finales para dicho clasificador (ver Figura 4.10) [116].

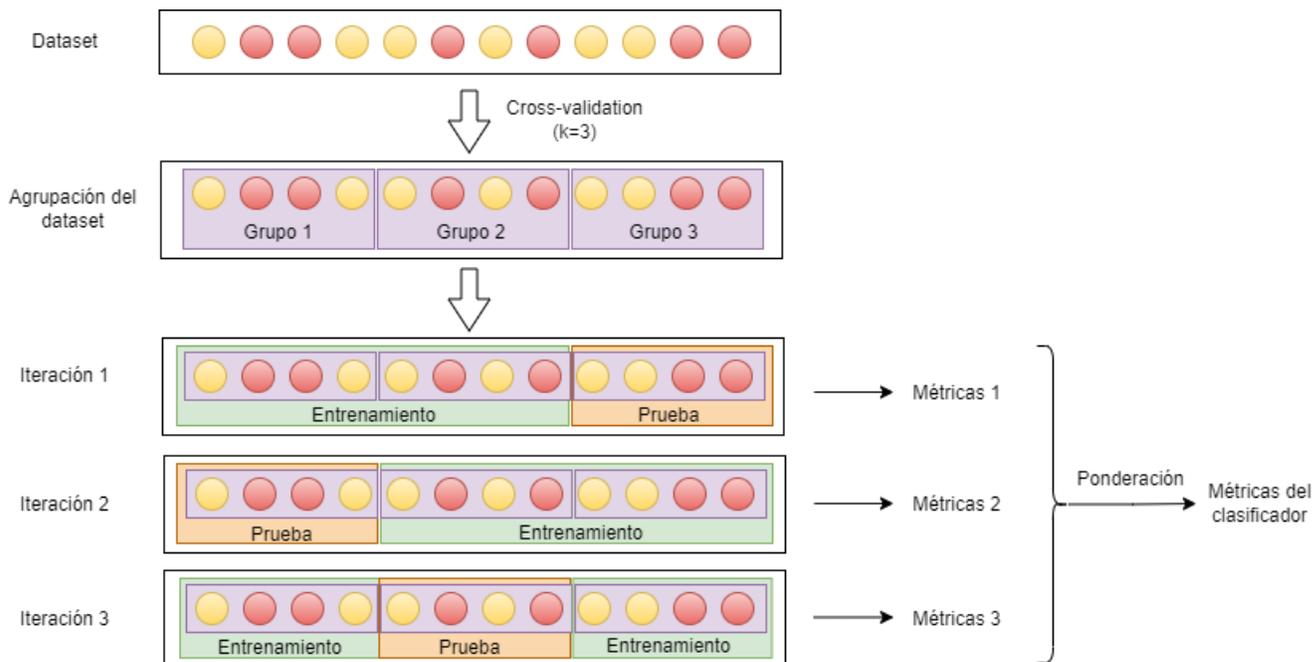


Figura 4.10: Proceso de Cross-validation con  $k = 3$

En cuanto a esta tesis, debido a la poca cantidad de datos con los que contó para realizar esta tesis, se tuvo que utilizar la técnica de *cross-validation* para la conformación de los set de datos utilizados durante la etapa de entrenamiento y testing del modelo, descartando así el uso clásico de subdivisión del set de datos en dos grupos diferentes y fijos desde un inicio.

Esta técnica suele utilizarse en el conjunto de entrenamiento para realizar la subdivisión entre conjunto de validación y entrenamiento, pero dado a la pequeña cantidad de datos con los cuales se contaba, se eliminó la elección de un set de validación y se aplicó la técnica en el conjunto completo de datos para determinar un set de entrenamiento y otro de prueba.

La razón por la cual se descartó el uso de un conjunto de validación fue debido al tamaño

del set de datos, el cual va de 27 a 128 datos, puesto que en caso de usarse y aplicar cross-validation sobre el 80% de los datos provocaría que el set de prueba esté conformado por una cantidad de datos fijos muy pequeños (de 5 a 25 datos) como para ser suficientes para determinar que las métricas de evaluación del clasificador sean válidas y generalizables. Por eso, se optó por aplicar cross-validation para conformar un set de entrenamiento y uno de prueba, omitiendo el set de validación.

El método de cross-validation utilizado en esta tesis, en esencia sigue el mismo patrón señalado previamente pero tenía dos problemas debido a la naturaleza de los datos, por lo cual no podía ser utilizado tal cual como se propone sino que tuvo que modificarse para poder ser compatible con los diferentes tipos de datos a usar.

El método utilizado consta en dividir los datos en  $k$  grupos con igual proporción de clases para luego elegir uno de estos como set de training y los otros  $k - 1$  como set de entrenamiento, entrenar el modelo con este último set y realizar la predicción sobre el set de prueba elegido. Luego repetir este mismo algoritmos hasta haber utilizado una vez cada uno de los  $k$  grupos de datos formados como set de prueba. Finalmente, se realiza una ponderación de las métricas obtenidas por cada uno de los  $k$  modelos evaluados para obtener las métricas finales generalizadas.

Lo anterior sería equivalente a entrenar y evaluar múltiples veces una misma configuración de red pero utilizando diferentes divisiones de los conjuntos de entrenamiento y prueba. Es importante notar que cada vez que se elige un nuevo set de prueba (al empezar cada iteración del cross-validation), se resetean los pesos del modelo, es decir, se utiliza un modelo no entrenado. Esto es para evitar que los datos de entrenamiento elegidos en una iteración previa genere problemas de aprendizaje del modelo en iteraciones posteriores y, por ende, resultados mejores que los reales.

Por otra parte, la conformación de los grupos utiliza un algoritmo especial debido a la naturaleza de los datos.

Para el caso en que el clasificador utilice solo una de las pruebas mencionadas (STB, DBF, RS) la conformación de los grupos se realiza de forma tal que los ERPs que pertenezcan a cada uno de los grupos contengan todos los ERPs posibles a los cuales se puede ver enfrentado el clasificador según la prueba.

Lo anterior se realiza para evitar que al hacer la subdivisión de los datos uno de los grupos pueda quedar conformado únicamente con un tipo de ERP (y no con todos los ERPs factibles que se le deben pasar al clasificador), pues esto generaría problemas ya sea durante el proceso de entrenamiento o de prueba del clasificador generado.

Para llevar a cabo dicho arreglo lo que se hizo fue aplicar cross-validation sobre un conjunto de datos ficticios que representaban a los sujetos (en teoría esto provoca que divida a los sujetos en  $k$  grupos diferentes); una vez obtenidos estos grupos, lo que se hace es hacer una especie de mapeo sujeto-ERPs para poder conformar los grupos finales, es decir, por cada grupo se toman los sujetos que lo conforman y luego se forma un grupo nuevo en el cual están los ERPs asociados a los sujetos de dicho grupo (ver Figura 4.11).

Con esto se evita que el problema que uno de los grupos pueda ser conformado con solo

una clase de ERPs, provocando que cada grupo tenga todos los tipos de ERPs posibles según la prueba utilizada; y además, esto también evita que un sujeto estuviese presente en dos grupos distintos del cross-validation, asegurando de esta forma que los datos de pruebas son datos y sujetos nunca antes vistos.

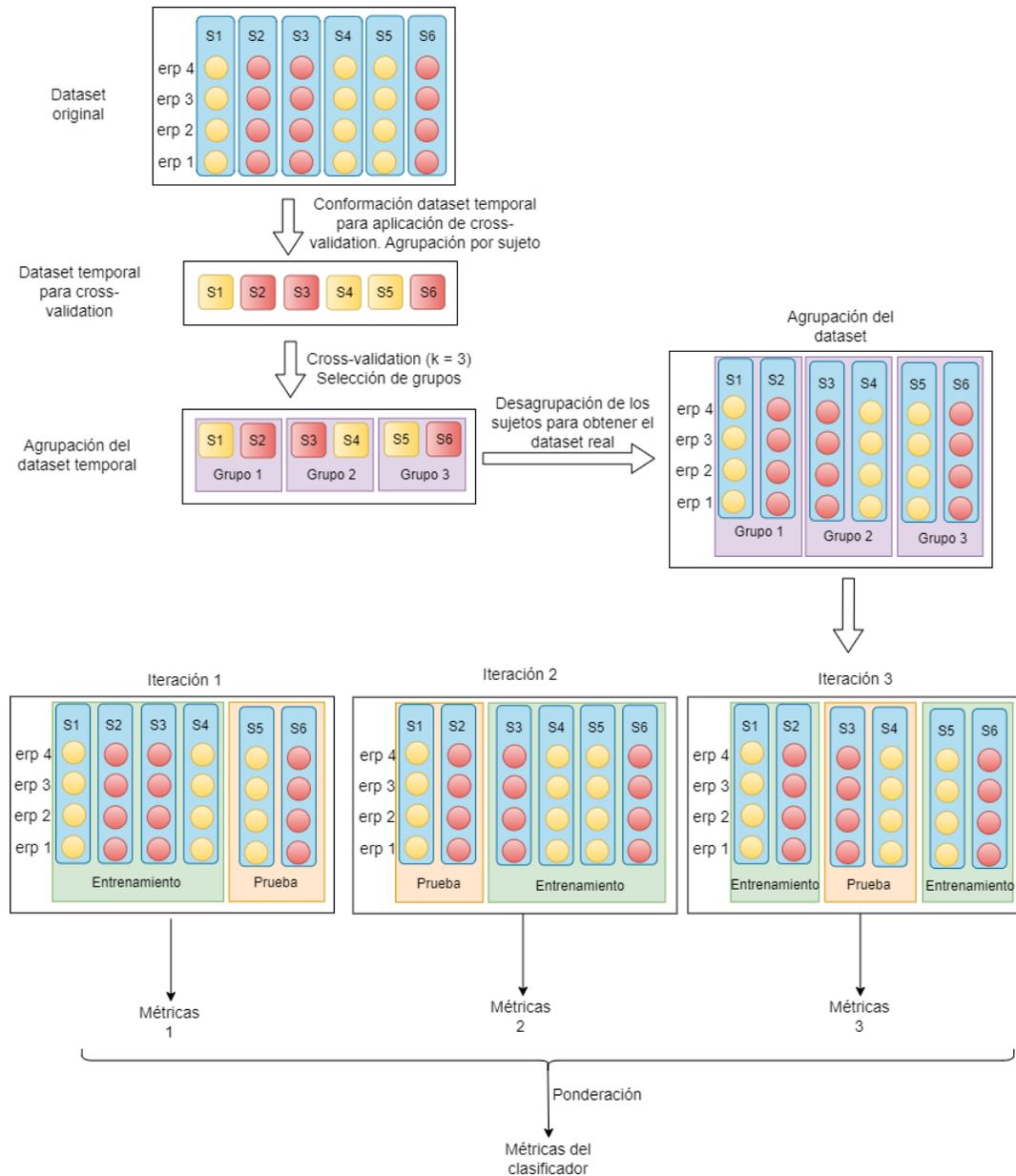


Figura 4.11: Proceso de Cross-validation con  $k = 3$ . Donde  $erp\ i$  corresponde al  $i$ -ésimo ERP del sujeto y  $S_i$  corresponde al  $i$ -ésimo sujeto del cual se extrajeron los ERP.

Por otra parte, en caso que el clasificador a conformar considere más de una prueba también existía un problema el cual tuvo que abordarse. En este caso, podía ocurrir que los grupos conformados no consideraran datos de todas las pruebas a las cuales se sometería el clasificador sino que únicamente a una de estas. Por esta razón, además de aplicar el método anterior para la conformación de los grupos para considerar todos los ERPs, se tuvo que desarrollar un método similar para asegurar que los grupos tuviesen datos de cada una de

las pruebas a utilizar.

Para llevar a cabo lo anterior lo que se hizo fue realizar el mismo proceso anterior para la conformación de grupos, pero haciendo la discriminación por prueba.

Por ejemplo, si se quiere evaluar STB con DBF entonces todos los datos pertenecientes a la prueba STB pasan por la conformación de grupo obteniéndose el set para STB y, a su vez, todos los datos de la prueba DBF pasan por el mismo proceso pero de forma separada, obteniéndose el set para DBF; una vez terminado este proceso se van a obtener 2 set de datos (uno para STB y otro para DBF) con  $k$  grupos cada uno.

Luego de obtener los distintos set de datos subdivididos en  $k$  grupos, lo que se hace es unir todos los  $i$ -ésimos grupos de los set de datos en uno solo, generando finalmente  $k$  grupos en total los cuales serán utilizados para realizar el resto del proceso del método cross-validation (ver Figura 4.12). De esta forma, los grupos finales contarán con data de cada una de las pruebas.

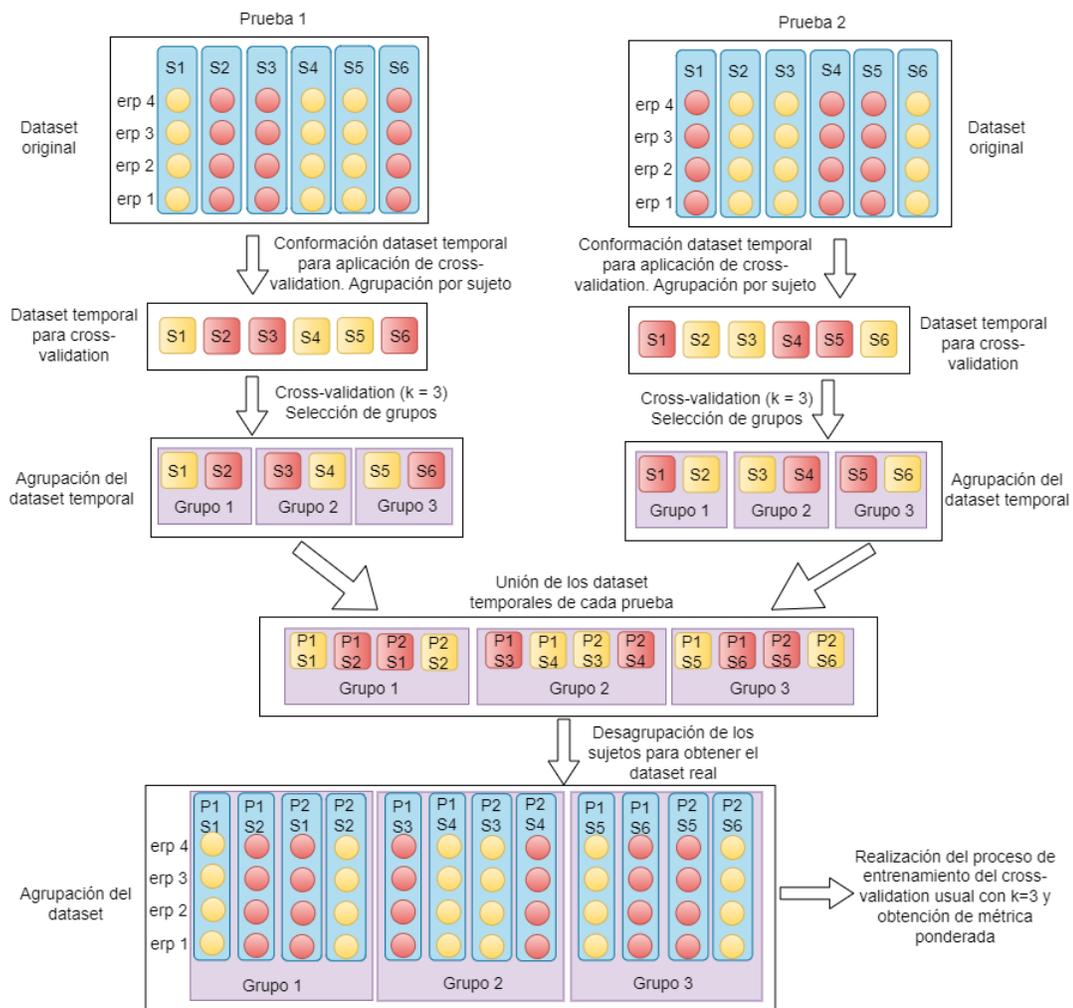


Figura 4.12: Proceso de Cross-validation con  $k = 3$  para el caso de uso de 2 pruebas. Donde  $erp\ i$  corresponde al  $i$ -ésimo ERP del sujeto;  $S_i$  corresponde al  $i$ -ésimo sujeto del cual se extrajeron los ERP y  $P_j S_i$  corresponde al  $i$ -ésimo sujeto de la  $j$ -ésima prueba.

### 4.4.2. Evaluar el sobreajuste del modelo

Dado que en este caso se eliminó el uso de un set de validación debido a la cantidad de datos tan pequeña del dataset, se tuvo que buscar una forma alternativa para asegurar que el modelo elegido no presentara overfitting o sobreajuste.

El sobreajuste ocurre cuando un modelo estadístico se ajusta exactamente a sus datos de entrenamiento, es decir, se aprende muy bien dichos datos; provocando un muy buen desempeño en los datos de entrenamiento pero un mal desempeño en los datos de validación, lo cual conlleva a que el modelo no sea generalizable dado que funcionaría mal frente a datos nunca antes vistos [31, 67].

El overfitting o sobreajuste suele determinarse mediante la visualización gráfica de la pérdida (loss) por época (ver Figura 4.13) [80, 128].

Otra forma para evaluar la existencia de overfitting es mediante el *accuracy*, si el *accuracy* del entrenamiento es muy bueno pero el del set de validación es muy bajo, entonces existe sobreajuste de los datos [80, 128].

En este caso, dado que solo se contaba con un set de entrenamiento y uno de prueba se optó por hacer el mismo análisis explicado anteriormente pero utilizando únicamente los datos de prueba. Es decir, se realizará la gráfica del número de épocas vs la pérdida del modelo (loss) con el set de prueba en vez del de validación. Para esto se grafica la pérdida durante el entrenamiento y la pérdida de la predicción cada cierto número de épocas. En caso que exista overfitting, la pérdida de la predicción presentará un punto a partir del cual la pérdida comienza a subir (la gráfica se asemeja a la de una función cuadrática).

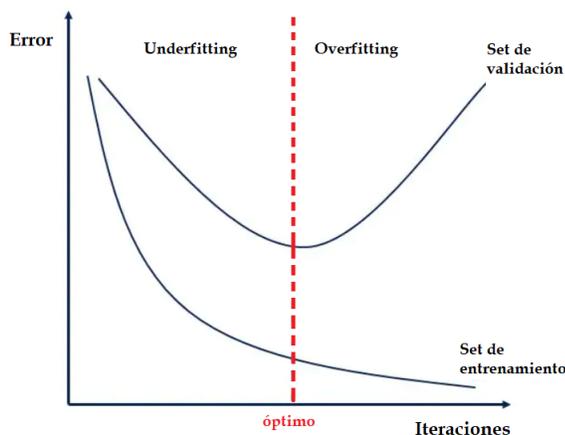


Figura 4.13: Overfitting y Underfitting

### 4.4.3. Métricas y métodos de evaluación

Para evaluar el desempeño de un clasificador se utilizan diferentes métricas y métodos de evaluación, cuyo uso dependerá completamente del objetivo y contexto bajo el cual fue elaborado el clasificador.

Uno de los métodos de evaluación a utilizar en esta tesis se llama matriz de confusión, la cual permite obtener información respecto a las predicciones y a partir de la cual se extraen las diferentes métricas para evaluar el desempeño del clasificador. Esta matriz tiene un tamaño de  $N \times N$ , con  $N$  la cantidad de clases que existen y además muestra en profundidad las predicciones correctas e incorrectas por clase que realizó el clasificador.

Para el caso de esta tesis, dado que existen 2 clases (esquizofrénico, no esquizofrénico), la matriz de confusión se vería de la siguiente manera:

1	TP	FN
0	FP	TN
	1	0

Donde  $TP$  (True Positive o verdaderos positivos) indica la cantidad de predicciones que el clasificador señaló como positivas (pertenecientes a la clase 1) que realmente son positivas;  $FP$  (False Positive o falsos positivos) indica la cantidad de predicciones señaladas como negativas (clase 0) que en realidad eran positivas;  $FN$  (False Negative o falsos negativos) indica la cantidad de predicciones señaladas como positivas (clase 1) que en realidad eran negativas y  $TN$  (True Negative o verdaderos negativos) indica la cantidad de predicciones clasificadas como negativas (clase 0) que realmente eran negativas.

Para analizar el rendimiento de los diferentes clasificadores entrenados en esta tesis, se utilizaron las siguientes métricas:

- **Accuracy:** Esta medida representa el número de predicciones correctas que realizó el clasificador, por ejemplo, si un modelo presenta un accuracy de 87% eso indica que de cada 100 datos el modelo clasifica correctamente 87 de ellos y se calcula de la siguiente manera [195, 8, 94]:

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

Generalmente, esta métrica representa que tan bueno es un modelo pero según el contexto esta puede no ser suficiente por si sola.

- **Precision:** Representa el porcentaje de casos positivos detectados y responde a la pregunta ¿Qué porcentaje de los clasificados como esquizofrénicos son realmente esquizofrénicos? Su fórmula es la siguiente [195, 8]:

$$Precision = \frac{TP}{TP + FP}$$

- **Recall:** Es la tasa de verdaderos positivos y corresponde a la proporción de casos positivos que fueron correctamente identificadas por el algoritmo. Su fórmula es la siguiente [195, 8]:

$$Recall = \frac{TP}{TP + FN}$$

En el área de la salud se dice que esta métrica es la capacidad de poder detectar correctamente la enfermedad entre los enfermos y por tanto responde a la pregunta ¿Qué porcentaje del total de esquizofrénicos el modelo pudo clasificar como esquizofrénicos?

- **F1:** Combina las métricas precision y recall en un único valor y se utiliza cuando las clases se encuentran desequilibradas. Se calcula [195, 8]:

$$F1 = \frac{2 * recall * precision}{recall + precision}$$

Para este trabajo, dado el contexto del mismo, las métricas a utilizar son: *accuracy*, por ser la métrica que generaliza el desempeño del modelo; y *Recall*, debido al contexto del problema puesto que corresponde a un clasificador para detectar una enfermedad (a mayor recall mejor es la probabilidad de detectar correctamente la enfermedad).

# Capítulo 5

## Resultados y discusión

Las pruebas realizadas para obtener un modelo a partir del cual se pudiesen extraer biomarcadores se pueden observar en la tabla 5.3

Clasificador	Características	Prueba/s
Red Recurrente (RNN)	ERP	STB
Red Recurrente (RNN)	ERP	DBF
Red Recurrente (RNN)	ERP	STB y DBF
Red Feedforward (FNN)	Conectividad	STB
Red Feedforward (FNN)	Conectividad	DBF
Red Feedforward (FNN)	Conectividad	RS
Red Feedforward (FNN)	Conectividad	STB y DBF
Red Feedforward (FNN)	Conectividad	STB y RS
Red Feedforward (FNN)	Conectividad	RS y DBF
Red Feedforward (FNN)	Conectividad	STB, RS y DBF
Concatenado	ERP y conectividad	STB
Concatenado	ERP y conectividad	DBF
Concatenado	ERP y conectividad	STB y DBF

Tabla 5.1: Pruebas realizadas para obtención de resultados

Estas pruebas se realizaron para el caso de la data concatenada y sin concatenar.

Además, cada prueba se realizó con diferentes configuraciones debido a la cantidad de parámetros que podían variar. La elección de los qué parámetros utilizar para las diferentes configuraciones fue completamente arbitrario. Los parámetros a variar fueron los siguientes:

- **Épocas:** Probándose 2, 6, 10. En un inicio se estimó realizar también 50 y 100 épocas pero al revisar el gráfico de entrenamiento para cada caso a probar, se verificó que utilizar dicha cantidad de épocas era redundante debido a que el punto óptimo de épocas de la curva variaba entre las 5 a 20 épocas.
- **Número de neuronas en las capas ocultas:** Se probaron 100, 50, 10 y también la mitad y el doble de neuronas que la primera capa. Esto también aplica para el output de la red neuronal (SimpleRNN).
- **Dropout:** Se probó un valor de 0 y 0,5.
- **Cantidad de grupos para cross-validation:** Se probaron 3, 4, 5 y 6 grupos. Con 3 grupos la proporcionalidad entre el grupo usado de testing con los de entrenamiento es aproximadamente 30 % y 70 %, respectivamente.
- **Número de capas ocultas de FNN:** Se probó 1 y 2 capas.

Con dichas configuraciones, se entrenaron 360 clasificadores diferentes utilizando solo datos ERPs (eran 120 clasificadores por prueba y en total eran 3 pruebas que utilizaban solo datos de potencial evocado); además de 1008 clasificadores para datos de conectividad cerebral (144 clasificadores por prueba, con un total de 7 pruebas que utilizaban únicamente datos de conectividad cerebral) y 6 clasificadores para el caso del modelo concatenado (solo se consideraron los mejores resultados por época para cada prueba para entrenar este clasificador). Con esto se obtuvieron más de 300.000 resultados diferentes.

Además, debido a la baja cantidad de datos con los que se contaba para realizar el entrenamiento de los clasificadores, cada configuración fue ejecutada 3 veces y luego promediadas sus métricas para así obtener métricas más representativas.

## 5.1. Resultados generales

Comenzando con los resultados se empezará señalando que solo un clasificador con datos de conectividad cerebral obtuvo buenos resultados ( $acc \geq 75\%$ ). Este fue para la prueba DBF y los resultados de dicho clasificador se detallan a continuación

Prueba	Configuración	CC	Accuracy	F1	Recall	Precision
DBF	6 épocas Sin dropout 100 neuronas en cada capa oculta 4 grupos de cross-validation 2 capas ocultas	1 - 2 - 11	75.4 %	64.29 %	66.67 %	73.71 %

Tabla 5.2: Resultados DBF para modelo Feedforward.

Además de ese clasificador, ningún otro logró obtener el resultado esperado para esta tesis. En promedio su el accuracy obtenido por los clasificadores que utilizan datos de conectividad cerebral fue de 55 % aproximadamente, con resultados que iban desde el 30 % al 77 %. En otras palabras no se obtuvieron los resultados esperados para los casos detallados en la tabla 5.3

Clasificador	Características	Prueba/s	Accuracy promedio
Feedforward (FNN)	Conectividad	STB	53 %
Feedforward (FNN)	Conectividad	RS	52 %
Feedforward (FNN)	Conectividad	STB y DBF	54 %
Feedforward (FNN)	Conectividad	STB y RS	60 %
Feedforward (FNN)	Conectividad	DBF y RS	55 %
Feedforward (FNN)	Conectividad	STB, DBF y RS	52 %
Concatenado	ERP y Conectividad	STB	60 %
Concatenado	ERP y Conectividad	DBF	58 %
Concatenado	ERP y Conectividad	STBy DBF	56 %

Tabla 5.3: Pruebas realizadas para obtención de resultados

A su vez, la prueba DBF tuvo en promedio un accuracy de 58 % aproximadamente.

Por otra parte, usando únicamente datos de ERP sí se obtuvieron varios resultados con un accuracy mayor o igual a 75 %. Con un resultado general para la prueba STB de 68 % accuracy aproximadamente; la prueba DBF tuvo un 65 % de accuracy aproximadamente y la prueba DBF-STB obtuvo un accuracy aproximado de 60 %.

Por tanto, en términos generales los clasificadores con uso único de potenciales evocados obtuvieron un rendimiento de 65 %, los con uso único de conectividad cerebral obtuvieron un 55 % y los que usaron potencial evocado y conectividad tuvieron un accuracy de 58 %.

A continuación se detallarán los resultados particulares con un accuracy mayor o igual al 75 %.

Prueba	Configuración	ERPs	Accuracy	F1	Recall	Precision
DBF	6 épocas Sin dropout 100 neuronas en capa oculta 3 grupos de cross-validation	1 - 3 - 10 - 11 - 12	75.31 %	67.76 %	63.29 %	85.87 %
DBF	10 épocas Sin dropout 10 neuronas en capa oculta 4 grupos de cross-validation	2 - 10 - 11 - 12	75 %	63.25 %	54.86 %	80.56 %

Tabla 5.4: Resultados DBF para modelo RNN.

Prueba	Configuración	ERPs	Accuracy	F1	Recall	Precision
STB	6 épocas Sin dropout 100 neuronas en capa oculta 6 grupos de cross-validation	91 - 101	75.45 %	69.20 %	62.04 %	87.93 %
STB	6 épocas 0.5 dropout mitad neuronas de primera capa en capa oculta 4 grupos de cross-validation	101	76.85 %	72.75 %	63.33 %	90 %
STB	6 épocas 0.5 dropout 50 neuronas en capa oculta 6 grupos de cross-validation	91 - 101	75 %	66.75 %	61.11 %	80.55 %
STB	6 épocas 0.5 dropout 100 neuronas en capa oculta 6 grupos de cross-validation	91	75.93 %	65.69 %	57.41 %	85.65 %
STB	6 épocas 0.5 dropout 100 neuronas en capa oculta 5 grupos de cross-validation	91	76.31 %	73.29 %	66.66 %	83.89 %

Tabla 5.5: Resultados STB para modelo RNN (parte 1).

Prueba	Configuración	ERPs	Accuracy	F1	Recall	Precision
STB	10 épocas Sin dropout 50 neuronas en capa oculta 4 grupos de cross-validation	101 - 102	75 %	70.27 %	61.67 %	85.42 %
STB	10 épocas 0.5 dropout 50 neuronas en capa oculta 3 grupos de cross-validation	92 - 101	75 %	70.42 %	62.96 %	85.26 %
STB	10 épocas Sin dropout 50 neuronas en capa oculta 5 grupos de cross-validation	91 - 101	75 %	68.49 %	60.56 %	89.55 %
STB	10 épocas Sin dropout 50 neuronas en capa oculta 4 grupos de cross-validation	91 - 102	79.63 %	75.43 %	67.5 %	91.67 %

Tabla 5.6: Resultados STB para modelo RNN (parte 2).

Prueba	Configuración	ERPs	Accuracy	F1	Recall	Precision
STB	6 épocas Sin dropout 50 neuronas en capa oculta 6 grupos de cross-validation <b>Data concatenada</b>	92 - 101 - 102	75 %	68.61 %	61.11 %	86.67 %
STB	10 épocas 0.5 dropout 100 neuronas en capa oculta 6 grupos de cross-validation <b>Data concatenada</b>	92 - 101	75 %	71.14 %	64.81 %	84.72 %
STB	10 épocas Sin dropout 50 neuronas en capa oculta 4 grupos de cross-validation	91 - 102	79.63 %	75.43 %	67.5 %	91.67 %
STB	10 épocas 0.5 dropout 100 neuronas en capa oculta 6 grupos de cross-validation <b>Data concatenada</b>	92 - 101 - 102	75 %	71.79 %	66.67 %	84.26 %

Tabla 5.7: Resultados STB para modelo RNN (data concatenada).

Ahora se detallarán los mejores clasificadores por cada prueba realizada. Estos fueron determinados en base al accuracy y el recall obtenido, eligiendo aquel que tuviese un accuracy sobre 75 % (con la finalidad de validar la hipótesis de la tesis) pero con el mayor recall.

Lo anterior implica que si por ejemplo hay un clasificador A con accuracy 80 % y con un recall de 77 % y otro clasificador B de accuracy 75 % pero con recall 79 %, entonces se elegirá como mejor clasificador el B dado que cumple el umbral del accuracy pero presenta mejor recall. La razón de esto es que un mayor recall implica que el clasificador pudo detectar mejor quienes eran realmente esquizofrénicos, dada la importancia de detectar la enfermedad en salud entonces se priorizó esta métrica.

Además, se verificó que los modelos elegidos cumplieren el criterio establecido para evitar overfitting y underfitting, con el objetivo de asegurar que el modelo seleccionado se enfrenta mejor que otros a datos nunca antes vistos. Cabe destacar que por dicha razón, varios resultados con accuracy mayor o igual a 75 % no fueron considerados al momento de detallar los resultados anteriormente señalados, dado que estos contaban con overfitting o underfitting.

Para el caso de la prueba de resting state (RS) y las pruebas conjuntas de STB-DBF, STB-RS, DBF-RS y STB-DBF-RS, no hubieron resultados satisfactorios.

Por otra parte, para la prueba DBF se puede observar que la mejor configuración para el clasificador Feedforward (solo conectividad) obtuvo un 75,4 % de accuracy junto a un 66,67 % de recall, utilizándose para esto únicamente 3 conectividades cerebrales que fueron la 1, 2 y 11.

A su vez, la mejor configuración para el clasificador RNN (solo ERP) obtuvo un 75 % de accuracy y un 63,29 %, utilizando los ERPs 1, 3, 10, 11 y 12. Para el caso del clasificador concatenado (ERP más conectividades) no se obtuvieron resultados satisfactorios.

En base a lo anterior, se determina que el clasificador Feedforward obtuvo los mejores resultados para el caso de la prueba DBF debido a que el recall de este es mejor.

Para el caso de la prueba STB se observa que tampoco hay resultados satisfactorios para el caso del clasificador concatenado ni para el clasificador Feedforward.

En cambio, para el caso del clasificador RNN se puede observar que la mejor configuración para el uso de los datos sin concatenar se obtiene a partir de los ERP 91 y 102, con los cuales se tiene un accuracy de 79,63 % y un recall de 67,5 %.

A su vez para el caso de los datos concatenados, el mejor clasificador se obtiene a partir de los ERP 92, 101 y 102 con un accuracy de 75 % y un recall de 66,67 %.

En base a lo anterior, se determina que el clasificador RNN utilizando los ERP sin concatenar obtiene mejores resultados.

Un resumen de los resultados mencionados anteriormente se muestran en la tabla [5.8](#)

Prueba	Modelo	Configuración	ERPs o CC	Accuracy	Recall
<b>DBF</b>	Feed Forward	6 épocas Sin dropout 100 neuronas en cada capa oculta 4 grupos de cross-validation 2 capas ocultas	1 - 2 - 11	75.4%	<b>66.67 %</b>
DBF	Red Recurrente	6 épocas Sin dropout 100 neuronas en capa oculta 3 grupos de cross-validation	1 - 3 - 10 - 11 - 12	75.31 %	63.29 %
<b>STB</b>	Red Recurrente	10 épocas Sin dropout 50 neuronas en capa oculta 4 grupos de cross-validation	91 - 102	79.63 %	<b>67.5 %</b>
STB	Red Recurrente	10 épocas 0.5 dropout 100 neuronas en capa oculta 6 grupos de cross-validation Data concatenada	92 - 101 - 102	75 %	66.67 %

Tabla 5.8: Mejores resultados

Todos los resultados obtenidos se pueden encontrar en el siguiente [link](#)

### 5.1.1. Significancia estadística

La significación estadística es un juicio sobre si se produce un resultado debido a la casualidad. Cuando un resultado es estadísticamente significativo, nos referimos a que es poco probable que el resultado se produzca debido a la casualidad o a la fluctuación aleatoria [99].

Para poder realizar el cálculo de la probabilidad, también llamado nivel de significación, se utilizó el método ANOVA.

ANOVA o análisis de varianza de un factor. El análisis de la varianza permite contrastar la hipótesis nula de que las medias de  $K$  poblaciones ( $K > 2$ ) son iguales, frente a la hipótesis alternativa de que por lo menos una de las poblaciones difiere de las demás en cuanto a su valor esperado [48].

ANOVA permite analizar la variación en una variable de respuesta medida en circunstancias definidas por factores discretos (variables de clasificación). Se usa un ANOVA en cuatro situaciones [163]:

- Cuando hay más de dos grupos que necesitan ser comparados.
- Cuando hay mediciones repetidas en más de dos ocasiones o cuando hay dos o más grupos en quienes se hacen mediciones repetidas en dos ocasiones.
- Cuando los sujetos pueden variar en una o más características que afectan el resultado y se necesita ajustar su efecto.
- Cuando se desea analizar simultáneamente el efecto de dos tratamientos diferentes, cuando el efecto de cada uno por separado y su posible interacción es importante.

Para el caso de esta tesis se deben comparar más de dos grupos de datos diferentes, los cuales están diferenciados por las pruebas realizadas. Lo cual también es equivalente a comparar el efecto de cada uno de estos clasificadores o tratamientos en el resultados de la clasificación.

Se realizó el análisis para cada uno de los modelos considerados los mejores de su grupo. De forma tal que dado un modelo  $N$  se tomaron todas las combinaciones posibles de características para ver si existía diferencia significativa entre estos. Esto se realiza para analizar la diferencia significativa a nivel de rendimiento (accuracy) y recall de los modelos.

Para poder determinar si se rechaza o acepta la hipótesis nula ( $H_0$ ), se determina un límite de significancia para realizar la comparación de probabilidad de 0.05 lo cual indica un riesgo de 5% de concluir que existe una diferencia cuando no hay una diferencia real [9].

Primero se realizó para el modelo STB con 10 épocas, 0 de dropout, 50 neuronas en capa oculta, 4 grupos de cross-validation. Para esto se tomaron las 15 combinaciones posibles de características bajo este modelo y se compararon entre sí bajo ANOVA.

Lo anterior generó como resultado que el nivel de significancia era 0.001 ( $p < 0.05$ ), por lo cual se rechaza la hipótesis nula y se concluye que hay diferencia significativa en el rendimiento de los modelos dependiendo de las características (ERP) de ingreso.

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	1046.6226	14	74.7587	3.4969	0.0019	2.0374
Dentro de los grupos	641.3404	40	21.3780			
Total	1687.9630	44				

Tabla 5.9: ANOVA STB no concatenada: Accuracy

También se obtuvo que el nivel de significancia para el caso de los recall del modelo es de 0.013 ( $p < 0.05$ ), por tanto se rechaza la hipótesis nula y se concluye que hay diferencia estadística en el recall de los modelos dependiendo de las características (ERP) de ingreso.

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	1266.4193	14	90.4585	2.6187	0.013	2.0374
Dentro de los grupos	1036.2842	30	34.5428			
Total	2302.7035	44				

Tabla 5.10: ANOVA STB no concatenada: Recall

Para el caso del modelo de STB con data concatenada, 10 épocas, 0.5 de dropout, 100 neuronas en capa oculta y 6 grupos de cross validation. Se obtuvieron los siguientes resultados

Evaluando la significancia estadística del rendimiento del modelo, se obtuvo que el nivel de significancia era de 0.30 ( $p > 0.05$ ) por tanto se concluye que no hay diferencia significativa en el rendimiento de los modelos dado que se cumple la hipótesis nula.

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	253.9557	10	25.3955	1.2631	0.308	2.2966
Dentro de los grupos	442.3170	22	20.1053			
Total	696.2728	32				

Tabla 5.11: ANOVA STB concatenada: Accuracy

Además, para el caso del recall se obtiene un nivel de significancia de 0.27 ( $p > 0.05$ ) concluyendo que no existe diferencia significativa en el recall de los modelos evaluados.

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	469.4531	10	46.9453	1.3210	0.2796	2.2966
Dentro de los grupos	781.8108	22	35.5368			
Total	1251.2639	32				

Tabla 5.12: ANOVA STB concatenada: Recall

Para el caso de los modelos DBF, se realizó el mismo proceso mencionado anteriormente para los modelos STB. En el caso del modelo DBF feed forward para 6 épocas, sin dropout, 2 capas ocultas, 100 neuronas en cada capa oculta y 4 grupos de cross validation se obtuvieron los siguientes resultados.

Se puede concluir que el modelo no cuenta con significancia estadística a nivel de rendimiento del clasificador con respecto al resto de los modelos con similares configuraciones. Esto dado que el nivel de significancia es mayor a 0.05

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	4006.6421	62	64.6232	0.6892	0.9482	1.4192
Dentro de los grupos	11814.2797	126	93.7641			
Total	15820.9218	188				

Tabla 5.13: ANOVA DBF feed forward: Accuracy

Por otra parte, también es posible concluir que el modelo tampoco cuenta con significancia estadística para el recall dado que su nivel de significancia es mayor a 0.05.

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	12474.1832	62	201.1965	1.0910	0.3362	1.4192
Dentro de los grupos	23235.6388	128	184.4098			
Total	35709.822	188				

Tabla 5.14: ANOVA DBF feed forward: Recall

Para el caso del modelo DBF con red recurrente para 6 épocas, sin dropout, 100 neuronas en capa oculta y 3 grupos de cross validation, se obtienen los siguientes resultados.

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	4735.6078	62	76.3807	1.1581	0.2427	1.4192
Dentro de los grupos	8310.0813	126	65.9530			
Total	13045.6892	188				

Tabla 5.15: ANOVA DBF red recurrente: Accuracy

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	13465.3676	62	217.1833	1.3584	0.0750	1.4192
Dentro de los grupos	20144.0562	126	159.8734			
Total	33609.4238	188				

Tabla 5.16: ANOVA DBF red recurrente: Recall

De esto se puede concluir que tanto para el caso de rendimiento como recall, el modelo no presenta significancia estadística con respecto a otros modelos de idéntica configuración pero diferentes características de entradas.

Además, se realizó una comparación entre los modelos con datos de las pruebas STB (red recurrente) y DBF (red recurrente), ambos con data no concatenada, con la finalidad de comprobar la significancia estadística entre los diferentes modelos con diferentes características de entradas.

Lo anterior permitirá comparar estadísticamente diferentes métodos de análisis para la clasificación de esquizofrenia. Esto entregó lo siguientes resultados

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	9875.097	93	106.1838	1.4518	0.0163	1.3331
Dentro de los grupos	13749.45	188	73.1353			
Total	23624.55	281				

Tabla 5.17: ANOVA Red Recurrente: Rendimiento

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	34920.91	93	375.4937	2.6239	$1,13 * e^{-8}$	1.3331
Dentro de los grupos	26902.96	188	143.1009			
Total	61823.87	281				

Tabla 5.18: ANOVA Red Recurrente: Recall

De lo anterior podemos ver que tanto para el rendimiento de los clasificadores como para el recall de los mismos se obtiene un nivel de significancia menor a 0.05. Por ende en ambos casos se rechaza la hipótesis nula y se concluye que hay diferencia significativa en el rendimiento y recall de los modelos.

## 5.2. Elección del mejor modelo

Finalmente, en base a los resultados ya señalados se puede obtener el mejor de todos los modelos evaluados.

Este modelo se eligió en base a las condiciones previamente señaladas; en este caso, como todos los clasificadores de los resultados mostrados anteriormente presentan sobre 75% de accuracy, se escogió aquel con el mayor recall.

En base a lo anterior se determina que el modelo con mayor recall que cumple con los criterios establecidos es el clasificador RNN para la prueba STB sin data concatenada, con un recall de 67,5%.

Prueba	Modelo	Configuración	ERPs o CC	Accuracy	Recall
STB	Red Recurrente	10 épocas Sin dropout 50 neuronas en capa oculta 4 grupos de cross-validation	91 - 102	79.63%	67.5%

Tabla 5.19: Mejor clasificador

Además, este modelo sí presentada una significancia estadística tanto a nivel de rendimiento (accuracy) como en la métrica de recall al momento de compararse con otros clasificadores similares pero con diferentes combinaciones de entrada.

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	1046.6226	14	74.7587	3.4969	0.0019	2.0374
Dentro de los grupos	641.3404	40	21.3780			
Total	1687.9630	44				

Tabla 5.20: ANOVA STB no concatenada: Accuracy

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	1266.4193	14	90.4585	2.6187	0.013	2.0374
Dentro de los grupos	1036.2842	30	34.5428			
Total	2302.7035	44				

Tabla 5.21: ANOVA STB no concatenada: Recall

También, para verificar mejor su significancia estadística se comparó el modelo seleccionado con otros modelos de diferente configuración pero con igual datos de entradas. Esto para comprobar si además existía significancia estadística a nivel de configuración. Esto conllevó al siguiente resultado

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	1240.022	16	77.5013	2.0285	0.0409	1.9515
Dentro de los grupos	1298.985	34	38.2054			
Total	2539.008	50				

Tabla 5.22: ANOVA STB no concatenada: Accuracy

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	2346.651	16	146.6657	1.8491	0.0650	1.9515
Dentro de los grupos	2696.662	34	79.3135			
Total	5043.313	50				

Tabla 5.23: ANOVA STB no concatenada: Recall

Estos resultados se traducen a que sí existe significancia estadística a nivel de rendimiento para los modelos STB con data no concatenada que utilizan iguales parámetros de entradas (ERP 91 - ERP 102) pero diferente configuración.

Pero no existe significancia estadística a nivel de recall dado que el nivel de significancia obtenido es mayor a 0.05

Los análisis anteriores, sumado a los análisis de significancia estadística realizados tanto para los modelos de red recurrente con datos STB y DBF, es posible mencionar que es poco probable que su buen resultado (superior a 75%) haya sido producido por casualidad a diferencia del resto de los clasificadores seleccionados.

### 5.3. Obtención de Biomarcadores

Para la selección de características se utilizaron todas las combinaciones posibles de entradas a la red neuronal. Esto implica que se usaron todas las combinaciones de ERP posibles pudiendo comprobar así el impacto de tiene cada uno de estos en la clasificación de esquizofrenia.

Dado lo anterior, es posible elegir cuál es la combinación que generaba los mejores resultados y que además perteneciese a un clasificador que se ajustara a los parámetros especificados en la hipótesis para determinar un clasificador válido. Esto determinaría la combinación más relevante a la hora de detectar si una persona padece de esquizofrenia.

Por tanto, para la obtención de los biomarcadores únicamente era necesario fijarse en cuales fueron los ERP, eventos y/o pruebas que se utilizaron como datos de entradas para el mejor clasificador.

En base al mejor clasificador previamente señalado, es posible señalar que los biomarcadores buscados en esta tesis serían los ERP 91 y ERP 102 (en conjunto).

# Capítulo 6

## Conclusión

En base a lo desarrollado a lo largo de este trabajo se puede concluir que es posible obtener un clasificador que permita determinar esquizofrenia en pacientes, con un accuracy de a lo menos 75 %. Y, por tanto, es posible la extracción de biomarcadores para la esquizofrenia a partir de los ERPs y conectividades cerebrales calculadas en base a las pruebas de resting state (RS), double flash (DBF) y stemberg (STB).

A partir de los resultados, se concluye que el mejor clasificador obtenido presenta un accuracy de 79,63% y un recall de 67,5%. Este clasificador se obtiene a partir de una red neuronal recurrente con datos únicamente de la prueba Stemberg, y con una configuración de 10 épocas, sin dropout, 50 neuronas en la capa oculta de la red y utilizando un cross-validation de 4 grupos.

Además se concluye que los biomarcadores para la esquizofrenia serían los ERP 91 y ERP 102, obtenidos a partir de los eventos 91 y 102 de la prueba STB respectivamente, dado que estos fueron los ERPs utilizados en la red previamente señalada.

También es importante señalar que los clasificadores obtenidos con rendimiento dentro del esperando en marco de la hipótesis de esta tesis y por tanto el mejor clasificador señalado cumplen con presentar diferencia estadística en comparación con otros modelos.

Por lo anterior es factible señalar que se cumplió la hipótesis formulada al inicio de esta tesis dado a que sí fue posible la obtención de biomarcadores para la esquizofrenia que permitiesen desarrollar un clasificador con una tasa de acierto superior al 75%.

Además, se logró cumplir con el objetivo general de este trabajo puesto que es posible demostrar la importancia que tiene el potencial evocado y la conectividad cerebral para la detección de esquizofrenia, determinándose que tanto el potencial evocado como la conectividad cerebral permiten determinar si un paciente presenta o no esquizofrenia.

En adición, dado que los modelos con características de ingreso de potencial evocado obtuvieron mejores resultados generales a comparación de la conectividad cerebral, es posible señalar que estos tiene mayor relevancia a la hora de determinar esquizofrenia en pacientes.

Por otra parte, es posible señalar que se cumplieron los objetivos específicos señalados. El

primer y segundo objetivo específico se cumplió en capítulo 3 de esta tesis, puesto que ahí es posible observar el trabajo de pre-procesamiento y limpieza realizada a los datos para luego poder extraer de las señales las distintas conectividades cerebrales y potenciales evocados que estas presentaban, los cuales pasaron a conformar el conjunto de características de los modelos evaluados.

A su vez, los objetivos 3, 4 y 5 se dan como cumplidos en base a lo realizado en el capítulo 4 de esta tesis, puesto que aquí se señalan las distintas arquitecturas elegidas para dar paso a los diversos clasificadores a evaluar, así como los porqué de cada elección y las técnicas utilizadas para generar los distintos grupos de características ingresadas en los modelos diseñados para la clasificación de esquizofrenia.

Finalmente los objetivos 6 y 7 se cumplieron al momento de definir una forma que permitiera determinar el grado de importancia de cada potencial evocado y conectividad cerebral, lo que llevaría a obtener los biomarcadores buscados (capítulo 5 sección 5.3).

También es importante señalar que, en vista de los resultados obtenidos, la conectividad cerebral no presenta buenos resultados a la hora de clasificar esquizofrenia pues que si bien existe un clasificador que cumple con los resultados esperados, en promedio el rendimiento de los clasificadores que utilizaron únicamente componentes de la conectividad cerebral fue de alrededor de un 55 % de accuracy, 20 puntos porcentuales por debajo de la cota inferior establecida en esta tesis.

Por el contrario, el uso del potencial evocado sí presenta buenos resultados generales dado que en promedio, el accuracy obtenido de los distintos modelos evaluados utilizando componentes del potencial evocado fue de aproximadamente un 65 %.

Además, en base a los resultados de los modelos de conectividad cerebral se pudo determinar que si se utilizan las conectividades extraídas a partir de las señales obtenidas bajo la prueba DBF se obtienen mejores resultados (60 % accuracy) en la detección de esquizofrenia frente a aquellas conectividades extraídas bajo las pruebas STB y RS (53 % y 52 % de accuracy, respectivamente). A su vez, la prueba RS y STB presentan resultados prácticamente similares.

De esta forma se puede concluir que el uso de las características extraídas en base a pruebas que utilizan la memoria de trabajo frente a estímulo son mejores para determinar si un paciente presenta o no esquizofrenia que aquellas extraídas de un paciente en estado de reposo.

Por otro lado, basándose en los resultados obtenidos para el uso de potencial evocado se puede determinar que la prueba STB genera resultados mejores que si se utilizan los potenciales evocados extraídos a partir de las señales cerebrales obtenidas de la prueba DBF.

De mano con lo anterior, también es posible señalar que los resultados utilizando únicamente una prueba (solo STB, solo DBF o solo RS) son mejores que los obtenidos al utilizar los potenciales evocados y/o conectividades cerebral de más de una prueba. De esta forma, se puede concluir que utilizar únicamente los potenciales evocados extraídos a partir de la prueba Stemberg permiten clasificar con mejor exactitud si un paciente presenta o no esquizofrenia.

En cuanto al clasificador que concatena la red neuronal recurrente con la red feed forward se puede mencionar que este obtiene mejores resultados generales que los clasificadores evaluados en base únicamente a la red feed forward, pero peores resultados que aquellos que utilizaban únicamente la red recurrente. Es decir, utilizar los datos de los potenciales evocados en conjunto con los de la conectividad cerebral permite clasificar esquizofrenia de mejor forma que si solo se utilizan los datos de la conectividad cerebral, pero peores resultados que si se utilizan únicamente los datos de los potenciales evocados.

Con esto se demuestra además, que el uso únicamente de los potenciales evocados permite clasificar de mejor forma si un paciente presenta esquizofrenia o no.

En base a los resultados obtenidos a partir del uso de las conectividades cerebrales se pudo observar también que la dimensionalidad de los datos de entrada en la red (y por tanto la cantidad de características) afecta negativamente en la clasificación sobretodo en un set de datos tan pequeño como el utilizado en esta tesis, puesto que a mayor número de neuronas en la capa de entrada de la red menor era el rendimiento obtenido por el clasificador.

Esto también concuerda con el hecho de que el uso de potencial evocado generara mejores resultados que el uso de la conectividad cerebral, debido a que la cantidad de neuronas en la primera capa de la red recurrente era cientos de miles de veces menor que la primera capa de la red feed forward, por tanto utilizando la misma cantidad de sujetos de prueba la red Feedforward debía aprender miles de características más que la red recurrente para poder determinar si un paciente presentaba esquizofrenia o no.

Por otra parte, en cuanto a la cantidad de sujetos utilizados en las distintas prueba, es posible señalar que en base al análisis de los resultados la inestabilidad de estos era bastante grande provocando que las distintas configuraciones de clasificadores presentados resultados bastante diversos frente a los mismos datos, obteniéndose desviaciones estándar grandes. Por dicha razón se tuvo que repetir una misma prueba, bajo las mismas condiciones (misma configuración, mismos datos de entrada) varias veces para poder obtener un resultado promedio que fuese más representativo para una prueba, un clasificador y una configuración en particular.

Finalmente, se pueden tener otras conclusiones a partir de análisis realizados a partir de las distintas configuraciones realizadas (número de épocas, cantidad de grupos del cross-validation, concatenación de los datos).

En base a lo anterior se puede mencionar que los clasificadores que utilizaron los datos concatenados presentaron, en promedio, un menor rendimiento que aquellos que utilizaron la data sin concatenar; pese a esto, las diferencias de rendimiento no eran significativamente grandes (aproximadamente una diferencia porcentual de 2% en términos de accuracy) como para concluir plenamente que el uso de data sin concatenar sea absolutamente mejor que el uso de data concatenada.

También es posible señalar que la variación en el número de épocas, más allá del overfitting o underfitting del modelo, no generaba ningún impacto en el rendimiento del clasificador. Es decir, se obtuvieron clasificadores con 6 épocas que presentaban mejores resultados que el mismo clasificador con 10 épocas o 50 épocas, así como también clasificadores con 10 o 50 épocas que presentaban mejores resultados que aquellos con 6 épocas, siendo en ambos casos

resultados válidos (sin overtiffing y sin underfitting).

De esta misma forma, el número de grupos utilizados en la técnica de cross-validation tampoco afectaba en los resultados obtenidos, así como habían buenos resultados para un cross-validation con 3 grupos también habían buenos resultados para un cross-validation con 4, 5 o 6 grupos.

Esta conclusión también aplica para la cantidad de neuronas en la capa oculta, el número de capas de la red Feedforward y el uso de dropout en la red.

## Contribuciones de la tesis

Los aportes de este trabajo se pueden diferenciar en dos grupos.

Por una parte, se tienen los aportes desde un punto de vista científico en base a la literatura señalada con anterioridad. En primer lugar, se destaca el uso conjunto de los potenciales evocados y la conectividad cerebral, para medir su impacto en el proceso de clasificación, dado que no se ha encontrado ningún trabajo que aborde el problema desde dicho punto de vista.

Además, los estudios actuales utilizan aquellos potenciales evocados de interés para el investigador, por lo que se destaca de este trabajo, el utilizar todos los potenciales evocados presentes en las señales EEG, para luego discriminar la importancia de estos en base a los resultados obtenidos.

Por otra parte, el cálculo de la conectividad cerebral, y por ende su uso en el proceso de clasificación, mediante datos obtenidos por señales EEG, es poco utilizado en la literatura actual, siendo más frecuente usar imágenes de resonancia magnética funcional o fMRI<sup>1</sup>, método que conlleva un alto costo operacional y monetario, a diferencia del uso de EEG.

En base a los fundamentos anteriores, es posible notar que el enfoque de esta tesis es relativamente nuevo siendo, a su vez, un estudio con un menor costo de implementación y mayor accesibilidad, debido al uso de EEG.

Por otra parte, se tienen los aportes desde un punto de vista de la utilidad de los resultados esperados en el mundo clínico. Esta contribución se fundamenta en tres focos principales.

En primer lugar, al detectar biomarcadores y características interpretables en el mundo clínico, será posible utilizarlos para atacar el tema de detección de esquizofrenia en un tiempo menor de diagnóstico, provocando que se pueda aplicar un tratamiento temprano al paciente. Lo anterior se fundamenta bajo el hecho de que, en la actualidad, el proceso de diagnóstico contempla un período de evaluación de varios meses para poder determinar si un paciente presenta esquizofrenia o no, antes de iniciar el tratamiento.

Al definir biomarcadores y características interpretables, provocará que este período de

---

<sup>1</sup>funcional magnetic resonance imaging, procedimiento clínico y de investigación que permite mostrar en imágenes las regiones cerebrales activas.

evaluación se reduzca al tiempo de la aplicación del examen para obtener las señales EEG (que son un par de horas [136]), sumado a lo que demore su análisis.

Con esto se puede observar que los tres focos en donde se fundamenta esta tesis para ser de utilidad son: la detección precoz de pacientes con esquizofrenia, la reducción del tiempo de diagnóstico para aplicar un tratamiento y la disminución de costos de aplicación frente al procedimiento actual. A los cuales se le suman los aportes en el área de investigación.

## Trabajo futuro

A partir del estudio realizado en esta tesis se pueden determinar una serie de mejoras y extensiones de la misma.

- **Realización del estudio con una población mayor (mayor cantidad de pacientes).** Lo que permitiría generar resultados más estables y, por tanto, más certeros. Una forma podría ser la implementación de datos sintéticos, los cuales se generan a partir de la data existente.
- **Utilización de los distintos grados de esquizofrenia.** En esta tesis únicamente se diseñó un clasificador que permitiese determinar si una persona presentaba o no esquizofrenia, pero no evaluaba el nivel de avance de la enfermedad. Debido a la importancia en temas médicos de esta diferenciación, sería fundamental el poder clasificar no solo esquizofrenia sino que también el nivel de avance de la misma.
- **Evaluar otras características del potencial evocado y/o conectividad cerebral.** En esta tesis únicamente se utilizaron los potenciales evocados y conectividades cerebrales “puros”, pero también es posible extraer otras características de los mismos, como por ejemplo las características propias de una señal (amplitud de onda, pick, frecuencia, etc), que podrían incidir significativamente en la detección de esquizofrenia.
- **Reducir la dimensionalidad de los datos de conectividad cerebral.** En esta tesis se utilizó una forma específica (la más común) para calcular la conectividad, así como también se utilizaron todos los electrodos disponibles para su cálculo. Pero también es factible cambiar tanto la forma de calcular la conectividad, por ejemplo utilizando otro método, como también los electrodos a utilizar para su cálculo, por ejemplo, en lugar de utilizar los 64 electrodos disponibles utilizar únicamente uno por zona cerebral o elegir las zonas cerebrales que más impactan en la esquizofrenia y utilizar dichos electrodos. Otra forma factible es realizar un downsampling menor a 256Hz.
- **Probar el efecto del padding.** En este caso, si bien se mencionó el uso del zero-padding no se probó a cabalidad debido a que se probó solo una forma de concatenación de datos (aumentando cantidad de filas), la cual no necesitaba el uso del zero-padding dado que no había diferencia en la cantidad de columnas de los datos. Por lo cual se puede aún probar el efecto de concatenar los datos aumentando número de columnas y también probando dos tipos de zero-padding: agregando ceros al inicio y agregando ceros al final. Esto podría, como no, generar nuevos resultados y conclusiones.

- **Identificar y probar otros modelos.** En este caso se utilizó únicamente una red feed forward y una recurrente, además de la concatenación de ambas; pero también podría ser factible el uso de otro tipo de red, por ejemplo una convolucional para el caso de la conectividad cerebral y/o los potenciales evocados o el diseño de un nuevo tipo de red.

# Bibliografía

- [1] Priyanka A. Abhang, Bharti W. Gawali, and Suresh C. Mehrotra. Technological basics of EEG recording and operation of apparatus. In *Introduction to EEG- and Speech-Based Emotion Recognition*, pages 19–50. Elsevier, 2016.
- [2] Jane Adams, Steven F. Faux, Paul G. Nestor, Martha Shenton, Brian Marcy, Scott Smith, and Robert W. McCarley. ERP abnormalities during semantic processing in schizophrenia. *Schizophrenia Research*, 10(3):247–257, October 1993.
- [3] Majid Al-gburi. Can someone recommend what is the best percent of divided the training data and testing data in neural network 75:25 or 80:20 or 90:10 ?
- [4] Rubiales Alberto. ¿qué es underfitting y overfitting?
- [5] Mental Health America. Esquizofrenia: Lo que usted necesita saber.
- [6] Xiu An, Deping Kuang, Xiaojiao Guo, Yilu Zhao, and Lianghua He. A deep learning method for classification of EEG data based on motor imagery. In *Intelligent Computing in Bioinformatics*, pages 203–210. Springer International Publishing, 2014.
- [7] Mohammad Arbabshirani, Kent A Kiehl, Godfrey Pearlson, and Vince Calhoun. Classification of schizophrenia patients based on resting-state functional network connectivity. *Frontiers in neuroscience*, 7:133, 07 2013.
- [8] Juan Ignacio Barrios Arce. La matriz de confusión y sus métricas.
- [9] M Molina Arias. ¿qué significa realmente el valor de p?
- [10] Arimetrics. Qué es framework.
- [11] Juan Arribas, Vince Calhoun, and Tülay Adalı. Automatic bayesian classification of healthy controls, bipolar disorder, and schizophrenia using intrinsic connectivity maps from fmri data. *IEEE transactions on bio-medical engineering*, 57:2850–60, 09 2010.
- [12] Mohamed Attia, Imali Hettiarachchi, Mohammed Hossny, and Saeid Nahavandi. A time domain classification of steady-state visual evoked potentials using deep recurrent-convolutional neural networks. In *2018 IEEE 15th International Symposium on Bio-medical Imaging (ISBI 2018)*, pages 766–769, 2018.
- [13] AWS. Validación cruzada.

- [14] Gi-Yeul Bae, Carly J. Leonard, Britta Hahn, James M. Gold, and Steven J. Luck. Assessing the information content of ERP signals in schizophrenia using multivariate decoding methods. *NeuroImage: Clinical*, 25:102179, 2020.
- [15] Yi Ji Bae, Midan Shim, and Won Hee Lee. Schizophrenia detection using machine learning approach from social media content. *Sensors*, 21(17):5924, September 2021.
- [16] Sara Bagherzadeh, Mohsen Sadat Shahabi, and Ahmad Shalhaf. Detection of schizophrenia using hybrid of deep learning and brain effective connectivity image from electroencephalogram signal. *Computers in Biology and Medicine*, 146:105570, July 2022.
- [17] Karim Pichara Baksai. Introducción a la minería de datos, 2019.
- [18] Instituto Psiquiátrico “Dr. José Horwitz Barak”. Esquizofrenia.
- [19] Carla Barros, Brian Roach, Judith M. Ford, Ana P. Pinheiro, and Carlos A. Silva. From sound perception to automatic detection of schizophrenia: An EEG-based deep learning approach. *Frontiers in Psychiatry*, 12, February 2022.
- [20] Carla Barros, Carlos A. Silva, and Ana P. Pinheiro. Advanced EEG-based learning approaches to predict schizophrenia: Promises and pitfalls. *Artificial Intelligence in Medicine*, 114:102039, April 2021.
- [21] Pouya Bashivan, Irina Rish, M Yeasin, and Noel Codella. Learning representations from eeg with deep recurrent-convolutional neural networks. 11 2015.
- [22] Kerstin Bendfeldt, Renata Smieskova, Nikolaos Koutsouleris, Stefan Klöppel, André Schmidt, Anna Walter, Fabienne Schönborn-Harrisberger, Johannes Wrege, Andor Simon, Bernd Taschler, Thomas Nichols, Anita Riecher-Rössler, Undine Lang, Ernst-Wilhelm Radue, and Stefan Borgwardt. Classifying individuals at high-risk for psychosis based on functional brain activity during working memory processing. *Clinical neuroimaging*, 9:555–563, 10 2015.
- [23] Lukas Biewald. Recurrent neural networks.
- [24] Top BigData. Cómo elegir una función de activación para el aprendizaje profundo.
- [25] Nicholas M. Blauch, Marlene Behrmann, and David C. Plaut. A connectivity-constrained computational account of topographic organization in primate high-level visual cortex. *Proceedings of the National Academy of Sciences*, 119(3), January 2022.
- [26] Reza Boostani and Malihe Sabeti. Optimising brain map for the diagnosis of schizophrenia. *International Journal of Biomedical Engineering and Technology*, 28(2):105, 2018.
- [27] Reza Boostani, Khadijeh Sadatnejad, and Malihe Sabeti. An efficient classifier to diagnose of schizophrenia based on the eeg signals. *Expert Syst. Appl.*, 36:6492–6499, 04 2009.

- [28] Nashaat Boutros, George Zouridakis, Terry Rustin, Cecilia Peabody, and Dhyanne Warner. The p50 component of the auditory evoked potential and subtypes of schizophrenia. *Psychiatry Research*, 47(3):243 – 254, 1993.
- [29] Jason Brownlee. A gentle introduction to k-fold cross-validation.
- [30] Jason Brownlee. A gentle introduction to the rectified linear unit (relu).
- [31] Jason Brownlee. How to avoid overfitting in deep learning neural networks.
- [32] Jason Brownlee. Softmax activation function with python.
- [33] Jason Brownlee. What is the difference between test and validation datasets?
- [34] Ed Burns. Aprendizaje profundo (deep learning).
- [35] Diego Calvo. Red neuronal convolucional cnn, 2017.
- [36] Diego Calvo. Red neuronal recurrente - rnn, 2018.
- [37] Bo Cao, Raymond Y. Cho, Dachun Chen, Meihong Xiu, Li Wang, Jair C. Soares, and Xiang Yang Zhang. Treatment response prediction and individualized identification of first-episode drug-naïve schizophrenia using brain functional connectivity. *Molecular Psychiatry*, 25(4):906–913, June 2018.
- [38] Fernando Sancho Caparrini. Entrenamiento de redes neuronales: mejorando el gradiente descendente.
- [39] Vasiliki Chatzi, Rui Pedro Teixeira, John Shawe-Taylor, Andre Altmann, Owen G. O’ Daly, Daan Christiaens, Jessica Schrouff, and Jacques-Donald Tournier. Grey matter biomarker identification in schizophrenia: detecting regional alterations and their underlying substrates. *NeuroImage: Clinical First*, 06 2018.
- [40] Zineb Cheker, Saad Chakkor, and Ahmed El Oualkadi. Visual evoked potential classification support with convolutional neural network and recurrent neural network - a comparative study. In *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, pages 1612–1617, 2022.
- [41] Leo Chu, Robert Qiu, Haichun Liu, Zenan Ling, and Xin Shi. Individual recognition in schizophrenia using deep learning methods with random forest and voting classifiers: Insights from resting state eeg streams. *CoRR*, 06 2017.
- [42] Claudio Ciprian, Kirill Masychev, Maryam Ravan, Akshaya Manimaran, and AnkitaAmol Deshmukh. Diagnosing schizophrenia using effective connectivity of resting-state EEG data. *Algorithms*, 14(5):139, April 2021.
- [43] Jose A. Cortes-Briones, Nicolas I. Tapia-Rivas, Deepak Cyril D'Souza, and Pablo A. Estevez. Going deep into schizophrenia with artificial intelligence. *Schizophrenia Research*, 245:122–140, July 2022.
- [44] Georgina Cosma, David Brown, Matthew Archer, Masood Khan, and A. Graham Pocoley. A survey on computational intelligence approaches for predictive modeling in prostate cancer. *Expert Systems with Applications*, 70:1–19, March 2017.

- [45] I. Torres Courchoud and J.I. Pérez Calvo. Biomarcadores y práctica clínica.
- [46] Alexander Craik, Yongtian He, and Jose Luis (Pepe) Contreras-Vidal. Deep learning for electroencephalogram (eeg) classification tasks: A review. *Journal of Neural Engineering*, 16, 02 2019.
- [47] Centros Comunitarios de Aprendizaje. Proceso de clasificación.
- [48] Universitat de Barcelona. Análisis de la varianza con un factor (anova).
- [49] NIH: Instituto Nacional de la Salud Mental. Esquizofrenia, 2019.
- [50] Comunicaciones Hospital Clínico de la Universidad de Chile. Clínica psiquiátrica implementa innovador programa de detección precoz de la esquizofrenia, 2016.
- [51] Facultad de Medicina UNAM. Potenciales relacionados con eventos.
- [52] Ministerio de Salud (Minsal) Gobierno de Chile. Diseño de provisión para la atención de personas con trastornos psiquiátricos severos: Esquizofrenia.
- [53] Ministerio de Salud (Minsal) Gobierno de Chile. Orientaciones técnicas: Tratamiento y rehabilitación de las personas afectadas de esquizofrenia.
- [54] Ministerio de Salud (Minsal) Gobierno de Chile. Guías clínicas auge: Tratamiento de personas desde el primer episodio de esquizofrenia, 2017.
- [55] Scott L. Decker, Paul T. Fillmore, and Alycia M. Roberts. Coherence: The measurement and application of brain connectivity. *NeuroRegulation*, 4(1):3–13, March 2017.
- [56] DeepAI. Feature selection.
- [57] DeepAI. Feed forward neural network.
- [58] Lorenzo Del Fabro, André Schmidt, Lydia Fortea, Giuseppe Delvecchio, Armando D’Agostino, Joaquim Radua, Stefan Borgwardt, and Paolo Brambilla. Functional brain network dysfunctions in subjects at high-risk for psychosis: A meta-analysis of resting-state functional connectivity. *Neuroscience & Biobehavioral Reviews*, 128:90–101, 2021.
- [59] Dra. Ma. del Pilar Gómez Gil. C261-69 tópicos avanzados: Redes neuronales artificiales conceptos básicos.
- [60] Arnaud Delorme. Infomax independent component analysis for dummies.
- [61] Christ Devia, Rocío Mayol Troncoso, Javiera Parrini, Gricel Orellana, Aida Ruiz A, Pedro Maldonado, and Josi Ignacio Egana. Eeg classification during scene free-viewing for schizophrenia detection. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, PP:1–1, 04 2019.
- [62] Science Direct. Phase classification.
- [63] Tali Ditman and Gina Kuperberg. The time course of building discourse coherence in schizophrenia: An erp investigation. *Psychophysiology*, 44:991–1001, 12 2007.

- [64] Daniel Durstewitz, Georgia Koppe, and Andreas Meyer-Lindenberg. Deep neural networks in psychiatry. *Molecular Psychiatry*, page 1, 02 2019.
- [65] Jaime Durán. Técnicas de regularización básicas para redes neuronales.
- [66] Hansen EC, Battaglia D, Spiegler A, Deco G, and Jirsa VK. La conectividad funcional dinámica, nuevo biomarcador de enfermedades mentales, 2015.
- [67] IBM Cloud Education. What is overfitting?
- [68] EEGlab. Background on eeg references.
- [69] EEGlab. Preprocess data.
- [70] Emotiv. Evoked response potential.
- [71] Elsa Santos Febles, Marlis Ontivero Ortega, Michell Valdés Sosa, and Hichem Sahli. Machine learning techniques for the diagnosis of schizophrenia based on event related potentials. March 2022.
- [72] Elsa Santos Febles, Marlis Ontivero Ortega, Michell Valdés Sosa, and Hichem Sahli. Machine learning techniques for the diagnosis of schizophrenia based on event related potentials. March 2022.
- [73] FieldTrip. Event-related averaging and meg planar gradient.
- [74] FieldTrip. ft\_freqanalysis.
- [75] FieldTrip. ft\_timelockanalysis.
- [76] FieldTrip. Independent component analysis (ica) to remove eeg artifacts.
- [77] Fieldtrip. Introduction on dealing with artifacts.
- [78] todo sobre fisioterapia Fisioonline. Definición - qué es estímulos.
- [79] Judith M. Ford, Holly K. Hamilton, Katiah Llerena, Brian J. Roach, Daniel H. Mathalon, Judith M. Ford, Holly K. Hamilton, Katiah Llerena, Brian J. Roach, and Daniel H. Mathalon. 225C26Neurophysiologic biomarkers of psychosis: Event-related potential biomarkers. In *Psychotic Disorders: Comprehensive Conceptualization and Treatments*. Oxford University Press, 11 2020.
- [80] Erica Gabriel. How to treat overfitting in convolutional neural networks.
- [81] Wen-Jun Gao, Sha-Sha Yang, Nancy R. Mack, and Linda A. Chamberlin. Aberrant maturation and connectivity of prefrontal cortex in schizophrenia—contribution of NMDA receptor development and hypofunction. *Molecular Psychiatry*, 27(1):731–743, June 2021.
- [82] Evgenia Gatov, Laura Rosella, Maria Chiu, and Paul Kurdyak. Trends in standardized mortality among individuals with schizophrenia, 1993-2012: A population-based, repeated cross-sectional study. *Canadian Medical Association Journal*, 189:E1177–E1187, 09 2017.

- [83] Yulia Gavrilova. Machine learning testing: A step to perfection.
- [84] Majed Ghanbari, Abdol Hamid Pilevar, and Narges Bathaeian. Diagnosis of schizophrenia using brain resting-state fMRI with activity maps based on deep learning. *Signal, Image and Video Processing*, June 2022.
- [85] Kelsey Graber, Michelle Bosquet Enlow, Frank H Duffy, Eugene D’Angelo, Georgios Sideridis, Damon E Hyde, Nicholas Morelli, Sahil Tembulkar, and Joseph Gonzalez-Heydrich. P300 amplitude attenuation in high risk and early onset psychosis youth. *Schizophrenia Research*, 01 2019.
- [86] J. L. Guiñón, E. Ortega, Jordi García-Antón Aviñó, and Valentín Pérez Herranz. Filtrado de señales (i).
- [87] Andres H Neuhaus, Florin C Popescu, John A Bates, Terry E Goldberg, and Anil K Malhotra. Single-subject classification of schizophrenia using event-related potentials obtained during auditory and visual oddball paradigms. *European archives of psychiatry and clinical neuroscience*, 263, 05 2012.
- [88] Andres H Neuhaus, Florin C Popescu, Cristian Grozea, Eric Hahn, Constanze Hahn, Carolin Opgen-Rhein, Carsten Urbanek, and Michael Dettling. Single-subject classification of schizophrenia by event-related potentials during selective attention. *Neuro-Image*, 55:514–21, 03 2011.
- [89] HHA Halim, SB Samdin, CM Ting, SH Salleh, YF Low, and AM Noor. Estimating brain connectivity in schizophrenia patients using electroencephalography signals. *International Medical Device and Technology Conference (IMEDITEC)*, 2017.
- [90] Arthur Hamilton and Georg Northoff. Abnormal ERPs and brain dynamics mediate basic self disturbance in schizophrenia: A review of EEG and MEG studies. *Frontiers in Psychiatry*, 12, April 2021.
- [91] Tong He, Ru Kong, Avram J. Holmes, Minh Nguyen, Mert R. Sabuncu, Simon B. Eickhoff, Danilo Bzdok, Jiashi Feng, and B.T. Thomas Yeo. Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics. *NeuroImage*, 206:116276, February 2020.
- [92] Tong He, Ru Kong, Avram J. Holmes, Mert R. Sabuncu, Simon B. Eickhoff, Danilo Bzdok, Jiashi Feng, and B. T. Thomas Yeo. Is deep learning better than kernel regression for functional connectivity prediction of fluid intelligence? In *2018 International Workshop on Pattern Recognition in Neuroimaging (PRNI)*, pages 1–4, 2018.
- [93] HeavyAI. Feature selection.
- [94] Jose Martinez Heras. Accuracy (exactitud).
- [95] Shannon Hilbert. Fft zero padding.
- [96] Tsung Hao Hsieh, Kuan Yi Wu, and Sheng Fu Liang. Identification of schizophrenic patients and healthy controls based on musical perception using AEP analysis. *Neuropsychiatry*, 08(04), 2018.

- [97] Mengjiao Hu, Xing Qian, Siwei Liu, Amelia Jialing Koh, Kang Sim, Xudong Jiang, Cuntai Guan, and Juan Helen Zhou. Structural and diffusion MRI based schizophrenia classification using 2d pretrained and 3d naive convolutional neural networks. *Schizophrenia Research*, 243:330–341, May 2022.
- [98] Heinz Häfner and Kurt Maurer. Early detection of schizophrenia: current evidence and future perspectives. *World psychiatry : official journal of the World Psychiatric Association (WPA)*, 5(3):130–138, 10 2006.
- [99] IBM. Significación estadística.
- [100] IBM. What are recurrent neural networks?
- [101] IBM. ¿qué es machine learning?
- [102] Artificial Intelligence. How can neural networks deal with varying input sizes?
- [103] Matías Irrarrázaval, Fernanda Prieto, and Julio Armijo. Prevención e intervenciones tempranas en salud mental: una perspectiva internacional. *Acta bioethica*, 22:37 – 50, 06 2016.
- [104] Frick Janek, Thilo Rieg, and Ricardo Buettner. Detection of schizophrenia: A machine learning algorithm for potential early detection and prevention based on event-related potentials. In *Proceedings of the Annual Hawaii International Conference on System Sciences*. Hawaii International Conference on System Sciences, 2021.
- [105] Jing Jin and Andrzej Cichocki. Evoked response potentials. In *Encyclopedia of Animal Cognition and Behavior*, pages 1–3. Springer International Publishing, October 2017.
- [106] TZYU-PING JUNG, SCOTT MAKEIG, COLIN HUMPHRIES, TE-WON LEE, MARTIN J. McKEOWN, VICENTE IRAGUI, and TERRENCE J. SEJNOWSKI. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(2):163–178, 2000.
- [107] Tzyy-Ping Jung, Scott Makeig, Marissa Westerfield, Jeanne Townsend, Eric Courchesne, and Terrence J. Sejnowski. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, 111(10):1745–1758, October 2000.
- [108] Chrissy Kidd. Deep learning vs machine learning: What’s the difference?, 2019.
- [109] Palani Thanaraj Krishnan, Alex Noel Joseph Raj, Parvathavarthini Balasubramanian, and Yuanzhu Chen. Schizophrenia detection using Multivariate Empirical mode decomposition and entropy measures from multichannel EEG signal. *Biocybernetics and Biomedical Engineering*, 40(3):1124–1139, July 2020.
- [110] V Santhosh Kumar and TVK Rao. Functional brain connectivity analysis using coherent measures. In *EMBEC & NBC 2017*, pages 779–782. Springer, 2017.
- [111] Miguel Díaz Kusztrich. Redes neuronales recurrentes y series temporales, 2016.

- [112] No-Sang Kwak, Klaus-Robert Müller, and Seong-Whan Lee. A convolutional neural network for steady state visual evoked potential classification under ambulatory environment. *PLOS ONE*, 12:e0172578, 02 2017.
- [113] Chi-Yu Lai, Elizabeth Scarr, Madhara Udawela, Ian Everall, Wei Chen, and Brian Dean. Biomarkers in schizophrenia: A focus on blood based diagnostics and theranostics. *World Journal of Psychiatry*, 6:102, 03 2016.
- [114] Nicole Laskowski. recurrent neural networks.
- [115] Jorne Laton, Jeroen Van Schependom, Jeroen Gielen, Jeroen Decoster, Tim Moons, Jacques De Keyser, Marc Hert, and Guy Nagels. Single-subject classification of schizophrenia patients based on a combination of oddball and mismatch evoked potential paradigms. *Journal of the Neurological Sciences*, 347, 10 2014.
- [116] Scikit learn. 3.1. cross-validation: evaluating estimator performance.
- [117] Aprende Machine Learning. Qué es overfitting y underfitting y cómo solucionarlo.
- [118] Seung-Hwan Lee, M Thiam Chye Lim, Jung Seo Yi, Dowon Kim, and Chang-Hwan Im. Disruptions in small-world cortical functional connectivity network during an auditory oddball paradigm task in patients with schizophrenia. *Schizophrenia Research*, 153:s201–s202, 2014.
- [119] Servicio Médico Legal. El suicidio en Chile: Análisis del fenómeno desde los datos médicos legales. período 2000-2010, 2018.
- [120] A.D. Legatt. Evoked potentials. In *Encyclopedia of the Neurological Sciences*, pages 228–231. Elsevier, 2014.
- [121] Cecilia Valenzuela León. Determinan potenciales marcadores de esquizofrenia en población de riesgo, 2018.
- [122] S.J. Luck. *An Introduction to the Event-Related Potential Technique, second edition*. A Bradford Book. MIT Press, 2014.
- [123] S.J. Luck. *An Introduction to the Event-Related Potential Technique, second edition*. A Bradford Book. MIT Press, 2014.
- [124] S.J. Luck. *An Introduction to the Event-Related Potential Technique, second edition*. A Bradford Book. MIT Press, 2014.
- [125] S.J. Luck. *An Introduction to the Event-Related Potential Technique, second edition*. A Bradford Book. MIT Press, 2014.
- [126] Mary-Ellen Lynall, Danielle S. Bassett, Robert Kerwin, Peter J. McKenna, Manfred Kitzbichler, Ulrich Muller, and Ed Bullmore. Functional connectivity and brain networks in schizophrenia. *Journal of Neuroscience*, 30(28):9477–9487, 2010.
- [127] Sambit Mahapatra. Why deep learning over traditional machine learning?, 2018.

- [128] Kasra Manshaei. How to know if a model is overfitting or underfitting by looking at graph.
- [129] MathWorks. Deep learning, tres cosas que es necesario saber.
- [130] MathWorks. Redes neuronales convolucionales.
- [131] Joseph Patrick McEvoy. The importance of early treatment of schizophrenia. *Behavioral Healthcare*, 27(4):40–43, 4 2007.
- [132] información de salud para usted MedlinePlus. Electroencefalograma.
- [133] Gobierno de Chile Ministerio de Salud. Enfermedad de parkinson, 2010.
- [134] Minitab. ¿qué es anova?
- [135] Organización médica colegial de España. La esperanza de vida de personas con esquizofrenia es 15 años menor a la general, 2018.
- [136] David Saceda Corralo (Médico). Estudio de potenciales evocados: Preparación para el estudio de potenciales evocados, 2018.
- [137] Instituto nacional del cáncer. Electrodo.
- [138] M A Niznikiewicz, B F O'Donnell, P G Nestor, L Smith, S Law, M Karapelou, M E Shenton, and R W McCarley. Erp assessment of visual and auditory language processing in schizophrenia. *Journal of abnormal psychology*, 106:85–94, 2 1997.
- [139] Brian Odonnell, Dean Salisbury, M.A. Niznikiewicz, Colleen Brenner, and Jenifer Vohs. Abnormalities of event-related potential components in schizophrenia. *The Oxford Handbook of Event-Related Potential Components*, 01 2012.
- [140] NIH: National Institute of Mental Health. La esquizofrenia, 2017.
- [141] Shu Lih Oh, Yuki Hagiwara, Raghavendra U, Yuvaraj Rajamanickam, Arunkumar, and U Rajendra Acharya. A deep learning approach for parkinson's disease diagnosis from eeg signals. *Neural Computing and Applications*, 08 2018.
- [142] Pierre Orban, Martin Desseilles, Adrianna Mendrek, Josiane Bourque, Pierre Bellec, and Emmanuel Stip. Altered brain connectivity in patients with schizophrenia is consistent across cognitive contexts. *Journal of Psychiatry and Neuroscience*, 42(1):17–26, 1 2017.
- [143] María Graciela López Ordieres. Tratamientos farmacológicos de la esquizofrenia, 2017.
- [144] Martijn P van den Heuvel, Olaf Sporns, Guusje Collin, Thomas Scheewe, René Mandl, Wiepke Cahn, Joaquín Goñi, Hilleke Pol, and René S Kahn. Abnormal rich club organization and functional brain dynamics in schizophrenia. *JAMA psychiatry (Chicago, Ill.)*, 70:1–10, 06 2013.

- [145] Pinkal Patel, Priya Aggarwal, and Anubha Gupta. Classification of schizophrenia versus normal subjects using deep learning. In *Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, ICVGIP '16*, pages 28:1–28:6, New York, NY, USA, 2016. ACM.
- [146] David Weenink Paul Boersma. Feedforward neural networks 1. what is a feedforward neural network?
- [147] Abdelmalik Moujahid Pedro Larranaga, Iñaki Inza. Tema 8. redes neuronales.
- [148] Veronica Perez, Scott Woods, Brian Roach, Judith Ford, Thomas Mcglashan, Vinod Srihari, and Daniel Mathalon. Automatic auditory processing deficits in schizophrenia and clinical high-risk patients: Forecasting psychosis risk with mismatch negativity. *Biological psychiatry*, 75, 09 2013.
- [149] Arthur Petrosian, Danil Prokhorov, Richard Homan, Richard Dasheiff, and Donald Wunsch. Recurrent neural network based prediction of epileptic seizures in intra- and extracranial eeg. *Neurocomputing*, 30:201–218, 01 2000.
- [150] Chun-Ren Phang, Chee-Ming Ting, Fuad Noman, and Hernando Ombao. Classification of eeg-based brain connectivity networks in schizophrenia using a multi-domain connectome convolutional neural network, 2019.
- [151] J. M. Phillips, C. R. Maxwell, R. S. Ehrlichman, and S. J. Siegel. Event-related potentials (ERPs) in the study of schizophrenia: How preclinical ERP studies have contributed to our understanding of schizophrenia. In *Handbook of Neurochemistry and Molecular Neurobiology*, pages 525–543. Springer US, 2009.
- [152] quora. How can neural networks deal with varying input sizes?
- [153] Md Abdur Rahaman, Jiayu Chen, Zening Fu, Noah Lewis, Armin Iraj, and Vince D. Calhoun. Multi-modal deep learning of functional and structural neuroimaging and genomic data to predict mental illness. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, November 2021.
- [154] Md Abdur Rahaman, Jiayu Chen, Zening Fu, Noah Lewis, Armin Iraj, and Vince D. Calhoun. Multi-modal deep learning of functional and structural neuroimaging and genomic data to predict mental illness. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, November 2021.
- [155] T. K. Rajji, Z. Ismail, and B. H. Mulsant. Age at onset and cognition in schizophrenia: meta-analysis. *British Journal of Psychiatry*, 195(4):286–293, 2009.
- [156] Paloma Recuero. Artificial intelligence, machine learning y deep learning: ¿conoces las diferencias?, 2017.
- [157] DIEGO ANDRÉS CANDIA RIVERA. Caracterización de la respuesta emocional ante estímulos visuales en registros electroencefalográficos.

- [158] Joaquín Amat Rodrigo. Kruskal-wallis test, 2016.
- [159] José Andrés Rodríguez. La esquizofrenia disminuye la esperanza de vida, 2012.
- [160] Edmund T. Rolls, Wei Cheng, and Jianfeng Feng. Brain dynamics: the temporal variability of connectivity, and differences in schizophrenia and ADHD. *Translational Psychiatry*, 11(1), January 2021.
- [161] Yannick Roy, Hubert Banville, Isabela Albuquerque, Alexandre Gramfort, Tiago H Falk, and Jocelyn Faubert. Deep learning-based electroencephalography analysis: a systematic review. *Journal of Neural Engineering*, 05 2019.
- [162] Patricia Ryser-Welch. Is there an ideal ratio between a training set and validation set? which trade-off would you suggest?
- [163] Jorge Dagnino S. Análisis de varianza.
- [164] Janssen-Cilag S.A. ¿cuánto tiempo tiene que pasar para recibir el diagnóstico?
- [165] Janssen-Cilag S.A. La esquizofrenia puede confundirse con otras enfermedades, 2019.
- [166] Janssen-Cilag S.A. Test de esquizofrenia, 2019.
- [167] Janssen-Cilag S.A. Tratando la esquizofrenia, 2019.
- [168] Janssen-Cilag S.A. ¿cómo descubre el psiquiatra lo que le ocurre?, 2019.
- [169] Janssen-Cilag S.A. ¿quienes pueden tener riesgo de padecerla?, 2019.
- [170] Janssen-Cilag S.A. ¿qué enfermedades producen el trastorno psicótico?, 2019.
- [171] Mehreen Saeed. A gentle introduction to sigmoid function.
- [172] I. Sagdinc, S. Kirac, M. Engin, K. Erkan, and E. Butun. Dynamic eeg modeling and single-evoked potential extraction using real-time recurrent neural network. In *Proceedings of the 1998 IEEE International Conference on Control Applications (Cat. No.98CH36104)*, volume 1, pages 358–362 vol.1, 1998.
- [173] N.J. Sairamya, M.S.P. Subathra, and S. Thomas George. Automatic identification of schizophrenia using EEG signals based on discrete wavelet transform and RLNDiP technique with ANN. *Expert Systems with Applications*, 192:116230, April 2022.
- [174] N.J. Sairamya, M.S.P. Subathra, and S. Thomas George. Automatic identification of schizophrenia using EEG signals based on discrete wavelet transform and RLNDiP technique with ANN. *Expert Systems with Applications*, 192:116230, April 2022.
- [175] Jerónimo Saiz Ruiz, Diego C. de la Vega Sánchez, and Patricia Sánchez Páez. Bases Neurobiológicas de la Esquizofrenia. *Clínica y Salud*, 21:235 – 254, 2010.
- [176] S. Charles Schulz. Esquizofrenia, 2016.
- [177] Noushath Shaffi. What is the difference between validation set and test set?

- [178] Rahil Shaikh. Feature selection techniques in machine learning with python.
- [179] Nadezhda Shanarova, Marina Pronina, Mikhail Lipkovich, and Juri Kropotov. Machine learning based diagnostics of schizophrenia patients. In *2022 6th Scientific School Dynamics of Complex Networks and their Applications (DCNA)*. IEEE, September 2022.
- [180] Han Shaoqiang, Wei Huang, Yan Zhang, Jingping Zhao, and Huafu Chen. Recognition of early-onset schizophrenia using deep-learning method. *Applied Informatics*, 4, 12 2017.
- [181] Hui Shen, Lubin Wang, Yadong Liu, and Dewen Hu. Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fmri. *NeuroImage*, 49:3110–21, 11 2009.
- [182] Siuly Siuly, Smith K. Khare, Varun Bajaj, Hua Wang, and Yanchun Zhang. A computerized method for automatic detection of schizophrenia using EEG signals. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(11):2390–2400, November 2020.
- [183] Pawel Skudlarski, Kanchana Jagannathan, Karen Anderson, Michael Stevens, Vince Calhoun, Beata Skudlarska, and Godfrey Pearlson. Brain connectivity is not only lower but different in schizophrenia: A combined anatomical and functional approach. *Biological psychiatry*, 68:61–9, 05 2010.
- [184] Julius O. Smith. *Spectral Audio Signal Processing*. <http://ccrma.stanford.edu/~jos/sasp/>, accessed <date>. online book, 2011 edition.
- [185] Luis Carlos Torres Soler. Redes neuronales artificiales.
- [186] Marco Solmi, Joaquim Radua, Miriam Olivola, Enrico Croce, Livia Soardo, Gonzalo Salazar de Pablo, Jae Il Shin, James B. Kirkbride, Peter Jones, Jae Han Kim, Jong Yeob Kim, André F. Carvalho, Mary V. Seeman, Christoph U. Correll, and Paolo Fusar-Poli. Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. *Molecular Psychiatry*, 27(1):281–295, June 2021.
- [187] Dr. Jesús Paniagua Soto. Potenciales evocados.
- [188] Srivathsan Srinivasagopalan, Justin Barry, Varadraj Gurupur, and Sharma Thankachan. A deep learning approach for diagnosing schizophrenic patients. *Journal of Experimental & Theoretical Artificial Intelligence*, pages 1–14, 01 2019.
- [189] Srivathsan Srinivasagopalan, Justin Barry, Varadraj Gurupur, and Sharma Thankachan. A deep learning approach for diagnosing schizophrenic patients. *Journal of Experimental & Theoretical Artificial Intelligence*, 0(0):1–14, 2019.
- [190] Luca Steardo, Elvira Anna Carbone, Renato de Filippis, Claudia Pisanu, Cristina Segura-Garcia, Alessio Squassina, Pasquale De Fazio, and Luca Steardo. Application of support vector machine on fMRI data as biomarkers in schizophrenia diagnosis: A systematic review. *Frontiers in Psychiatry*, 11, June 2020.

- [191] Lisha Sun, Ying Liu, and P.J. Beadle. Independent component analysis of eeg signals. In *Proceedings of 2005 IEEE International Workshop on VLSI Design and Video Technology, 2005.*, pages 219–222, 2005.
- [192] Rebeca Sánchez. Que es combinacion en matematicas.
- [193] Shawhin Talebi. Independent component analysis (ica).
- [194] Yan Tang, Lifeng Wang, Fang Cao, and Liwen Tan. Identify schizophrenia using resting-state functional connectivity: an exploratory research and analysis. *BioMedical Engineering OnLine*, 11(1):50, 08 2012.
- [195] Christopher Tao. How to evaluate a classification machine learning model. an introduction of accuracy, precision, roc/auc and logistic loss.
- [196] J.A. Taylor, N. Matthews, P.T. Michie, M.J. Rosa, and M.I. Garrido. Auditory prediction errors as individual biomarkers of schizophrenia. *NeuroImage: Clinical*, 15:264–273, 2017.
- [197] Alaa Tharwat. Independent component analysis: An introduction. *Applied Computing and Informatics*, 17(2):222–249, August 2020.
- [198] John Thomas, Tomasz Maszczyk, Nishant Sinha, Tilmann Kluge, and Justin Dauwels. Deep learning-based classification for brain-computer interfaces. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, October 2017.
- [199] Mirjam Van Tricht, Stephan Ruhrmann, Martijn Arns, Ralf Müller, Mitja Bodatsch, Eva Velthorst, Johannes H. T. M. Koelman, Lo J. Bour, Katharina Zurek, Frauke Schultze-Lutter, Joachim Klosterkötter, Don Linszen, Lieuwe de Haan, Anke Brockhaus-Dumke, and Dorien Nieman. Can quantitative eeg measures predict clinical outcome in subjects at clinical high risk for psychosis? a prospective multicenter study. *Schizophrenia Research*, 153, 03 2014.
- [200] Christos Vasios, Charalabos Papageorgiou, George Matsopoulos, Konstantina Nikita, and Nikolaos Uzunoglu. A decision support system of evoked potentials for the classification of patients with first-episode schizophrenia. *German Journal of Psychiatry*, 5, 01 2002.
- [201] Baihan Wang, Eirini Zartaloudi, Jennifer F. Linden, and Elvira Bramon. Neurophysiology in psychosis: The quest for disease biomarkers. *Translational Psychiatry*, 12(1), March 2022.
- [202] Rui Wang, Min S. H. Aung, Saeed Abdullah, Rachel Brian, Andrew T. Campbell, Tanzeem Choudhury, Marta Hauser, John Kane, Michael Merrill, Emily A. Scherer, Vincent W. S. Tseng, and Dror Ben-Zeev. CrossCheck. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, September 2016.
- [203] Wikipedia. Coherence (signal processing).
- [204] Wikipedia. Multitaper.

- [205] Wikipedia. Spectral concentration problem.
- [206] Wikipedia. Red neuronal recurrente, 2019.
- [207] Jason Zalev and Michael C. Kolios. Detecting abnormal vasculature from photoacoustic signals using wavelet-packet features. In Alexander A. Oraevsky and Lihong V. Wang, editors, *SPIE Proceedings*. SPIE, February 2011.
- [208] Aiyong Zhang, Jian Fang, Faming Liang, Vince Calhoun, and Yuping Wang. Aberrant brain connectivity in schizophrenia detected via a fast gaussian graphical model. *IEEE Journal of Biomedical and Health Informatics*, PP:1–1, 07 2018.
- [209] Junhao Zhang, Vishwanatha M. Rao, Ye Tian, Yanting Yang, Nicolas Acosta, Zihan Wan, Pin-Yu Lee, Chloe Zhang, Lawrence S. Kegeles, Scott A. Small, and Jia Guo. Detecting schizophrenia with 3d structural brain mri using deep learning, 2022.
- [210] Lei Zhang. EEG signals classification using machine learning for the identification and diagnosis of schizophrenia. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, July 2019.
- [211] Min Zhao, Weizheng Yan, Na Luo, Dongmei Zhi, Zening Fu, Yuhui Du, Shan Yu, Tianzi Jiang, Vince D. Calhoun, and Jing Sui. An attention-based hybrid deep learning framework integrating brain connectivity and activity of resting-state functional MRI data. *Medical Image Analysis*, 78:102413, May 2022.
- [212] Scott Makeig zyy Ping Jung. Removing artifacts from eeg.

# Anexo

En este anexo se encontrarán tablas con mayor información respecto a la limpieza de cada uno de los datos utilizados. Así como también en mayor detalle resultados que no cumplieron con los valores mínimos de las métricas calculadas.

Sujeto	Total trials	<i>jumps</i>	ruido	cruces	Eliminados	Observaciones
CNTF_002	292	3	5	0	5	-
CNTF_003	291	3	16	0	19	-
CNTF_005	290	7	2	0	9	-
CNTF_007	289	14	3	0	17	-
CNTF_010	142	2	0	0	2	-
CNTF_030	293	6	2	0	8	-
CNTF_031	291	20	4	0	24	-
CNTF_032	292	5	5	0	10	-
CNTF_033	289	12	30	0	42	-
CNTF_035	289	10	31	0	41	-
CNTF_036	298	8	0	0	8	-
CNTF_037	288	5	4	0	9	-
CNTF_039	290	17	0	0	17	-
CNTF_040	300	8	1	0	9	-
CNTF_041	289	7	1	0	8	-
CNTF_044	289	1	0	0	1	-
CNTF_045	288	2	11	6	19	-

Tabla 6.1: Análisis visual para pacientes sanos en prueba STB

Sujeto	Total trials	<i>jumps</i>	ruido	cruces	Eliminados	Observaciones
FEP_001	309	20	0	0	20	-
FEP_004	321	3	0	0	3	-
FEP_005	365	0	50	1	51	-
FEP_006	287	6	0	0	6	-
FEP_007	288	3	0	0	3	-
FEP_008	287	36	0	0	36	-
FEP_009	291	8	0	0	8	-
FEP_010	335	31	0	5	36	-
FEP_011	290	20	0	8	28	-
FEP_012	287	2	0	1	3	-
FEP_014	292	3	1	0	4	-
FEP_016	292	12	0	0	12	-
FEP_017	313	5	1	146	7	-
FEP_018	303	6	12	0	18	-
FEP_022	316	3	0	0	3	-
FEP_023	289	8	0	0	8	-

Tabla 6.2: Análisis visual para pacientes con esquizofrenia en prueba STB

Sujeto	Total trials	<i>jumps</i>	ruido	cruces	Eliminados	Observaciones
CNTF_001	497	0	0	1	1	-
CNTF_004	498	0	20	10	30	-
CNTF_007	502	0	9	0	9	-
CNTF_008	501	0	22	1	23	-
CNTF_009	498	0	17	2	19	-
CNTF_010	353	1	18	0	19	-
CNTF_021	498	0	2	0	2	-
CNTF_031	502	0	0	2	2	-
CNTF_035	500	0	4	0	4	-

Tabla 6.3: Análisis visual para pacientes sanos en prueba DBF

Sujeto	Total trials	<i>jumps</i>	ruido	cruces	Eliminados	Observaciones
EQZ_009_R1	69	0	2	0	2	-
EQZ_009_R2	69	0	3	0	3	-
EQZ_010_R1	69	0	5	7	9	-
EQZ_010_R2	69	0	1	1	1	-
EQZ_013_R2	69	0	2	6	8	-
EQZ_015_R2	69	0	4	4	8	-
EQZ_1058_R1	69	0	1	0	1	-
EQZ_1058_R2	69	0	3	0	3	-
EQZ_1059_R1	69	0	0	5	5	-
EQZ_GDP_R1	69	0	17	17	17	Trials >15%. Dato eliminado
EQZ_GDP_R2	69	-	-	-	-	Dato eliminado
FEP_002_R2	69	0	0	4	4	-
FEP_009_R2	69	0	1	0	1	-
FEP_011_R1	69	-	-	-	-	Dato eliminado
FEP_011_R2	69	-	-	14	14	Trials >15%. Dato eliminado
FEP_012_R2	69	0	0	2	2	-

Tabla 6.4: Análisis visual para pacientes con esquizofrenia en prueba RS

Sujeto	Total trials	<i>jumps</i>	ruido	cruces	Eliminados	Observaciones
EQZ_001	500	2	0	1	3	-
FEP_001	292	0	0	3	3	-
FEP_005	503	18	3	7	27	-
FEP_008	497	0	1	1	1	-
FEP_011	498	0	4	0	4	-
FEP_014	497	1	0	0	1	-
FEP_018	497	2	11	1	13	-

Tabla 6.5: Análisis visual para pacientes con esquizofrenia en prueba DBF

Sujeto	Total trials	<i>jumps</i>	ruido	cruces	Eliminados	Observaciones
CNTF_001_R1	69	0	1	0	1	-
CNTF_002_R1	69	0	0	11	11	Trials >15%. Dato eliminado
CNTF_002_R2	69	-	-	-	-	Dato eliminado
CNTF_007_R2	69	0	4	0	4	-
CNTF_008_R1	69	0	7	0	7	-
CNTF_008_R2	69	0	3	0	3	-
CNTF_009_R1	69	0	2	0	2	-
CNTF_010_R1	69	0	2	0	2	-
CNTF_010_R2	69	0	1	0	1	-
CNTF_012_R1	69	0	1	0	1	-
CNTF_012_R2	69	0	0	3	3	-
CNTF_016_R2	69	0	1	0	1	-
CNTF_019_R2	69	0	1	0	1	-
CNTF_035_R1	69	0	4	0	4	-
CNTF_040_R1	69	0	1	2	3	-
CNTF_041_R1	69	0	1	1	2	-
CNTF_045_R2	69	0	2	0	2	-
CNTF_074_R1	69	0	2	0	2	-

Tabla 6.6: Análisis visual para pacientes sanos en prueba RS

Sujeto	Total rechazados	pestañeo	sacada	Observaciones
CNTF_002	2	2	0	-
CNTF_003	2	1	1	-
CNTF_005	1	1	0	-
CNTF_007	1	1	0	-
CNTF_010	2	2	0	-
CNTF_030	2	1	1	-

Tabla 6.7: Análisis visual para el rechazo de componentes (ICA) en pacientes sanos para la prueba STB Parte 1

Sujeto	Total rechazados	pestañeo	sacada	Observaciones
CNTF_031	1	1	0	-
CNTF_032	2	1	1	-
CNTF_033	1	1	0	-
CNTF_034	2	1	1	-
CNTF_035	1	1	0	-
CNTF_036	2	1	1	-
CNTF_037	3	2	1	-
CNTF_039	1	0	1	-
CNTF_041	1	1	0	-
CNTF_044	1	0	1	-

Tabla 6.8: Análisis visual para el rechazo de componentes (ICA) en pacientes sanos para la prueba STB Parte 2

Sujeto	Total rechazados	pestañeo	sacada	Observaciones
FEP_001	1	1	0	-
FEP_002	2	2	0	-
FEP_004	3	2	1	-
FEP_005	1	1	0	-
FEP_006	1	1	0	-
FEP_007	1	1	0	-
FEP_008	3	3	0	-
FEP_009	1	1	0	-
FEP_010	1	1	0	-
FEP_011	1	1	0	-
FEP_012	1	0	1	-
FEP_016	1	0	1	-
FEP_017	1	0	1	-
FEP_023	1	1	0	-

Tabla 6.9: Análisis visual para el rechazo de componentes (ICA) en pacientes con esquizofrenia para la prueba STB

Sujeto	Total rechazados	pestañeo	sacada	Observaciones
EQZ_001	1	1	0	-
EQZ_004	1	1	0	-
FEP_001	1	1	0	-
FEP_002	1	1	0	-
FEP_005	2	1	1	-
FEP_006	1	1	0	-
FEP_007	1	1	0	-
FEP_008	1	1	0	-
FEP_010	1	1	0	-
FEP_011	1	1	0	-

Tabla 6.10: Análisis visual para el rechazo de componentes (ICA) en pacientes con esquizofrenia para la prueba DBF

Sujeto	Total rechazados	pestañeo	sacada	Observaciones
CNTF_001	3	2	1	-
CNTF_004	1	1	0	-
CNTF_005	1	0	1	-
CNTF_006	2	2	0	-
CNTF_007	1	1	0	-
CNTF_008	1	1	0	-
CNTF_009	3	2	1	-
CNTF_010	1	1	0	-
CNTF_015	1	1	0	-
CNTF_017	1	1	0	-
CNTF_021	1	1	0	-
CNTF_031	1	1	0	-
CNTF_035	1	1	0	-
CNTF_037	1	1	0	-

Tabla 6.11: Análisis visual para el rechazo de componentes (ICA) en pacientes sanos para la prueba DBF

Sujeto	Total rechazados	pestañeo	sacada	Observaciones
CNTF_001_R1	3	2	1	-
CNTF_001_R2	2	1	1	-
CNTF_003_R1	3	1	2	-
CNTF_003_R2	1	1	0	-
CNTF_004_R2	1	0	1	-
CNTF_005_R1	-	-	0	Componentes inválidos, canal malo. Dato eliminado
CNTF_005_R2	-	-	0	Dato eliminado
CNTF_006_R1	2	2	0	-
CNTF_007_R1	1	1	0	-
CNTF_007_R2	3	3	0	-
CNTF_008_R1	3	1	2	-
CNTF_008_R2	1	1	0	-
CNTF_009_R1	1	0	1	-
CNTF_010_R1	3	0	3	-
CNTF_010_R2	1	0	1	-
CNTF_011_R1	1	0	1	-
CNTF_012_R1	1	1	0	-
CNTF_012_R2	1	0	1	-
CNTF_015_R1	1	1	0	-
CNTF_016_R1	1	1	0	-
CNTF_018_R1	2	1	1	-
CNTF_019_R2	1	0	1	-
CNTF_021_R1	1	0	1	-
CNTF_021_R1	1	0	1	-
CNTF_022_R1	-	-	0	Componentes inválidos, canal malo. Dato eliminado
CNTF_022_R2	-	-	0	Dato eliminado

Tabla 6.12: Análisis visual para el rechazo de componentes (ICA) en pacientes sanos para la prueba RS Parte 1

Sujeto	Total rechazados	pestañeo	sacada	Observaciones
CNTF_023_R1	1	1	0	-
CNTF_023_R2	1	0	1	-
CNTF_030_R1	3	0	3	-
CNTF_031_R2	2	0	2	-
CNTF_033_R1	3	1	2	-
CNTF_035_R2	1	1	0	-
CNTF_041_R1	3	0	3	-
CNTF_071_R1	1	1	0	-
CNTF_071_R2	1	1	0	-

Tabla 6.13: Análisis visual para el rechazo de componentes (ICA) en pacientes sanos para la prueba RS Parte 2

Sujeto	Total rechazados	pestañeo	sacada	Observaciones
EQZ_003_R1	1	0	1	-
EQZ_005_R1	-	-	0	Componentes inválidos, canal malo. Dato eliminado
EQZ_005_R2	-	-	0	Dato eliminado
EQZ_013_R1	1	1	1	-
EQZ_013_R2	1	0	1	-
EQZ_015_R1	1	0	1	-
EQZ_1058_R1	2	0	2	-
EQZ_1059_R2	1	0	1	-
FEP_002_R1	3	3	0	-
FEP_005_R1	2	1	1	-
FEP_006_R1	2	2	0	-
FEP_006_R2	1	1	0	-
FEP_008_R1	1	1	0	-
FEP_009_R1	3	1	2	-

Tabla 6.14: Análisis visual para el rechazo de componentes (ICA) en pacientes con esquizofrenia para la prueba RS Parte 1

Sujeto	Total rechazados	pestañeo	sacada	Observaciones
FEP_009_R2	2	0	2	-
FEP_010_R1	2	2	0	-
FEP_015_R1	1	0	1	-
FEP_016_R1	1	0	1	-
FEP_017_R1	1	0	1	-
FEP_023_R1	2	1	1	-

Tabla 6.15: Análisis visual para el rechazo de componentes (ICA) en pacientes con esquizofrenia para la prueba RS Parte 2

Prueba	Configuración	ERPs	Accuracy	F1	Recall	Precision
STB	2 épocas 0.5 dropout doble de neuronas en capa oculta 4 grupos de cv <b>Data concatenada</b>	92 - 101 - 102	75.93 %	69.4 %	61.67 %	88.33 %
STB	10 épocas 0.5 dropout 50 neuronas en capa oculta 6 grupos de cv <b>Data concatenada</b>	92 - 101	75 %	64.74 %	57.41 %	82.41 %
STB	10 épocas 0.5 dropout doble de neuronas en capa oculta 6 grupos de cv <b>Data concatenada</b>	91 - 92 - 101	76.85 %	72.3 %	66.67 %	86.11 %

Tabla 6.16: Resultados con  $acc \geq 75\%$  pero con overfitting (parte 1).

Prueba	Configuración	ERPs	Accuracy	F1	Recall	Precision
STB	10 épocas 0.5 dropout doble de neuronas de 1ra capa en capa oculta 3 grupos de cross-validation <b>Data concatenada</b>	91 - 92 - 101	75.93 %	73.06 %	70.37 %	80.22 %
STB	6 épocas 0.5 dropout mitad neuronas de primera capa en capa oculta 4 grupos de cross-validation	91	75.93 %	71.17 %	63.75 %	86.8 %
STB	6 épocas 0.5 dropout doble de neuronas de primera capa en capa oculta 3 grupos de cross-validation	91 - 92 - 102	75 %	77.24 %	74.21 %	88.81 %

Tabla 6.17: Resultados con  $acc \geq 75\%$  pero con overfitting (parte 2).

Prueba	Configuración	ERPs	Accuracy	F1	Recall	Precision
STB	2 épocas Sin dropout 100 neuronas en capa oculta 5 grupos de cross-validation	91 - 101	75.83 %	72.87 %	68.89 %	81.89 %
STB	2 épocas Sin dropout doble neuronas de primera capa en capa oculta 6 grupos de cross-validation	91	75 %	70.21 %	62.96 %	87.04 %
STB	2 épocas Sin dropout mitad neuronas de primera capa en capa oculta 4 grupos de cross-validation	101 - 102	75 %	67.09 %	64.58 %	83.45 %
STB	2 épocas 0.5 dropout doble de neuronas de 1ra capa en capa oculta 6 grupos de cross-validation <b>Data concate-</b> <b>nada</b>	92 - 101	75 %	73.9 %	72.22 %	79.26 %

Tabla 6.18: Resultados con  $acc \geq 75\%$  pero con underfitting.

Prueba	Configuración	ERPs	Accuracy	F1	Recall	Precision
STB	6 épocas Sin dropout mitad de neuronas de la 1ra capa en capa oculta 3 grupos de cv	92 - 101	76.85 %	71.85 %	64.82 %	88.68 %
STB	10 épocas 0.5 dropout 100 neuronas en capa oculta 6 grupos de cv	101 - 102	75.93 %	73.10 %	68.52 %	83.89 %
STB	10 épocas 0.5 dropout mitad neuronas en capa oculta 3 grupos de cv	92 - 101 - 102	75.93 %	74.54 %	70.83 %	83.3 %

Tabla 6.19: Resultados con  $acc \geq 75\%$  pero con overfitting (parte 3).

amplcorr (amplitude correlation)	coh (coherence)	csd (cross-spectral density matrix)
dtf (directed transfer function)	granger (granger causality)	pdc (patial directed coherence)
plv (phase-locking value)	powcorr (power correlation)	ppc (pairwise phase consistency)
psi (phaseslope index)	wpli (weighted phase lag index)	plm (phrase linearity measurement)
wpli_debiased (debiased weighted phase lag index)	corr (pearson correlation)	laggedcoherence (lagged coherence estimate)
wppc (weighted pairwise phase consistency)	powcorr_ortho (power correlation with single trials orthogonalisation)	

Tabla 6.20: Métodos estadísticos utilizados para calcular la conectividad cerebral