



UNIVERSIDAD DE VALLADOLID

ESCUELA TÉCNICA SUPERIOR

DE INGENIEROS DE TELECOMUNICACIÓN

TRABAJO FIN DE MÁSTER

MÁSTER EN INGENIERÍA DE TELECOMUNICACIÓN

**Análisis de la estructura temporal de la dinámica
asociada a la red neuronal funcional para ayudar
a la caracterización de la enfermedad de
Alzheimer**

Autor:

D. Pablo Carretero Calvo

Tutores:

Dr. D. Carlos Gómez Peña

Dr. D. Jesús Poza Crespo

Valladolid, 2023

TÍTULO: **Análisis de la estructura temporal de la dinámica asociada a la red neuronal funcional para ayudar a la caracterización de la enfermedad de Alzheimer**

AUTOR: **D. Pablo Carretero Calvo**

TUTOR/ES: **Dr. D. Carlos Gómez Peña**
Dr. D. Jesús Poza Crespo

DEPARTAMENTO: **Teoría de la Señal y las Comunicaciones e Ingeniería Telemática**

Tribunal

PRESIDENTE: **Dr. D. Roberto Hornero Sánchez**

SECRETARIA: **Dra. D^a. María García Gadañón**

VOCAL: **Miriam Antón Rodríguez**

P. SUPLENTE: **Dr. D. Ignacio de Miguel Jiménez**

S. SUPLENTE: **Dr. D. Javier Aguiar Pérez**

V. SUPLENTE: **Dr. D. Juan Pablo Casaseca de la Higuera**

FECHA: **2023**

CALIFICACIÓN:

Dedicado a mi familia.

Agradecimientos

Agradezco a mis compañeros de laboratorio, quienes, aunque no todos aparezcan como co-tutores, me han brindado ayuda y enseñanza a lo largo de mis primeros pasos en la investigación y la realización de este trabajo. Agradezco a mis tutores por abrirme las puertas al mundo de la investigación. Especial agradecimiento a mi familia: Rosa, Joaquín, Luis, Marta, Kira; así como a todos mis amigos y a mi pareja, Ainara. Han sabido brindarme su apoyo de diferentes formas en los momentos difíciles y estresantes que he enfrentado a lo largo de mi carrera. Sé que seguirán respaldándome en todo momento, como lo han hecho hasta ahora. Gracias a todos.

Resumen

La enfermedad de Alzheimer (EA) es una de las principales causas de demencia. Los pacientes con EA sufren alteraciones cognitivas, funcionales y conductuales. El principal factor de riesgo es la edad, lo cual hace que esta enfermedad tenga un impacto cada vez más notable debido al aumento en la esperanza de vida. Sin embargo, su diagnóstico sigue siendo un desafío. A pesar de los avances en técnicas diagnósticas respaldadas por evidencia biológica, su disponibilidad generalizada está limitada por su alto coste. En el contexto del rápido desarrollo tecnológico reciente, el análisis de señales de electroencefalografía (EEG), impulsado por diversos algoritmos de computación, está ganando importancia al proponer enfoques innovadores para caracterizar la enfermedad y ayudar en su diagnóstico. En esta línea, este estudio se enfoca en la aplicación de métodos cronectómicos, que cuantifican las propiedades dinámicas de la conectividad funcional cerebral.

El análisis se llevó a cabo utilizando dos bases de datos. La primera de ellas incluyó a 160 participantes, entre los cuales 67 eran pacientes con EA, 50 tenían deterioro cognitivo leve (DCL) y 43 eran individuos sanos desde el punto de vista cognitivo. La segunda base de datos constó de 126 sujetos, de los cuales 50 eran pacientes con EA, 38 padecían DCL y 38 eran individuos de control. El proceso comenzó con la reconstrucción de la actividad neuronal a nivel de fuente utilizando el algoritmo sLORETA (*Standardized Low-Resolution Brain Electromagnetic Tomography*). Luego, se identificaron meta-estados cerebrales mediante la aplicación del algoritmo IAC (*Instantaneous Amplitude Correlation*) en las bandas de frecuencia delta, zeta, alfa y beta-1 para obtener tensores de conectividad funcional instantánea en diversos momentos temporales. A partir de estas matrices de conectividad, se generaron diagramas de recurrencia que mostraban la presencia de meta-estados recurrentes en el tiempo. Utilizando el algoritmo de detección de comunidades Louvain-GJA, se obtuvieron tres meta-estados, que posteriormente se ordenaron según su similitud topológica. Los datos se normalizaron utilizando señales subrogadas para descartar que los resultados surgieran de oscilaciones de origen no neuronal. El proceso de subrogación se optimizó utilizando archivos en C/C++ y procesamiento de datos en la unidad de procesamiento gráfico (GPU). Mediante este enfoque, se generaron tensores de correlación instantánea (ICT) y la secuencia de activación temporal (TAS) que cuantifican la semejanza de las topologías observadas con los meta-estados a lo largo del tiempo. Para resumir las propiedades dinámicas de la red neuronal funcional, se calcularon diversas medidas cronectómicas: singularidad del meta-estado dominante, área bajo el ICT, diferencia de permanencia, la probabilidad de ocurrencia del meta-estado nulo, percentil 5°, grado de antagonismo, *dwell time*, *leap size* y complejidad de TAS.

El análisis detallado de los tensores de correlación instantánea reveló diversos cambios asociados con las alteraciones cerebrales que sufren los pacientes con DCL y EA, principalmente en las bandas zeta y alfa. Estas perturbaciones implican cambios en la comunicación entre regiones cerebrales, aparición de oscilaciones más aleatorias y deterioro de la estabilidad y estructura temporal. Además, se detectaron cambios en la persistencia temporal de los meta-estados, que estaban vinculados a la pérdida de estructura y al aumento de la aleatoriedad. Los resultados de las medi-

das implementadas se utilizaron para clasificar los tres grupos presentes en la base de datos y para diferenciar entre individuos de control y pacientes (combinando pacientes con DCL y EA). Se emplearon algoritmos de selección de características como el filtro rápido basado en la correlación (FCBF), selección secuencial de características (SFS) y eliminación recursiva de características (RFE). Para elegir el modelo de clasificación, se utilizó una validación cruzada con 10 particiones de la base de datos. Los modelos de clasificación que mejor rendimiento obtuvieron fueron los árboles de decisión y las redes neuronales. Los resultados de clasificación entre los tres alcanzaron valores de kappa de 0.81 para cada base de datos por separado y 0.63 para la combinación de las bases de datos. En el caso de la clasificación entre controles y la agrupación de pacientes con DCL y EA, se obtuvo valores kappa de 1 para la primera base de datos, de 0.7 para la segunda y 0.68 para la combinación de ambas. En su conjunto, los resultados sugieren que las medidas conectómicas podrían ser herramientas valiosas para ayudar en el diagnóstico y la caracterización de la EA.

Palabras clave

Conectividad funcional dinámica, conectómica, deterioro cognitivo leve, electroencefalografía, enfermedad de Alzheimer, meta-estados.

Abstract

Alzheimer's disease (AD) is one of the main causes of dementia. This condition leads to cognitive, functional, and behavioral changes. The main risk factor is aging, which makes this disease have an increasingly notable impact due to the increase in life expectancy. However, its diagnosis remains a challenge. Despite advances in diagnostic techniques supported by biological evidence, their widespread availability is limited by their high cost. In the context of recent rapid technological development, the analysis of electroencephalography (EEG) signals, driven by various computational algorithms, is gaining importance by proposing innovative approaches to characterize AD. In this line, this study focuses on the application of chronotomic methods, which quantify the dynamic properties of functional connectivity in neuronal associations.

The analysis was carried out using two databases. The first one included 160 participants, divided into 67 AD patients, 50 mild cognitive impairment patients, and 43 cognitively healthy individuals. The second database consisted of 126 subjects, among whom 50 were AD patients, 38 had mild cognitive impairment, and 38 were cognitively healthy controls. Initially, the neural activity was reconstructed at source level using the sLORETA (Standardized Low-Resolution Brain Electromagnetic Tomography) algorithm. Then, brain meta-states were identified thanks to the application of the IAC (Instantaneous Amplitude Correlation) algorithm in the frequency bands delta, theta, alpha, and beta-1 to obtain instantaneous functional connectivity tensors at various time points. From these connectivity matrices, recurrence plot diagrams showing the presence of recurrent meta-states in time were generated. Using the Louvain-GJA community detection algorithm, three meta-states were obtained; they were subsequently ordered according to their topological similarity. The data were normalized using surrogated signals to rule out that the results arose from random oscillations. The surrogate process was optimized using C/C++ files and GPU data processing. Using this approach, the instantaneous correlation tensors (ICT) and the temporal activation sequence (TAS) were generated indicating the similarity of the observed topologies to the meta-states over time. To summarize the dynamic properties of the functional neural network, multiple chronotomic measures were calculated: uniqueness of the dominant meta-state, area under the ICT, difference in permanence, probability of occurrence of the null meta-state, 5th percentile, degree of antagonism, dwell time, leap size, and TAS complexity.

Detailed analysis of the instantaneous correlation tensors revealed several alterations related to MCI and AD, mainly in the theta and alpha bands. These perturbations involve changes in communication between brain regions, appearance of more random oscillations, and impaired temporal stability and structure. In addition, changes in the temporal persistence of meta-states were detected, which were linked to a loss of structure and an increased randomness. The results of the implemented measures were used to classify the three groups present in the database and to differentiate between controls and patients (combining MCI and AD patients). Feature selection algorithms such as fast correlation-based filter (FCBF), sequential forward selection (SFS), and recursive feature elimination (RFE) were used. To choose the classification model, a cross-validation with 10 partitions of the database was used. The best performing classification models were deci-

sion trees and neural networks. The classification results among the three groups obtained kappa values of 0.81 for each database separately and 0.63 for the combination of databases. In the case of classification between controls and the grouping of patients with MCI and AD, kappa values of 1 were obtained for the first database, 0.7 for the second, and 0.68 for the combination of both databases. Taken together, the results suggest that chronotomic measures could be valuable tools to support the diagnosis and characterization of AD.

Keywords

Dynamic functional connectivity, chronotomics, mild cognitive impairment, electroencephalography, Alzheimer's disease, classification, meta-states.

Índice general

1. Introducción	1
1.1. Enfermedad de Alzheimer	2
1.2. Actividad cerebral y conectómica	2
1.3. Motivación	2
1.4. Hipótesis	3
1.5. Objetivos	3
1.6. Estructura del documento	4
2. La enfermedad de Alzheimer	5
2.1. Introducción	6
2.2. Demencia por enfermedad de Alzheimer	6
2.3. Evolución de la enfermedad de Alzheimer	6
2.4. Diagnóstico	7
3. Electroencefalografía	9
3.1. Introducción	10
3.2. Generación de señales	10
3.3. Oscilaciones neuronales y su significado	13
3.4. Alteraciones del EEG asociadas a la enfermedad de Alzheimer	13
4. Materiales y métodos	15
4.1. Introducción	16
4.2. Conjunto de datos	16
4.2.1. Registros EEG	16
4.2.2. Localización de fuentes	17
4.3. Detección de meta-estados	17
4.3.1. Correlación de amplitud instantánea (IAC)	17
4.3.2. Diagramas de recurrencias y ventanas adaptativas	18
4.3.3. Algoritmo de detección de comunidades para obtención de meta-estados	18
4.4. Medidas conectómicas	19
4.4.1. Singularidad de meta-estado dominante	19
4.4.2. Área bajo el ICT	20
4.4.3. Diferencia de permanencia	20
4.4.4. Probabilidad de ocurrencia del meta-estado nulo	20
4.4.5. Percentil 5	20
4.4.6. Grado de antagonismo	20
4.4.7. <i>Dwell time</i>	21
4.4.8. <i>Leap size</i>	21

4.4.9. Complejidad de TAS	21
4.5. Normalización	22
4.6. Optimización del algoritmo	22
4.7. Métodos estadísticos	23
4.8. Clasificación	24
4.8.1. Selección de características	24
4.8.2. Modelos de clasificación	25
4.8.3. Métricas de clasificación	26
5. Resultados	28
5.1. Introducción	29
5.2. Optimización	29
5.3. Singularidad de meta-estado dominante	30
5.4. Área bajo el ICT	31
5.5. Diferencia de permanencia	32
5.6. Probabilidad de ocurrencia del meta-estado nulo	33
5.7. Percentil 5°	34
5.8. Grado de antagonismo	34
5.9. <i>Dwell time</i>	35
5.10. <i>Leap size</i>	36
5.11. Complejidad de TAS	37
5.12. Clasificación	38
5.12.1. K-mejores	39
5.12.2. FCBF	40
5.12.2.1. Clasificación con 3 clases	41
5.12.2.2. Clasificación con 2 clases	43
5.12.3. SFS	44
5.12.3.1. Clasificación con 3 clases	44
5.12.3.2. Clasificación con 2 clases	45
5.12.4. RFE	47
5.12.4.1. Clasificación con 3 clases	49
5.12.4.2. Clasificación con 2 clases	50
6. Discusión	53
6.1. Introducción	54
6.2. Optimización	54
6.3. Medidas cronectómicas	54
6.4. Clasificación	57
6.5. Limitaciones y líneas futuras	59
7. Conclusiones	62
7.1. Introducción	63
7.2. Consecución de los objetivos del trabajo	63
7.3. Conclusiones	63
A. Glosario de Siglas y Acrónimos	65
Bibliografía	66

Índice de figuras

2.1. Progresión de la EA. Figura adaptada de (Alzheimer’s Association, 2023)	7
3.1. Esquema resumen de las partes de una neurona (Noback et al., 1993).	10
3.2. Esquema ilustrativo de los campos magnéticos y eléctricos producidos en la neurona. Imagen adaptada de (Lopes da Silva, 2013). Las flechas rojas ilustran los campos magnéticos. Las flechas negras ilustran los campos eléctricos. Las gráficas a la izquierda ilustran la polaridad asociada a los campos en la zona de la neurona donde se encuentran. EPSP: potencial postsináptico excitatorio.	11
3.3. Esquema ilustrativo de un sistema 10-5. Los electrodos etiquetados se corresponden a un sistema 10-20. “P” ilustra los puntos preauriculares. “A1” y “A2” representan los lóbulos de la oreja (Müller-Putz, 2020).	12
4.1. Meta-estados detectados por el método de detección de comunidades Louvain-GJA; A: meta-estados en la banda delta; B: meta-estados en la banda zeta; C: meta-estados en la banda alfa; D: meta-estados en la banda beta-1. Color Rojo: Meta-estados obtenidos para la BBDD del Hospital Universitario Río Hortega; Color verde: Meta-estados obtenidos para la BBDD del proyecto POCTEP.	18
4.2. Secuencia de activación temporal (diagrama superior) y tensor de correlación instantánea (diagrama inferior) para un paciente de muestra con EA durante los primeros 10 segundos de la grabación en la banda alfa (Núñez et al., 2021).	19
4.3. Esquema ilustrativo de la medida “probabilidad de ocurrencia del meta-estado nulo”. El círculo resalta la región en la que los tres meta-estados presentan valores de correlación cercanos a cero durante varias muestras (Carretero, 2022).	21
4.4. Porcentaje de tiempo invertido en cada etapa de la obtención de meta-estados.	23
5.1. Resultado de la medida “singularidad de meta-estado dominante”. A: Base de datos del Hospital Universitario Río Hortega B: Base de datos del proyecto POCTEP. El cuadro azul indica diferencias estadísticamente significativas globales ($p < 0,05$, Kruskal-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).	30
5.2. Resultado de la medida “área bajo el ICT” en el meta-estado dominante. A: Base de datos del Hospital Universitario Río Hortega B: Base de datos del proyecto POCTEP. Los cuadros azules indican diferencias globales estadísticamente significativas ($p < 0,05$, Kruskal-Wallis, corrección FDR). Las llaves rojas indican diferencias entre pares de grupos estadísticamente significativas ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).	31

5.3. Resultado de la medida “diferencia de permanencia”. A: Base de datos del Hospital Universitario Río Hortega B: Base de datos del proyecto POCTEP. El cuadro azul indica diferencias globales estadísticamente significativas ($p < 0,05$, Kruskall-Wallis, corrección FDR). Las llaves rojas indican diferencias entre pares de grupos estadísticamente significativas ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).	32
5.4. Resultado de la medida “probabilidad de ocurrencia del meta-estado nulo”. A: Base de datos del Hospital Universitario Río Hortega B: Base de datos del proyecto POCTEP. Los cuadros azules indican resultados globales significativos ($p < 0,05$, Kruskall-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).	33
5.5. Resultado de la medida “percentil 5º”. Los cuadros azules indican resultados globales significativos ($p < 0,05$, Kruskall-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).	35
5.6. Resultado de la medida “grado de antagonismo”. Los cuadros azules indican resultados globales significativos ($p < 0,05$, Kruskall-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).	36
5.7. Resultado de la medida <i>dwell time</i> . Los cuadros azules indican resultados globales significativos ($p < 0,05$, Kruskall-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).	37
5.8. Resultado de la medida <i>leap size</i> . Los cuadros azules indican resultados globales significativos ($p < 0,05$, Kruskall-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).	38
5.9. Resultado de la medida “complejidad de TAS”. Los cuadros azules indican resultados globales significativos ($p < 0,05$, Kruskall-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).	39
5.10. Importancia asignada a cada parámetro por el algoritmo de K-mejores en las BBDD del Hospital Universitario Río Hortega (A) y POCTEP (B).	40
5.11. Importancia asignada a cada parámetro por el algoritmo de K-mejores en ambas BBDD.	40
5.12. Correlación asignada a cada parámetro por el algoritmo FCBF en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B).	41
5.13. Correlación asignada a cada parámetro por el algoritmo FCBF en ambas BBDD.	41
5.14. Precisión obtenida en la clasificación con tres clases (control, DCL y EA) en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B) con características seleccionadas por FCBF. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.	42
5.15. Precisión obtenida en la clasificación con tres clases (control, DCL y EA) en las BBDD de forma conjunta con características seleccionadas por FCBF. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.	42

5.16. Precisión obtenida en la clasificación con dos clases (control y demencia) en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B) con características seleccionadas por FCBF. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.	44
5.17. Precisión obtenida en la clasificación con dos clases (control y Demencia) en las BBDD de forma conjunta con características seleccionadas por FCBF. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.	44
5.18. Características seleccionadas por el algoritmo SFS en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B).	45
5.19. Características seleccionadas por el algoritmo SFS en ambas BBDD.	45
5.20. Precisión obtenida en la clasificación con tres clases (control, DCL y EA) en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B) con características seleccionadas por SFS. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.	46
5.21. Precisión obtenida en la clasificación con tres clases (control, DCL y EA) en las BBDD de forma conjunta con características seleccionadas por SFS. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.	47
5.22. Precisión obtenida en la clasificación con dos clases (control y demencia) en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B) con características seleccionadas por SFS. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.	48
5.23. Precisión obtenida en la clasificación con dos clases (control y demencia) en las BBDD de forma conjunta con características seleccionadas por SFS. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.	48
5.24. Características seleccionadas por el algoritmo RFE en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B).	48
5.25. Características seleccionadas por el algoritmo RFE en ambas BBDD.	49
5.26. Precisión obtenida en la clasificación con tres clases (control, DCL y EA) en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B) con características seleccionadas por RFE. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.	49
5.27. Precisión obtenida en la clasificación con tres clases (control, DCL y EA) en las BBDD de forma conjunta con características seleccionadas por RFE. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.	50
5.28. Precisión obtenida en la clasificación con dos clases (control y demencia) en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B) con características seleccionadas por RFE. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.	51
5.29. Precisión obtenida en la clasificación con dos clases (control y demencia) en las BBDD de forma conjunta con características seleccionadas por RFE. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.	52

Índice de tablas

2.1. Criterios de diagnóstico de la EA basado en los principales biomarcadores. A indica la presencia de $A\beta$. T indica la presencia de tau fosforilada. N indica daño neurodegenerativo y/o neuronal. El símbolo + o un - indica la presencia o no de los biomarcadores anteriores. Tabla adaptada de (Jack et al., 2018).	8
4.1. H: hombre; M: mujer. A: educación primaria o inferior; B: estudios superiores; m: mediana; IQR: rango intercuartil (<i>interquartile range</i>).	17
4.2. Estructura de la matriz de confusión. VP: verdadero positivo; FP: falso positivo; VN: verdadero negativo; FN: falso negativo	27
4.3. Estructura de la matriz de confusión englobando DCL y EA bajo el grupo “Pacientes” . VP: verdadero positivo; FP: falso positivo; VN: verdadero negativo; FN: falso negativo	27
5.1. Reducción del tiempo de ejecución conseguida en cada etapa del proceso de subrogación.	29
5.2. Tiempo empleado en el proceso de subrogación antes y después de aplicar los métodos de optimización.	29
5.3. Matriz de confusión utilizando el método de selección de características FCBF en la BBDD del Hospital Universitario Río Hortega.	42
5.4. Matriz de confusión utilizando el método de selección de características FCBF en la BBDD del proyecto POCTEP.	42
5.5. Matriz de confusión utilizando el método de selección de características FCBF en la combinación de ambas BBDD.	43
5.6. Matriz de confusión utilizando el método de selección de características FCBF en la BBDD del Hospital Universitario Río Hortega.	43
5.7. Matriz de confusión utilizando el método de selección de características FCBF en la BBDD del proyecto POCTEP.	43
5.8. Matriz de confusión utilizando el método de selección de características FCBF en la combinación de ambas BBDD.	43
5.9. Matriz de confusión utilizando el método de selección de características SFS en la BBDD del Hospital Universitario Río Hortega.	46
5.10. Matriz de confusión utilizando el método de selección de características SFS en la BBDD del proyecto POCTEP.	46
5.11. Matriz de confusión utilizando el método de selección de características SFS en la combinación de ambas BBDD.	46
5.12. Matriz de confusión utilizando el método de selección de características SFS en la BBDD del Hospital Universitario Río Hortega.	47
5.13. Matriz de confusión utilizando el método de selección de características SFS en la BBDD del proyecto del proyecto POCTEP.	47

5.14. Matriz de confusión utilizando el método de selección de características SFS en la combinación de ambas BBDD.	47
5.15. Matriz de confusión utilizando el método de selección de características RFE en la BBDD del Hospital Universitario Río Hortega.	50
5.16. Matriz de confusión utilizando el método de selección de características RFE en la BBDD del proyecto POCTEP.	50
5.17. Matriz de confusión utilizando el método de selección de características RFE en la combinación de ambas BBDD.	50
5.18. Matriz de confusión utilizando el método de selección de características SFS en la BBDD del Hospital Universitario Río Hortega.	51
5.19. Matriz de confusión utilizando el método de selección de características SFS en la BBDD del proyecto del proyecto POCTEP.	51
5.20. Matriz de confusión utilizando el método de selección de características RFE en la combinación de ambas BBDD.	51
5.21. Valor <i>kappa</i> obtenido en la clasificación para diferentes métodos de selección de características y diferentes combinaciones de las BBDD. AD: Árbol de Decisión; RF: <i>Random Forest</i> ; RN: Red Neuronal; RLL2: Regresión Logística con penalización <i>L2</i>	52
6.1. Resultados obtenidos en otros estudios de EEG para identificar controles frente a pacientes con DCL y EA. p: precisión; e: especificidad; s: sensibilidad; k: valor <i>kappa</i> ; AUC: área bajo la curva ROC; C: controles; DCL: deterioro cognitivo leve; EA: enfermedad de Alzheimer.	60

Capítulo 1

Introducción

1.1. Enfermedad de Alzheimer	2
1.2. Actividad cerebral y conectómica	2
1.3. Motivación	2
1.4. Hipótesis	3
1.5. Objetivos	3
1.6. Estructura del documento	4

1.1. Enfermedad de Alzheimer

La enfermedad de Alzheimer (EA) representa la forma más frecuente de demencia en nuestros días. Sus efectos incluyen la pérdida de memoria y otras funciones cognitivas, impactando directamente en la vida diaria de los afectados (Alzheimer's Association, 2023). Durante las primeras etapas de la enfermedad, los pacientes aún pueden mantener conversaciones y desenvolverse de forma normal. Sin embargo, a medida que progresa la enfermedad, los síntomas se intensifican dificultando cada vez más la vida diaria (Alzheimer's Association, 2023). La enfermedad se divide en tres etapas: (i) EA preclínica, sin síntomas de la enfermedad ; (ii) deterioro cognitivo leve (DCL), fase que se distingue del envejecimiento normal pero aún no es considerada demencia por EA; y (iii) demencia provocada por EA, la etapa más avanzada (Alzheimer's Association, 2023). Con el aumento de la esperanza de vida, esta enfermedad está cada vez más presente en la sociedad (Miraglia et al., 2017).

1.2. Actividad cerebral y cronectómica

La electroencefalografía (EEG) y la magnetoencefalografía (MEG) son dos técnicas empleadas para comprender el funcionamiento de la actividad neuronal en el cerebro humano y para estudiar las alteraciones que diversas patologías y condiciones anómalas provocan en la dinámica de los procesos cognitivos (Cohen, 2017). Una de las técnicas más frecuentemente utilizadas para la extracción de información a partir de registros cerebrales es el conectoma, que posibilita la construcción de un “mapa” de las conexiones existentes en el cerebro (Sporns et al., 2005) y aporta información útil para entender cómo las distintas patologías impactan en su comportamiento (Calhoun et al., 2014). El concepto de cronectoma lleva esta idea un paso más allá, focalizándose en las estructuras y procesos que evolucionan a lo largo del tiempo (Calhoun et al., 2014). En este sentido, posibilita un estudio del cerebro como un sistema interconectado y dinámico. La generación del cronectoma demanda técnicas de adquisición de actividad cerebral que sean sensibles a las fluctuaciones dinámicas. Esto es particularmente aplicable a EEG y MEG, que prácticamente reflejan en tiempo real las variaciones relacionadas con la actividad cerebral (Tewarie et al., 2019).

La información proveniente del cronectoma posibilita el desarrollo del concepto de “meta-estados”. Esta noción se fundamenta en la premisa de que existen patrones de conectividad que se repiten. Estos meta-estados pueden manifestarse de manera simultánea, aunque con distintos niveles de influencia (Calhoun et al., 2014), lo que facilita su asociación con la noción de atractor en sistemas no lineales (Núñez et al., 2021). En este Trabajo de Fin de Máster (TFM) se implementan y emplean medidas cronectómicas que evalúan las fluctuaciones entre meta-estados, con el propósito de describir la dinámica de la demencia atribuible a la EA y cuya capacidad se mide mediante la aplicación de diversos modelos de clasificación.

1.3. Motivación

Actualmente, la EA no tiene cura, pero se dispone de terapias que retrasan su progresión y diversas investigaciones están enfocadas en su análisis y detección (Alzheimer's Association, 2023). Gracias a los avances tecnológicos significativos de los últimos años, se han concebido investigaciones basadas en señales de EEG y MEG que permiten una exploración detallada de la conectividad funcional neuronal presente en el cerebro humano (Calhoun et al., 2014; Finn et al., 2015; Hutchison et al., 2013). El presente estudio se enfoca en analizar las características dinámicas de

las conexiones neuronales y trata de establecer vínculos entre los resultados obtenidos y el desarrollo de la EA. Tal como se explicó en la sección 1.2, la disciplina que investiga estas propiedades es conocida como cronectómica, que aplica nuevas metodologías de teoría de redes dinámicas para explorar la conectividad neuronal de manera similar a la conectómica, con la diferencia de que es sensible al comportamiento dinámico de las asociaciones neuronales observadas (Calhoun et al., 2014).

1.4. Hipótesis

Las personas con DCL y EA manifiestan cambios en la actividad cerebral que se traducen en alteraciones de los patrones de actividad neuronal en estado de reposo. Por tanto, el análisis de los registros EEG proporciona una manera de describir el comportamiento del cerebro y las modificaciones que experimenta debido a dichas patologías. La presente investigación plantea como hipótesis que las variaciones dinámicas en la actividad EEG son indicativas de las alteraciones en la red funcional cerebral relacionadas con la progresión de la EA. Por tanto, la aplicación de métodos cronectómicos podría ayudar a comprender las alteraciones que el DCL y la EA provocan progresivamente en la actividad cerebral.

1.5. Objetivos

El trabajo presentado tiene como objetivo principal realizar un análisis de las modificaciones en la estructura temporal de la actividad neuronal en estado de reposo asociadas con la progresión de la EA. Para lograr esto, se ha realizado un análisis cronectómico, el cual facilita la exploración de la dinámica de la red neuronal funcional mediante registros EEG en reposo. Además se han empleado diversos modelos de clasificación, para estudiar la capacidad de discriminación entre grupos que se consigue con la aplicación de las medidas. La implementación de estas metodologías implica el cumplimiento de los siguientes objetivos específicos:

- i. Realizar una etapa de documentación con el propósito de comprender las particularidades de los datos que se van a analizar, junto con las medidas cronectómicas previamente desarrolladas.
- ii. Familiarizarse con la metodología aplicada para comprender el procesado aplicado a las señales de EEG y el proceso de subrogación necesario para obtener resultados sin sesgo.
- iii. Llevar a cabo una fase de documentación sobre métodos de optimización para reducir el tiempo de ejecución necesario en el procesamiento de señal para su posterior implementación.
- iv. Adquirir conocimiento en el uso de la herramienta MATLAB, que se emplea para el procesamiento de señales y las subsiguientes comparaciones de los resultados obtenidos.
- v. Familiarizarse con la herramienta Python, que se emplea para el procesado de los resultados de aplicar las medidas para la extracción de diversos modelos de clasificación.
- vi. Realizar análisis estadísticos sobre los resultados obtenidos entre los diversos grupos de individuos para identificar disparidades a nivel poblacional.
- vii. Implementar métodos de selección de características y modelos de clasificación para identificar las características que permiten distinguir mejor entre controles, pacientes con DCL y pacientes con EA, así como para obtener el modelo con mejor rendimiento.

- viii. Examinar y contrastar los resultados obtenidos con los hallazgos documentados en investigaciones y publicaciones previas.
- ix. Extraer conclusiones a partir del estudio realizado, considerar las limitaciones presentes y explorar posibles direcciones futuras de investigación.

1.6. Estructura del documento

A continuación se presenta la estructura del presente trabajo. El capítulo actual corresponde a la introducción, mientras que el resto del documento se organiza de la siguiente manera:

- **Capítulo 2. La enfermedad de Alzheimer.** Este capítulo introduce la EA, abordando sus síntomas, su impacto en la vida cotidiana de los pacientes y el progreso de la enfermedad. También se introducen los métodos de diagnóstico vigentes.
- **Capítulo 3. Electroencefalografía.** En esta sección se detalla el tipo de señales registradas, el proceso de adquisición y su naturaleza. Además, se discuten los efectos de la enfermedad en las señales de EEG.
- **Capítulo 4. Materiales y métodos.** En este capítulo se exponen las características de los datos empleados en el estudio, se describen los algoritmos utilizados para la adquisición y preprocesamiento de las señales EEG, se detalla el método de extracción de meta-estados, se presentan las medidas aplicadas a las señales EEG, los métodos de optimización utilizados, y los modelos de clasificación implementados.
- **Capítulo 5. Resultados.** En este capítulo se presentan los resultados de aplicar los métodos de optimización, así como los resultados obtenidos mediante las medidas introducidas en el capítulo anterior. Se incluyen los resultados de los análisis estadísticos correspondientes y de los métodos de clasificación empleados.
- **Capítulo 6. Discusión.** Esta sección ofrece un análisis en profundidad de los resultados, respaldando posibles conclusiones a partir de comparaciones con estudios similares.
- **Capítulo 7. Conclusiones.** Finalmente, se presentan las principales conclusiones derivadas del desarrollo del trabajo. También se abordan líneas interesantes para futuras investigaciones y mejoras basadas en el trabajo realizado.

Capítulo 2

La enfermedad de Alzheimer

2.1. Introducción	6
2.2. Demencia por enfermedad de Alzheimer	6
2.3. Evolución de la enfermedad de Alzheimer	6
2.4. Diagnóstico	7

2.1. Introducción

La EA constituye en la actualidad el tipo más frecuente de demencia (Alzheimer's Association, 2023). El riesgo de desarrollar esta enfermedad está condicionado por diversos factores, siendo la edad el más relevante entre ellos (García Cobo, 2022). Con el incremento de la esperanza de vida en la sociedad actual, se observa un aumento progresivo de casos de esta patología (Alzheimer's Association, 2023). A lo largo del tiempo, la EA se agrava, los síntomas empeoran progresivamente y complican la vida cotidiana de los pacientes (García Cobo, 2022).

En este capítulo, se realiza una descripción general de la EA. Se describen sus síntomas, el impacto en la vida diaria, la evolución de la enfermedad y los métodos empleados hasta la fecha para su detección. En capítulos posteriores, se enfatiza en el creciente interés por investigar nuevas metodologías destinadas a la caracterización y diagnóstico de la EA. Entre estos enfoques innovadores se destaca la cronectómica.

2.2. Demencia por enfermedad de Alzheimer

La demencia se caracteriza por una pérdida o debilitamiento progresivo de las facultades mentales debido a la edad o una enfermedad, que resulta en alteraciones de la memoria, la razón y la conducta (Real Academia Española, 2023). La EA es una de las principales causas de demencia y produce cambios significativos en la conducta, la funcionalidad y las capacidades cognitivas de los pacientes. Estos cambios se desarrollan de manera progresiva, comenzando con la degeneración neuronal hasta décadas antes de que la enfermedad pueda ser detectada clínicamente (Alzheimer's Association, 2023).

La EA se caracteriza por tres etapas distintas: la etapa preclínica, sin síntomas de la enfermedad; la etapa de DCL, que precede a la demencia debida a EA; y finalmente, la demencia causada por EA (Alzheimer's Association, 2023). La diferencia entre estas etapas se entiende desde un punto de vista clínico, ya que el envejecimiento normal puede presentar cierto déficit cognitivo, pero el deterioro en casos de DCL es mayor que en personas cognitivamente sanas y no llega a ser comparable al de la demencia por EA. En esta etapa intermedia, se produce un deterioro de las neuronas en el hipocampo y el lóbulo temporal, áreas asociadas al almacenamiento y procesamiento de información (Cancino & Rehbein, 2016).

Con el aumento de la esperanza de vida, la demencia debida a EA se ha convertido en un importante problema de salud pública (Alzheimer's Association, 2023). La prevención de la demencia se ha convertido en una preocupación actual, y varios factores de riesgo han sido identificados, incluyendo un menor nivel educativo, hipertensión, tabaquismo, discapacidad auditiva, obesidad, depresión, inactividad física, diabetes, poco contacto social, consumo excesivo de alcohol, daños cerebrales por fuerzas físicas externas y la polución del aire (Livingston et al., 2020). No obstante, el principal factor de riesgo es la edad (Livingston et al., 2020). Se requiere una inversión adecuada en educación básica, servicios sociales, reducción de la polución y otras medidas para evitar un aumento significativo de los casos de demencia debida a la EA (Livingston et al., 2020).

2.3. Evolución de la enfermedad de Alzheimer

Como se mencionó previamente, la EA sigue un proceso continuo donde los cambios ocurren de manera progresiva (Poza, 2008). La figura 2.1 muestra el grado de los síntomas relacionados

con el avance de la enfermedad. Para analizar el desarrollo de la patología, se pueden identificar tres etapas principales (Alzheimer’s Association, 2023):

- **EA preclínico.** Es la fase inicial en la que aparecen las primeras alteraciones cerebrales. Por ejemplo las anomalías en la concentración de β -amiloide ($A\beta$), que pueden indicar un inicio de la enfermedad. No hay cambios en las capacidades cognitivas, ni se desarrolla DCL o demencia en esta etapa.
- **Deterioro cognitivo leve.** Como se mencionó anteriormente, esta etapa es lo suficientemente significativa como para distinguirla del envejecimiento normal, pero aún no se considera demencia. Los pacientes con DCL muestran niveles de $A\beta$ fuera del rango de control y se puede detectar mediante el uso de biomarcadores. Aunque el DCL es claramente distinguible, su progresión puede llevar a la demencia por EA o mantenerse en un estado de deterioro estable.
- **Demencia provocada por EA.** Esta es la etapa más avanzada de la enfermedad en la que la demencia afecta significativamente la vida diaria de los pacientes. Los síntomas en esta etapa se dividen en tres grados: leve, moderado y severo. A medida que la enfermedad avanza, los pacientes requieren cada vez más apoyo profesional para llevar a cabo actividades cotidianas (Alzheimer’s Association, 2023).

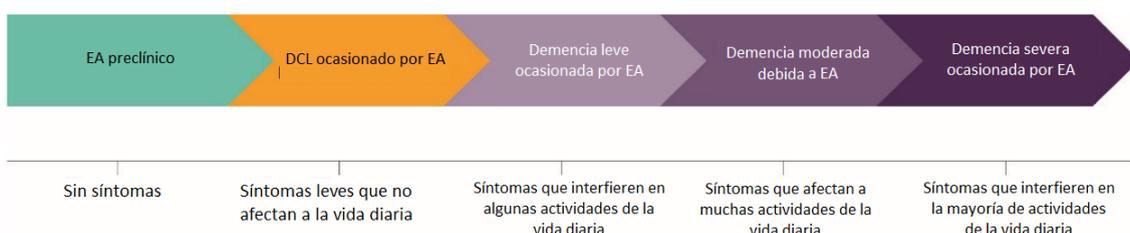


Figura 2.1. Progresión de la EA. Figura adaptada de (Alzheimer’s Association, 2023)

2.4. Diagnóstico

Inicialmente, el diagnóstico definitivo de la EA solo era posible mediante el examen histológico de muestras cerebrales y la aplicación de métodos tradicionales de diagnóstico. Estos métodos implican el análisis del historial médico en relación a cambios cognitivos, antecedentes psiquiátricos, la retroalimentación de familiares acerca del comportamiento del paciente y la realización de pruebas de memoria y resolución de problemas Alzheimer’s Association (2023). Sin embargo, la investigación se ha centrado en identificar rasgos biológicos en lugar de simplemente considerar los síntomas, lo que ha permitido un estudio más preciso de la aparición y progresión de la enfermedad (Jack et al., 2018). Como resultado, se han identificado varios biomarcadores que facilitan la detección de la enfermedad, incluyendo la deposición de proteína $A\beta$, la presencia de la proteína “tau” patológica y la neurodegeneración (Jack et al., 2018). La combinación de estos biomarcadores proporciona información valiosa sobre la patología y la etapa en la que se encuentra el paciente; ejemplos específicos se encuentran detallados en la tabla 2.1 (Jack et al., 2018).

Estado cognitivo				
Perfil de biomarcadores		Control	DCL	Demencia
	$A^{-}T^{-}(N)^{-}$	Biomarcadores en niveles normales para paciente de control	Biomarcadores en niveles normales para paciente con DCL	Biomarcadores en niveles normales para paciente con demencia
	$A^{+}T^{-}(N)^{-}$	Cambios asociados a EA preclínica	Paciente con DCL desarrollando cambios patológicos asociados a la EA	Paciente con demencia desarrollando cambios patológicos asociados a la EA
	$A^{+}T^{-}(N)^{+}$	Disparidad cognitiva asociadas a EA, a la par que factores ajenos a la enfermedad	Paciente con DCL con disparidad cognitiva asociadas a EA, a la par que factores ajenos a la enfermedad	Paciente con demencia con disparidad cognitiva asociadas a EA, a la par que factores ajenos a la enfermedad
	$A^{+}T^{+}(N)^{-}$	EA preclínica	Paciente con DCL previo a desarrollo de EA	Paciente con demencia debida a EA
	$A^{+}T^{+}(N)^{+}$			

Tabla 2.1. Criterios de diagnóstico de la EA basado en los principales biomarcadores. **A** indica la presencia de $A\beta$. **T** indica la presencia de tau fosforilada. **N** indica daño neurodegenerativo y/o neuronal. El símbolo + o un - indica la presencia o no de los biomarcadores anteriores. Tabla adaptada de (Jack et al., 2018).

En cuanto al DCL, los criterios son similares, pero la decisión resulta más compleja. Para diagnosticar esta etapa de la enfermedad, se emplean diversas pruebas similares a las mencionadas anteriormente, con el objetivo de detectar síntomas de deterioro cognitivo. Si estas pruebas no muestran deterioro, pero los biomarcadores revelan la presencia de la patología, el diagnóstico es DCL (Albert et al., 2011). A partir de este punto, se realiza un seguimiento de la evolución de la enfermedad, prestando especial atención a la información proporcionada por el biomarcador $A\beta$ (Albert et al., 2011).

Capítulo 3

Electroencefalografía

3.1. Introducción	10
3.2. Generación de señales	10
3.3. Oscilaciones neuronales y su significado	13
3.4. Alteraciones del EEG asociadas a la enfermedad de Alzheimer	13

3.1. Introducción

En la actualidad, tanto EEG como MEG son métodos no invasivos empleados para comprender la dinámica neuronal del cerebro humano, así como para establecer relaciones entre las dinámicas observadas y posibles patologías (Cohen, 2017). La historia de estas técnicas se remonta a 1929, cuando se realizó el primer registro EEG. No obstante, el primer registro MEG no se llevó a cabo hasta 1968 (Poza, 2008). Ambas técnicas están intrínsecamente vinculadas, ya que comparten un principio fundamental: la generación de corrientes iónicas como resultado de procesos bioquímicos a nivel neuronal (Lopes da Silva, 2013). Si se comparan estas dos metodologías, MEG ha sido ampliamente aceptada en diversos campos y es reconocida como una herramienta complementaria y superior al EEG (Poza, 2008), debido principalmente a su mayor resolución espacial (Lopes da Silva, 2013). Sin embargo, la adopción de MEG en el ámbito clínico es complicada, ya que su instalación y mantenimiento conllevan costes considerables. En contraste, EEG se ha convertido en el método neurofisiológico más utilizado en la práctica clínica en la actualidad, dada su mayor disponibilidad y menor complejidad logística (Lopes da Silva, 2013; Poza, 2008).

3.2. Generación de señales

Para comprender el origen de las señales EEG bajo análisis, se llevará a cabo una exploración del concepto de neurona, que constituye la unidad elemental del sistema nervioso. La neurona desempeña diversas funciones, siendo una de ellas la transmisión del impulso nervioso, cuyos componentes se encuentran detallados en la figura 3.1 (Noback et al., 1993). La transmisión resulta de un proceso bioquímico que se propaga a través de la sinapsis, el punto de conexión entre dos neuronas. La sinapsis tiene la función de regular la propagación de señales, transmitiendo una señal como una serie de impulsos repetitivos, con el propósito de generar un registro de “todo o nada”, a través de la modulación en frecuencia (Poza, 2008).

El fenómeno que origina las señales registradas tiene lugar cuando las neuronas piramidales ubicadas en la corteza cerebral son estimuladas (Sanei & Chambers, 2013). Esta estimulación, que está asociada a la transmisión del impulso nervioso, surge a raíz de variaciones temporales del potencial acumulado a lo largo de la membrana y que se propagan por el axón (Sanei & Chambers, 2013). Por tanto, la información se comunica a través de alteraciones en los potenciales. La propagación se inicia en la membrana (donde se origina el potencial) y se extiende a lo largo del axón hasta los botones terminales, donde tiene lugar la liberación de neurotransmisores que funcionan

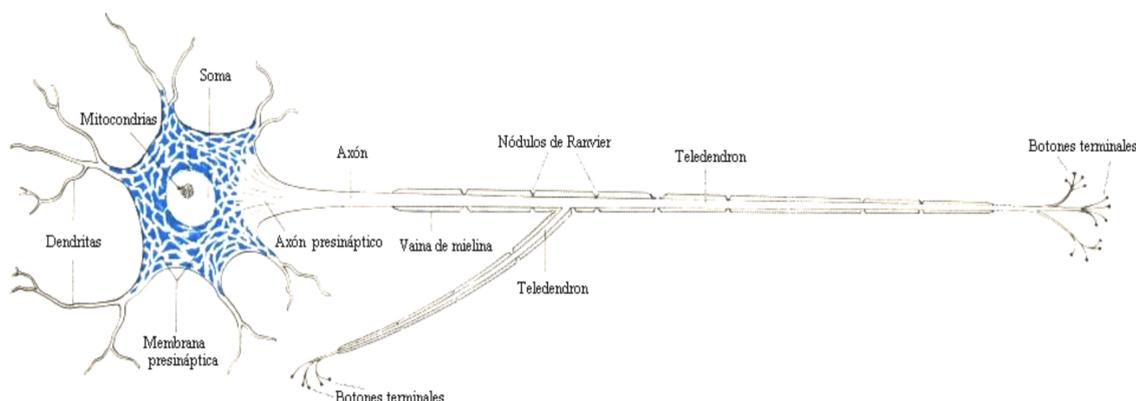


Figura 3.1. Esquema resumen de las partes de una neurona (Noback et al., 1993).

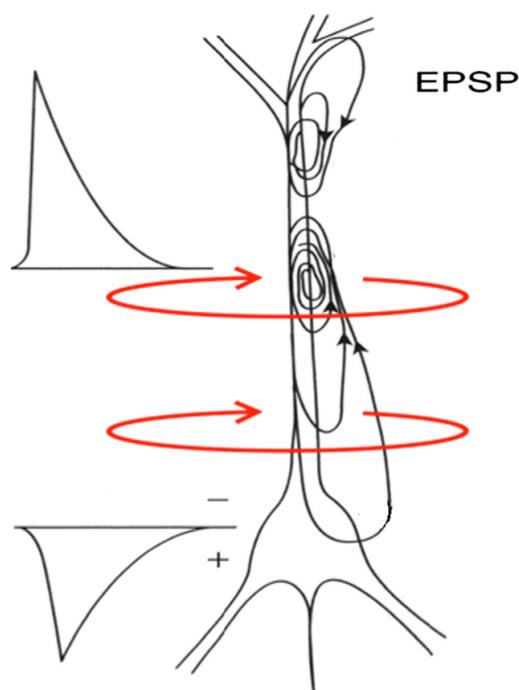


Figura 3.2. Esquema ilustrativo de los campos magnéticos y eléctricos producidos en la neurona. Imagen adaptada de (Lopes da Silva, 2013). Las flechas rojas ilustran los campos magnéticos. Las flechas negras ilustran los campos eléctricos. Las gráficas a la izquierda ilustran la polaridad asociada a los campos en la zona de la neurona donde se encuentran. EPSP: potencial postsináptico excitatorio.

como estímulos para las membranas circundantes (Poza, 2008).

Las neuronas pueden establecer diferentes tipos de asociaciones: una a una, una a varias y varias a varias (Carr & Brown, 2001). Cuando el proceso de transmisión previamente mencionado atraviesa un conjunto de neuronas suficientemente extenso y organizado espacialmente (dendritas dispuestas de forma paralela entre sí y perpendiculares a la corteza cerebral), los cambios en el potencial pueden detectarse en forma de campos electromagnéticos (Lopes da Silva, 2013). La disposición de estos campos se presenta en la figura 3.2: los campos magnéticos se distribuyen alrededor de la neurona, mientras que los campos eléctricos atraviesan la neurona siguiendo la dirección del eje principal. La diferencia de potencial que origina los campos electromagnéticos se denomina potencial postsináptico excitatorio (EPSP) (Lopes da Silva, 2013).

La técnica habitual para identificar las oscilaciones implica el uso de conjuntos de electrodos dispuestos en el cuero cabelludo de los individuos (Cohen, 2017). Sin embargo, la detección no resulta sencilla, ya que las señales deben atravesar diversas capas de tejidos, cada una con sus propias características eléctricas y geométricas, antes de llegar a los sensores. Por tanto, el desafío radica en estimar la ubicación de las fuentes cerebrales que generan los campos detectados fuera del cráneo y la intensidad con la que se originan dichos campos (Lopes da Silva, 2013). Para una detección adecuada, es fundamental contar con un número suficientemente grande de sensores. Aunque un mayor número de electrodos conlleva una mejor resolución espacial (Cassani et al., 2018), también incrementa el coste del sistema. Por tanto, se plantea una relación entre la resolución espacial y el coste.

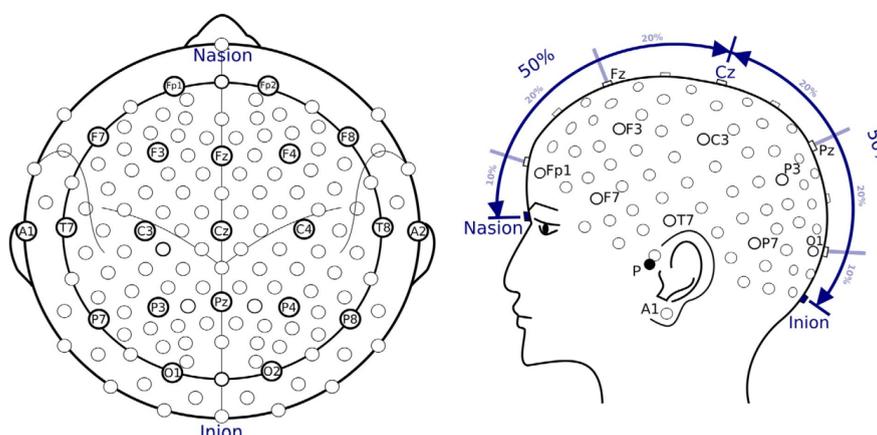


Figura 3.3. Esquema ilustrativo de un sistema 10-5. Los electrodos etiquetados se corresponden a un sistema 10-20. “P” ilustra los puntos preauriculares. “A1” y “A2” representan los lóbulos de la oreja (Müller-Putz, 2020).

La disposición de los sensores sigue un estándar establecido, siendo uno de los enfoques más comunes el sistema internacional 10-20. Este sistema garantiza una colocación uniforme de los sensores en ubicaciones situadas a un 10 % y 20 % de distancia de cuatro puntos de referencia anatómicos clave: nasión, inión y los puntos preauriculares izquierdo y derecho (Müller-Putz, 2020). Para lograr una mayor precisión espacial, es posible agregar más electrodos al sistema 10-20, que resultada en configuraciones más detalladas como los sistemas 10-10 o 10-5 (Müller-Putz, 2020). En la figura 3.3, se presenta un sistema 10-5 con etiquetas solo para los electrodos del sistema 10-20. La figura 3.3 también representa los puntos de referencia y la ubicación de los sensores siguiendo las pautas de nomenclatura propuestas por Müller-Putz (2020): la primera letra denota la región cortical en la que está ubicado el sensor (F=frontal, C=central, P=parietal; O=occipital y FP=fronto-polar); los sensores ubicados en áreas intermedias usan dos letras; los números siguientes a las letras indican el hemisferio en el que se encuentra el sensor, con números impares para el hemisferio izquierdo y pares para el derecho; la línea media de la cabeza se identifica con la letra “z”, y los números aumentan a medida que se alejan de esta línea.

La disposición de los electrodos se denomina montaje (Siuly et al., 2016). En la actualidad, los tipos de montaje más empleados son: **(i) montaje bipolar**, que evalúa la variación de voltaje entre dos electrodos adyacentes; **(ii) montaje referencial**, que mide la variación de voltaje entre un electrodo y otro electrodo de referencia con una posición diferente a los electrodos de registro; **(iii) montaje de referencia media**, en el cual se suma el voltaje de todos los electrodos, se promedia y se utiliza como referencia en cada canal; y finalmente, **(iv) montaje laplaciano**, donde cada canal representa la diferencia entre un electrodo y un promedio ponderado de los electrodos cercanos (Siuly et al., 2016).

Concluyendo la exploración de la detección de las señales electromagnéticas de la actividad cerebral, es esencial considerar el tamaño de las agrupaciones neuronales requeridas para detectar intercambios de potencial. Según Novo-Olivas (2010), se estima que se necesitan al menos 60 millones de neuronas activas simultáneamente para medir el potencial en el cuero cabelludo. Sin embargo, la sensibilidad de los sensores en uso podría implicar la detección de agrupaciones neuronales aún más extensas (Novo-Olivas, 2010).

3.3. Oscilaciones neuronales y su significado

Al explorar las señales EEG, se aprecia una característica esencial de las redes neuronales: actividad oscilatoria (Uhlhaas et al., 2008). Estas oscilaciones en distintas frecuencias desempeñan un papel crucial en la comprensión de los procesos cognitivos que subyacen a la actividad cerebral (Cohen, 2017). Hans Berger (1873-1941) desempeñó un papel fundamental al observar y denominar estas oscilaciones. A pesar de que en su estudio únicamente identificó las bandas de frecuencia alfa y beta, como consecuencia de estudios posteriores, se encontró un mayor número de oscilaciones. Se pueden clasificar en cinco bandas de frecuencia convencionales (Uhlhaas et al., 2008):

- **Delta (δ):** oscilaciones en el rango de frecuencia de 1-4 Hz. han sido identificadas en el tálamo y la corteza cerebral (Niedermeyer et al., 2011). Estas fluctuaciones se vinculan principalmente con estados de sueño, si bien también están presentes fuera del estado de reposo. Sin embargo, esta banda es vulnerable a interferencias debido a las contracciones musculares cercanas a las zonas de detección de las oscilaciones (Sanei & Chambers, 2013).
- **Zeta (θ):** oscilaciones entre 4-8 Hz. Son prominentes en el hipocampo y se registran también en otras áreas del cerebro (Uhlhaas et al., 2008). Estudios realizados en animales han validado la conexión de estas oscilaciones con funciones como el movimiento, la orientación, los procesos cognitivos y las respuestas emocionales (Novo-Olivas, 2010).
- **Alfa (α):** oscilaciones entre 8-13 Hz. Fueron las primeras en ser identificadas por Hans Berger. Estas oscilaciones son más prominentes en el lóbulo occipital y se presentan cuando el individuo está en un estado de relajación con los ojos cerrados (Uhlhaas et al., 2008). Al abrir los ojos, estas oscilaciones tienden a desaparecer, lo que ha dado lugar a debates sobre si están relacionadas con la inhibición o con diversos procesos cognitivos (Uhlhaas et al., 2008).
- **Beta (β):** oscilaciones entre 13-30 Hz. Estas oscilaciones son prominentes en toda la corteza cerebral y se encuentran relacionadas con una amplia gama de tareas cognitivas, además de tener un papel significativo en la coordinación de la actividad neuronal a nivel global (Uhlhaas et al., 2008). Por lo general, esta banda se subdivide en dos bandas: beta-1, que abarca el rango de 13-19 Hz, y beta-2, que se extiende en el rango de 19-30 Hz.
- **Gamma (γ):** oscilaciones que superan los 30 Hz. Se distribuyen de manera extensa por todo el cerebro. Estas oscilaciones tienen vínculos con diversas funciones cognitivas, incluyendo la memoria, la atención, la conciencia y la percepción (Uhlhaas et al., 2008).

3.4. Alteraciones del EEG asociadas a la enfermedad de Alzheimer

Después de haber expuesto la naturaleza de las señales EEG, en la próxima sección se examinan las principales alteraciones que la EA induce en la actividad cerebral. Diversos estudios señalan la presencia de cuatro cambios en el EEG como resultado de la progresión de esta enfermedad (Dauwels et al., 2010):

- **Ralentización del EEG.** Múltiples investigaciones han evidenciado que tanto el DCL como la EA ocasionan una disminución gradual en la frecuencia de oscilación de las señales EEG (Dauwels et al., 2010). Se observa un aumento en la potencia de las frecuencias bajas (delta

y zeta), mientras que las frecuencias más altas (alfa, beta y gamma) experimentan una disminución en su potencia (Jafari et al., 2020). Además, se ha encontrado una correlación entre la ralentización de las señales y la puntuación obtenida en la evaluación cognitiva conocida como *Mini-Mental State Examination* (MMSE) (Rossini et al., 2020), lo que podría permitir la identificación de individuos con DCL en riesgo de desarrollar EA (Jafari et al., 2020).

- **Disminución de la complejidad y regularidad.** Para evaluar estas características, se emplean métricas que buscan identificar los patrones existentes en las señales EEG y diversas propiedades no lineales. Ejemplos de estas medidas son la entropía muestral, la entropía multiescala o la complejidad de Lempel-Ziv (LZC). Los resultados obtenidos en varios estudios revelan que tanto los pacientes con DCL como aquellos con EA presentan una reducción en la complejidad y regularidad de las señales EEG en comparación con individuos sin trastornos cognitivos (Dauwels et al., 2010).
- **Alteraciones en la conectividad funcional entre distintas regiones.** Investigaciones centradas en la conectividad cerebral han revelado que los patrones de conexión entre diferentes áreas cerebrales sufren cambios durante el proceso de envejecimiento (Miraglia et al., 2017); asimismo, la EA conlleva modificaciones en estos patrones de conectividad. La teoría de grafos se utiliza para entender estas alteraciones, examinando la estructura de la red neuronal para simplificar el análisis de la conectividad funcional. Métricas como el coeficiente de agrupamiento, la longitud de camino y el grado de nodo son herramientas esenciales en este análisis (Miraglia et al., 2017). A medida que progresa la EA, se ha observado un aumento en la longitud de camino en la banda zeta y un incremento en el coeficiente de agrupamiento en las bandas zeta y alfa en comparación con los valores de individuos cognitivamente sanos (Jafari et al., 2020; Miraglia et al., 2017). Estos hallazgos podrían indicar que la EA conduce a una mayor aleatoriedad y a una disminución de la eficiencia en las conexiones neuronales (Miraglia et al., 2017).
- **Irregularidades en la conectividad dinámica funcional.** El EEG posibilita la evaluación de los patrones de conectividad funcional dinámica (dFC), que brinda una nueva perspectiva sobre la organización en redes neuronales que cambian en el tiempo. Se han identificado secuencias de meta-estados que se transforman a lo largo del tiempo con una estructura subyacente (Núñez et al., 2021). Estos patrones están alterados en los pacientes que padecen DCL y EA. En particular, se ha observado que las transiciones entre meta-estados son menos estables y más aleatorias debido a la disminución de la estructura dinámica conforme progresa la EA. Los individuos cognitivamente sanos, en comparación, presentan menos transiciones entre meta-estados en comparación con aquellos con DCL y EA (Núñez et al., 2021).

Capítulo 4

Materiales y métodos

4.1.	Introducción	16
4.2.	Conjunto de datos	16
4.2.1.	Registros EEG	16
4.2.2.	Localización de fuentes	17
4.3.	Detección de meta-estados	17
4.3.1.	Correlación de amplitud instantánea (IAC)	17
4.3.2.	Diagramas de recurrencias y ventanas adaptativas	18
4.3.3.	Algoritmo de detección de comunidades para obtención de meta-estados	18
4.4.	Medidas cronectómicas	19
4.4.1.	Singularidad de meta-estado dominante	19
4.4.2.	Área bajo el ICT	20
4.4.3.	Diferencia de permanencia	20
4.4.4.	Probabilidad de ocurrencia del meta-estado nulo	20
4.4.5.	Percentil 5	20
4.4.6.	Grado de antagonismo	20
4.4.7.	<i>Dwell time</i>	21
4.4.8.	<i>Leap size</i>	21
4.4.9.	Complejidad de TAS	21
4.5.	Normalización	22
4.6.	Optimización del algoritmo	22
4.7.	Métodos estadísticos	23
4.8.	Clasificación	24
4.8.1.	Selección de características	24
4.8.2.	Modelos de clasificación	25
4.8.3.	Métricas de clasificación	26

4.1. Introducción

En este capítulo se describen las bases de datos (BBDD) empleadas en esta investigación y la metodología aplicada en su análisis. En primer lugar, se exponen las características sociodemográficas de los individuos cuyas señales EEG conforman el conjunto de datos. Después, se presenta el método utilizado para la adquisición y el preprocesamiento de las señales EEG. La metodología para la detección de los meta-estados se basa en el enfoque previamente empleado por Núñez et al. (2021) en su estudio. Una vez que los meta-estados son extraídos para cada sujeto, se generan secuencias simbólicas y tensores de correlación, sobre los que se implementan y exploran nuevas medidas cronectómicas. Para garantizar la solidez de los resultados, se realiza la normalización mediante subrogación, una técnica que fue optimizada en el transcurso del TFM. A continuación, se detalla la metodología de optimización adoptada. Finalmente, se presentan las medidas cronectómicas empleadas y se explican los algoritmos de selección de características y los modelos de clasificación utilizados para discriminar entre los diversos grupos (controles, pacientes con DCL y pacientes con EA).

4.2. Conjunto de datos

4.2.1. Registros EEG

Los datos provienen de dos BBDD diferentes. La primera base de datos está compuesta por 160 sujetos: 43 sujetos cognitivamente sanos, 67 pacientes con DCL y 50 pacientes con demencia causada por EA. El diagnóstico se basó en los criterios establecidos por el *National Institute on Aging and Alzheimer's Association* (NIA-AA) (Albert et al., 2011). Cabe destacar que todos los participantes fueron adecuadamente informados sobre el estudio y otorgaron su consentimiento para el uso de sus datos. El estudio recibió la aprobación del Comité Ético del Hospital Universitario Río Hortega de Valladolid, en conformidad con el código ético de la Asociación Médica Mundial (AMM). La segunda base de datos está compuesta por 126 sujetos: 38 son sujetos cognitivamente sanos, 38 son pacientes con DCL y 50 pacientes con demencia causada por EA (que se dividen en 25 con demencia leve y 25 con demencia moderada). De forma similar a la primera base de datos, el criterio de diagnóstico fue el establecido por el NIA-AA (Jack et al., 2018). El protocolo seguido fue aprobado por el Comité de Ética de la Universidad de Oporto (Portugal) de acuerdo con la AMM. Estas señales fueron adquiridas en el marco del proyecto europeo POCTEP AD-EEGWA. Las características demográficas de las BBDD se encuentran descritas en la tabla 4.1.

Las señales de EEG fueron registradas mediante un sistema XLTEK® de Natus Medical para la primera base de datos y con un sistema Nihon Kohden Neurofax JE-921A para la segunda. En ambos casos, se utilizaron 19 canales (Fp1, Fp2, Fz, F3, F4, F7, F8, Cz, C3, C4, T3, T4, T5, T6, Pz, P3, P4, O1 y O2). Las señales de EEG del Hospital Universitario Río Hortega se registraron con una frecuencia de muestreo de 200 Hz, mientras que las de POCTEP se adquirieron a 500 Hz. La duración de los registros EEG fue de 5 minutos, durante los cuales los sujetos se mantuvieron despiertos, con los ojos cerrados y en estado de reposo. Para procesar las señales obtenidas, se aplicaron los siguientes pasos (Maturana-Candelas et al., 2020; Núñez et al., 2021): (i) se utilizó un análisis de fuentes independientes (ICA) para eliminar artefactos como parpadeos y latidos del corazón; (ii) se eliminó el ruido de la red eléctrica a 50 Hz mediante un filtro de rechazo de banda, y se limitó el espectro de las señales EEG al rango de actividad cerebral (1-70 Hz) mediante un filtro FIR paso banda con una ventana de Hamming de orden 2000; y (iii) se realizó una inspección

Hospital Universitario Río Hortega						
Grupo	Número de sujetos	de	Edad(años) (m[IQR])	Sexo (H:M)	Educación (A:B)	MMSE (m[IQR])
Control	43		75.8[74.0, 78.7]	13 : 30	16 : 27	29[28, 30]
DCL	67		77.2[72.2, 80.6]	29 : 38	41 : 26	27[26, 28]
EA	50		78.5[75.7, 82.4]	23 : 27	35 : 15	22[20, 24]
Universidad de Oporto						
Grupo	Número de sujetos	de	Edad(años) (m[IQR])	Sexo (H:M)	Educación (A:B)	MMSE (m[IQR])
Control	38		79[76.7, 86.2T]	19 : 19	29:9	29[27, 30]
DCL	38		86[80, 90]	11 : 27	33 : 5	23,5[21, 25]
EA	50		79[74.5, 85]	34 : 16	40 : 10	20[14, 23]

Tabla 4.1. H: hombre; M: mujer. A: educación primaria o inferior; B: estudios superiores; m: mediana; IQR: rango intercuartil (*interquartile range*).

visual de los resultados para seleccionar los primeros 60 segundos consecutivos sin artefactos.

4.2.2. Localización de fuentes

La reconstrucción de los patrones de activación neuronal a nivel de corteza cerebral que provocan las señales EEG se realizó mediante el método *Standardized Low-Resolution Brain Electromagnetic Tomography* (sLORETA) combinado con la información anatómica de la plantilla ICBM152 del Instituto Neurológico de Montreal. Como resultado, se obtuvo un modelo de tres capas de la cabeza que incluye el cerebro, el cráneo y el cuero cabelludo (Núñez et al., 2021). El modelo proporciona un total de 15000 fuentes de información que fueron divididas en 68 regiones de interés (*Region Of Interest, ROI*) de acuerdo con el atlas de Desikan-Killiany (Desikan et al., 2006).

4.3. Detección de meta-estados

El proceso de identificación de los meta-estados en la actividad EEG durante el estado de reposo se basó en el método descrito por Núñez et al. (2021). El proceso consta de tres fases principales: (i)método de correlación de amplitud instantánea; (ii)diagramas de recurrencias y ventanas adaptativas; y (iii)algoritmo de detección de comunidades.

4.3.1. Correlación de amplitud instantánea (IAC)

El método denominado correlación de amplitud instantánea (*Instantaneous Amplitude Correlation, IAC*), desarrollado por Tewarie et al. (2019), se utiliza para generar un tensor de conectividad funcional instantánea a partir de las señales de actividad neuronal reconstruidas a nivel de fuentes cerebrales. El algoritmo IAC calcula la correlación entre la actividad neuronal de diferentes regiones: cuando dos regiones presentan grandes amplitudes simultáneamente, la correlación entre dichas regiones es alta (Tewarie et al., 2019). Aplicar el algoritmo a las bandas convencionales de EEG, delta, zeta, alfa, beta1, beta2 y gamma, resulta en una matriz que muestra los patrones de conectividad funcional en cada instante temporal.

4.3.2. Diagramas de recurrencias y ventanas adaptativas

Los diagramas de recurrencias (*Recurrence Plots*, RPs) se obtuvieron utilizando las matrices de conectividad instantánea obtenidas a partir de la IAC. Estos gráficos revelan la existencia de meta-estados en cada instante temporal. Entre el conjunto de transiciones entre meta-estados es posible observar patrones de transición a lo largo del tiempo. Al realizar el promedio temporal del IAC, se obtienen ventanas adaptativas de los RPs contribuyendo a reducir el ruido y la carga computacional asociados (Núñez et al., 2021).

4.3.3. Algoritmo de detección de comunidades para obtención de meta-estados

El método de detección de comunidades Louvain-GJA se aplicó a las ventanas adaptativas de los RPs para obtener los meta-estados presentes en los patrones de conectividad funcional (FC) instantánea de forma automática (Núñez et al., 2021). El algoritmo fue aplicado a cada base de datos y a cada grupo. Se obtuvieron tres meta-estados distintos en cada banda de frecuencia, los cuales fueron ordenados de manera que sus topologías coincidieran con las observadas en los sujetos cognitivamente sanos, tanto en el grupo con DCL como en el de demencia ocasionada por EA. La figura 4.1 muestra los meta-estados para las bandas delta (figura 4.1-A), zeta (figura 4.1-B), alfa (figura 4.1-C) y beta-1 (figura 4.1-D). En este estudio, no se consideraron beta-2 y gamma debido al elevado coste computacional asociado.

Los meta-estados calculados se utilizaron para la obtención de dos series temporales para cada sujeto:

- La **Secuencia de actividad temporal** o “*Temporal Activation Sequence*” (TAS) es una representación discreta del meta-estado dominante en cada momento. El meta-estado domi-

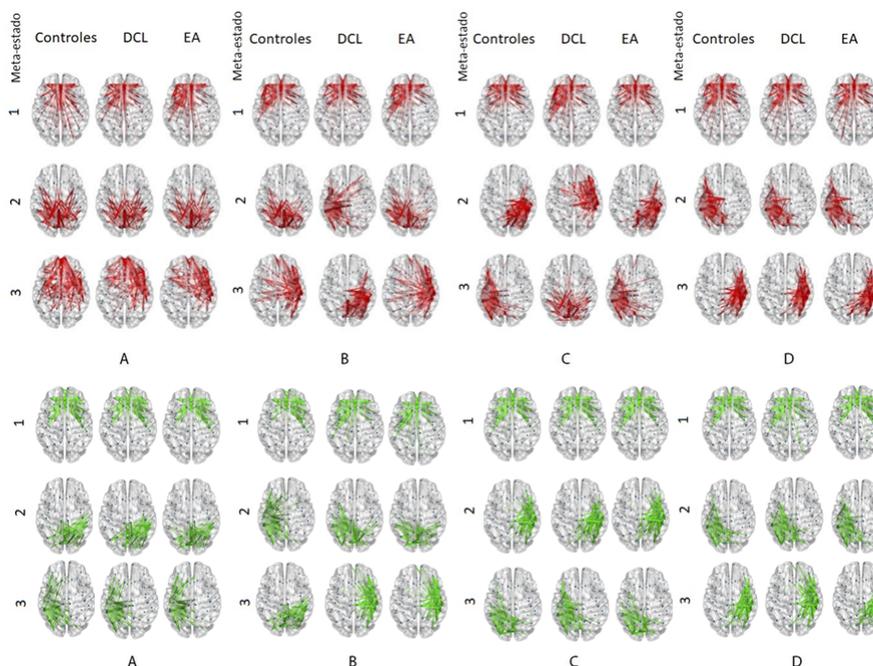


Figura 4.1. Meta-estados detectados por el método de detección de comunidades Louvain-GJA; A: meta-estados en la banda delta; B: meta-estados en la banda zeta; C: meta-estados en la banda alfa; D: meta-estados en la banda beta-1. Color Rojo: Meta-estados obtenidos para la BBDD del Hospital Universitario Río Hortega; Color verde: Meta-estados obtenidos para la BBDD del proyecto POCTEP.

nante se define como aquel con la correlación de Spearman más alta con respecto a la topología de red en un instante dado (Núñez et al., 2021). Se muestra en el diagrama superior de la figura 4.2.

- El **Tensor de correlación instantánea** o “*Instantaneous Correlation Tensor*” (ICT) es un vector que incluye la correlación de Spearman entre los patrones de conectividad instantánea en cada instante y cada meta-estado detectado. Se muestra en el diagrama inferior de la figura 4.2.

4.4. Medidas cronectómicas

Las medidas cronectómicas introducidas por Carretero (2022), Bautista-Salineró (2021) y Núñez et al. (2021), cuyo objetivo es caracterizar las alteraciones en el comportamiento dinámico de la actividad neuronal provocadas por el avance de la EA, fueron utilizadas como características de entrada en los modelos de clasificación supervisada implementados en este TFM. Todas las medidas fueron normalizadas de acuerdo con el método descrito en la sección 4.5.

4.4.1. Singularidad de meta-estado dominante

Esta medida permite cuantificar el grado de definición de los meta-estados. Se obtiene restando los dos valores de correlación más altos en cada muestra del ICT. Un valor alto indicaría que el meta-estado dominante se encuentra bien definido, mientras que un valor bajo correspondería a una secuenciación de topologías de red con distancias de correlación similar a dos meta-estados (Carretero, 2022). La singularidad S se define de acuerdo con la ecuación 4.1:

$$S = \frac{1}{N} \sum_{n=1}^N (x_1[n] - x_2[n]), \tag{4.1}$$

donde N es el número total de muestras del ICT, $x_1[n]$ es el valor de correlación más alto en la muestra n y $x_2[n]$ es el valor de correlación más alto inmediatamente después de $x_1[n]$.

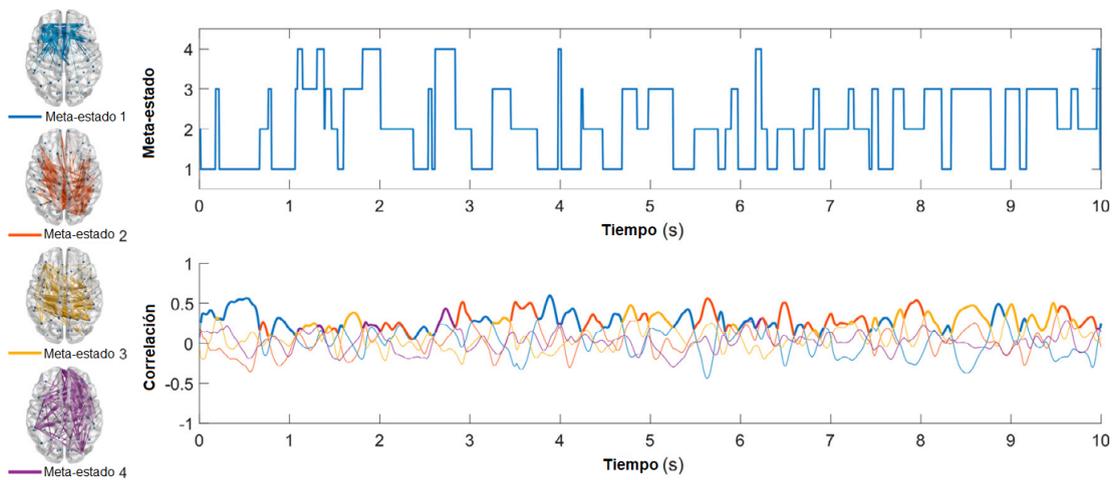


Figura 4.2. Secuencia de activación temporal (diagrama superior) y tensor de correlación instantánea (diagrama inferior) para un paciente de muestra con EA durante los primeros 10 segundos de la grabación en la banda alfa (Núñez et al., 2021).

4.4.2. Área bajo el ICT

Esta medida cuantifica la permanencia temporal y el grado de definición del meta-estado dominante. Resulta del cálculo del meta-estado dominante en cada banda. Únicamente se tuvieron en cuenta los valores de correlación positivos ya que un valor de correlación negativo para un meta-estado indicaría la presencia del meta-estado complementario (Carretero, 2022). La medida se define en la ecuación 4.2. Al utilizar variables discretas, el área se calcula aplicando la siguiente ecuación:

$$A_i = \sum_{n=1}^N x_i[n], \forall x(n) > 0, \quad (4.2)$$

donde N es el número total de muestras del ICT y $x_i[n]$ es el valor de correlación en la muestra n para el meta-estado i . La medida se aplica una vez para cada meta-estado en cada banda.

4.4.3. Diferencia de permanencia

Esta medida cuantifica las alteraciones en la persistencia temporal y la definición de los meta-estados. Se obtiene restando los dos resultados mayores de aplicar la medida “área bajo el ICT”. La medida se define en la ecuación 4.3.

$$D = A_1 - A_2, \quad (4.3)$$

donde A_1 es el resultado de la medida área bajo el ICT más grande y A_2 es el resultado inmediatamente más grande después de A_1 .

4.4.4. Probabilidad de ocurrencia del meta-estado nulo

Esta medida determina los momentos en los que la elección de un meta-estado sobre otro no es clara debido a una baja correlación. Para identificar las zonas que cumplen esta condición, el ICT se divide en segmentos conocidos como “lóbulos”. Cada lóbulo comprende los valores del ICT entre dos puntos de cruce del tensor de correlación, excepto el primer y último lóbulo, que abarcan los valores desde la primera muestra hasta el primer cruce cero, y desde el último cruce cero hasta la última muestra, respectivamente. Se genera una función de distribución que considera las áreas de los lóbulos en cada meta-estado, teniendo en cuenta solo las áreas positivas por el motivo descrito en la medida “área bajo el ICT”. Cuando las áreas de los lóbulos de los tres meta-estados coinciden temporalmente con valores por debajo del percentil 5, se identifica una situación de incertidumbre en la cual la topología observada no coincide con los meta-estados conocidos. A este fenómeno se le denomina “meta-estado nulo” (Carretero, 2022). Se presenta un ejemplo visual en la Figura 4.3. La medida consiste en contar la cantidad de veces que aparece el meta-estado nulo.

4.4.5. Percentil 5

Esta medida es el valor del percentil 5 de las áreas de la medida “probabilidad de ocurrencia del meta-estado nulo” para el meta-estado de mayor área. El objetivo de esta medida es caracterizar los cambios en la distribución de la permanencia temporal en el meta-estado dominante.

4.4.6. Grado de antagonismo

Esta medida es la diferencia entre el mayor y menor valor del ICT para una muestra. Su objetivo es mostrar las alteraciones provocadas por la EA en la arquitectura temporal de la actividad EEG espontánea (Bautista-Saliner, 2021). Se define en la ecuación 4.4.

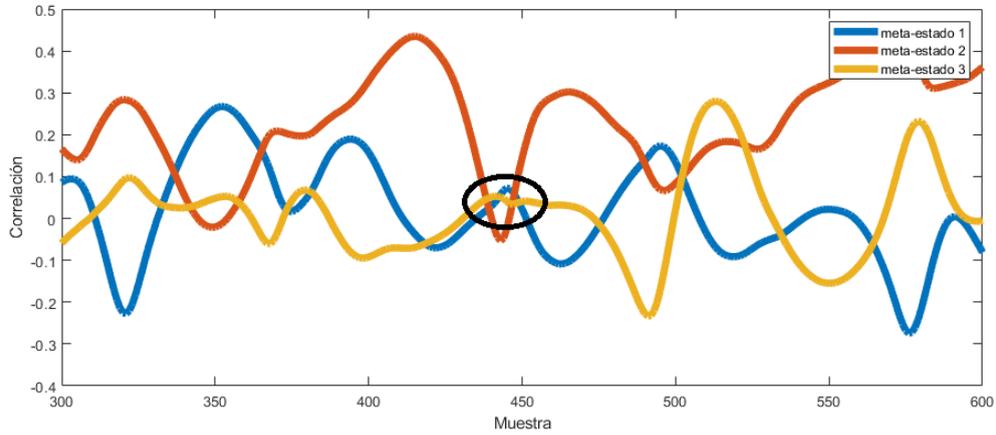


Figura 4.3. Esquema ilustrativo de la medida “probabilidad de ocurrencia del meta-estado nulo”. El círculo resalta la región en la que los tres meta-estados presentan valores de correlación cercanos a cero durante varias muestras (Carretero, 2022).

$$GA = \frac{1}{N} \sum_{n=1}^N (CA[n] - CR[n]), \quad (4.4)$$

donde GA es el valor de la medida, $CA[n]$ es el mayor valor del ICT en la muestra n y $CR[n]$ es el menor valor del ICT en la muestra n .

4.4.7. Dwell time

Este parámetro cuantifica el tiempo que el cerebro pasa en promedio en el mismo meta-estado dominante (Núñez et al., 2021). Ha sido ampliamente empleado en investigaciones sobre los cambios dinámicos en el estado cerebral (Vohryzek et al., 2020). Su definición viene dada por la ecuación 4.5:

$$D = \frac{\frac{d_1}{N} + \dots + \frac{d_m}{N}}{m} \quad (4.5)$$

donde N es el número total de muestras del ICT, d_i es el número de muestras donde domina el meta-estado i y m es el número de meta-estados.

4.4.8. Leap size

Esta medida cuantifica el coste metabólico de la transición entre meta-estados (Núñez et al., 2021). Se calcula como la “distancia” entre meta-estados consecutivos (Ramirez-Mahaluf et al., 2020). Su definición se encuentra en la ecuación 4.6.

$$L = 1 - \sigma_s(IAC) \quad (4.6)$$

donde $\sigma_s(IAC)$ es la correlación de Spearman de las matrices consecutivas de IAC.

4.4.9. Complejidad de TAS

Esta medida cuantifica la estructura subyacente en la secuencia temporal de metaestados. Para su cálculo, se emplea la métrica de complejidad de Lempel-Ziv (LZC) (Lempel & Ziv, 1976). Esta

métrica no paramétrica de complejidad se aplica a señales unidimensionales y está relacionada con la cantidad de subcadenas únicas presentes en los datos y su tasa de ocurrencia (Núñez et al., 2021). Un valor alto de LZC indica una complejidad alta de los datos.

4.5. Normalización

Para asegurar que las medidas aplicadas sobre las señales de EEG son atribuibles a la conectividad funcional dinámica y no a oscilaciones aleatorias o ruido, se emplearon señales subrogadas como método de normalización. Estas señales son réplicas de la señal EEG original, donde se conserva la forma de amplitud pero se modifica la fase de manera aleatoria en cada réplica. Para preservar la relación entre diferentes ROIs, se aplicó el mismo patrón de aleatorización en todas las réplicas de las señales subrogadas, manteniendo así la conectividad estática (Núñez et al., 2021). En el TFM realizado, se utilizó un total de 100 señales subrogadas para llevar a cabo dicha normalización. De acuerdo con Carretero (2022), para cada métrica se realizó la división del valor de la medida correspondiente por la media de los valores obtenidos en cada réplica subrogada. Si el valor normalizado se aleja de 1, se puede asociar con un comportamiento relacionado a la conectividad funcional dinámica de origen neural. Por otro lado, si el resultado de la normalización se acerca a 1, no se puede descartar que la medida esté influenciada por factores aleatorios o ruido.

4.6. Optimización del algoritmo

Para llevar a cabo el cálculo de las señales subrogadas fue necesario seguir el proceso de detección de meta-estados descrito en la sección 4.3 para cada señal subrogada. El cálculo de 100 señales subrogadas por cada sujeto tiene un coste computacional elevado. Se realizó un estudio del tiempo invertido en cada fase del proceso que permitió observar que el cuello de botella del proceso se encuentra en la obtención de los RPs. En la figura 4.4 se muestra un esquema con el porcentaje de tiempo invertido en cada fase del proceso.

El método de subrogación se encuentra implementado en MATLAB®. Se trata de un lenguaje interpretado e interactivo que permite un proceso de programación con mayor abstracción que un lenguaje compilado y, por tanto, una elaboración más rápida e intuitiva del código (Rackauckas, 2023). Sin embargo, la velocidad de ejecución de programas en lenguajes compilados supera la de los lenguajes interpretados (Merelo-Guervós et al., 2016). Una vez se comprobó que el código implementado en MATLAB se encontraba optimizado, se plantearon las siguientes posibilidades para aumentar el rendimiento del proceso:

- **Ficheros ejecutables.** Los lenguajes interpretados utilizan un intérprete que traduce el código fuente a código máquina y lo ejecuta línea por línea en cada instrucción de un programa (Kwame et al., 2017). El uso de 100 señales subrogadas implica un tiempo dedicado a la interpretación del código en cada iteración. El uso de archivos ejecutables evita el proceso de interpretación y traducción del código fuente a código máquina (Kwame et al., 2017).
- **Ficheros en C y C++.** Los lenguajes de programación C y C++ son lenguajes compilados. Este tipo de lenguaje permite la programación a bajo nivel, obteniendo un código de ejecución más eficiente que los generados por lenguajes interpretados. Los lenguajes compilados traducen a código máquina la totalidad de un programa antes de ejecutarlo (Dmitrović, 2023).

PORCENTAJE DE TIEMPO INVERTIDO EN CADA SEÑAL SUBROGADA

■ IAC ■ RPs ■ Detección de comunidades

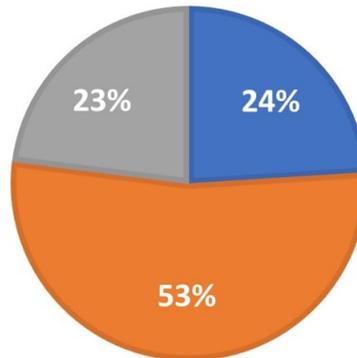


Figura 4.4. Porcentaje de tiempo invertido en cada etapa de la obtención de meta-estados.

- Cálculo paralelo.** El cálculo paralelo divide la ejecución de una tarea entre dos o más núcleos de procesado permitiendo reducir el tiempo total de ejecución (Scott et al., 2021). Dentro de las diversas opciones para paralelizar el cálculo, se destacan dos en particular para este TFM: i) paralelización en Unidad Central de Procesamiento (CPU), que permite paralelizar el cálculo en cualquier equipo que disponga de un procesador de más de un núcleo; y ii) paralelización en GPU, concretamente haciendo uso de Arquitectura Unificada de Dispositivos de Cómputo (*Compute Unified Device Architecture* o CUDA). CUDA es una interfaz de programación de aplicaciones (API) que permite la coordinación entre la CPU y la unidad de procesamiento gráfico (GPU) de forma exclusiva para tarjetas GPU Nvidia (Asaduzzaman et al., 2021). En aquellas tarjetas gráficas de marca distinta a Nvidia se utiliza la API OpenCL. En el TFM se ha optado por utilizar CUDA debido a la compatibilidad nativa de MATLAB con las GPU Nvidia y a la disponibilidad de computadoras equipadas con GPU de este tipo.

Para implementar la ejecución de ficheros escritos en C, C++ y CUDA en MATLAB se utilizó la API de ficheros MEX. Esta API permite utilizar estos ficheros desde la línea de comandos de MATLAB o desde scripts *.m como si se tratase de funciones estándar (Matusiak, 2021). Para conseguir una mayor optimización fue necesaria la elaboración de funciones implementadas en MATLAB no compatibles con la ejecución en C y C++. Para el TFM se programó una versión optimizada de los algoritmos de obtención de la correlación de Pearson, del algoritmo de Louvain-GJA y de la IAC.

4.7. Métodos estadísticos

Después de evaluar la normalidad y la homocedascidad de los parámetros calculados utilizando las pruebas de Lilliefors y la prueba de Levene, respectivamente, se llegó a la conclusión de que los datos no cumplían las suposiciones de los test paramétricos. Por lo tanto, se optó por utilizar pruebas no paramétricas. Una vez que se calcularon las medidas, dado que había tres grupos (control, DCL y EA), se aplicó el test estadístico de Kruskal-Wallis para detectar diferencias globales. En las bandas donde los resultados fueron significativos de manera global, se empleó la prueba de la U de Mann-Whitney para encontrar diferencias entre pares de grupos. Los resultados de ambos test estadísticos se corrigieron mediante el método del *false discovery rate* (FDR). Se consideró significativo aquel *p-valor* inferior a 0.05.

4.8. Clasificación

Las medidas descritas se utilizaron para la creación de un modelo de clasificación. Cada BBDD se dividió en dos conjuntos: entrenamiento y test, siendo las dimensiones del conjunto de entrenamiento un 80 % del número de sujetos de la base de datos completa y un 25 % en el caso del conjunto de test. Sobre las medidas del conjunto de entrenamiento se aplicó un proceso de selección de características para conservar las más relevantes y eficientes con el fin de mejorar el rendimiento de la clasificación (Alhenawi et al., 2022). A continuación, se utilizó validación cruzada en K grupos (*k-fold cross-validation*) para la elección del modelo de clasificación de mayor precisión de nuevo utilizando el conjunto de entrenamiento. De acuerdo con el estudio realizado por Zhang & Liu (2023), el valor de K elegido fue 10 debido a que minimiza el riesgo de error cometido en la predicción comparado con otros posibles valores de K . Una vez se encontró el mejor modelo para las características de entrada, se procedió a evaluar la capacidad diagnóstica del modelo sobre el conjunto de test.

4.8.1. Selección de características

En el proceso de selección de características, se utilizó el método univariable, eliminación recursiva de variables (*Recursive Feature Elimination* o RFE), *Sequential Forward Selection* (SFS) y *Fast-Correlation-Based-Filter* (FCBF). El primer método utilizado fue el método univariable. El método consiste en la creación de un árbol de decisión utilizando una única característica y la etiqueta objetivo. El algoritmo clasifica la calidad de todas las características según el área de la curva de la región de convergencia (ROC) del modelo y selecciona las características de valor más alto (Jain & Saha, 2022). En el TFM se utilizó un algoritmo de K -mejores univariable. Esta variante selecciona las K características de área de curva ROC más alto.

El siguiente método implementado fue RFE, que consiste en la utilización de un modelo de clasificación que inicialmente se entrena con el total de las características y determina de forma interna la importancia de cada una mediante una puntuación. Las características menos importantes se eliminan. Se vuelve a entrenar el modelo y se calculan de nuevo las puntuaciones de importancia de forma recursiva hasta que se conserva el número de características requerido (Kumari et al., 2023).

Asimismo, se aplicó el algoritmo SFS, que utiliza un estimador inicialmente entrenado con cero características. En cada iteración del algoritmo, se añade una única característica que maximiza el rendimiento. Una vez seleccionada una característica, se repite el proceso añadiendo una nueva característica al conjunto de características seleccionadas. El algoritmo conserva la combinación de características que consigue los mejores resultados (Tahir et al., 2023). El número final de características lo determina el usuario (Shafiee et al., 2021). El número de características se eligió utilizando el mínimo de características entre la raíz cuadrada del número de sujetos que componen la base de datos (Figueroa et al., 2012) y el número de características de mayor precisión de acuerdo al algoritmo RFE aplicado a una máquina de vectores de soporte o *Support Vector Machines* (SVM). El criterio del mínimo número se tomó para evitar la elección de un número demasiado bajo de características cuando las BBDD se analizan de forma individual y demasiado alto cuando se analiza su combinación.

Finalmente, se utilizó el método FCBF. Este se basa en la incertidumbre simétrica (*SU*), que mide la información mutua entre cada característica y las variables de pertenencia a un grupo (Yu & Liu, 2004). De acuerdo con Yu & Liu (2004), el método consta de dos fases:

- La primera fase consiste en el análisis de la importancia de las características. La SU entre cada característica X_i y la pertenencia a un grupo Y se calcula como:

$$SU(X_i, Y) = 2 \left[\frac{H(X_i) - H(X_i|Y)}{H(X_i) + H(Y)} \right], i = 1, 2, \dots, I, \quad (4.7)$$

donde H corresponde a la entropía de Shannon e I es el número de características. SU resulta en un número en el rango $[0,1]$; el valor 1 indica que conociendo el valor de X_i o Y , puede predecirse el valor de la otra variable; sin embargo, el valor 0 indica que no existe relación entre ambas variables.

- Para terminar, se descartan las variables irrelevantes. Para ello, se compara la SU obtenida para un par de variables X_i y X_j con la obtenida para la característica Y . Si la variable X_i aporta comparte más información con X_j que con Y , es decir $SU(X_i, X_j) > SU(X_i, Y)$, y $SU(X_i, Y) > SU(X_j, Y)$; entonces se descarta la característica j y no se tiene en cuenta para las posteriores comparaciones (Ruiz-Gómez et al., 2018).

4.8.2. Modelos de clasificación

Para llevar a cabo la elección del modelo de clasificación utilizado, se utilizó *K-fold cross-validation*. Este algoritmo divide la base de datos en K particiones del mismo tamaño. Se utilizan $K-1$ partes para entrenamiento y una para validación. El proceso se repite hasta que cada una de las K partes se haya utilizado como conjunto de validación. Finalmente, se realiza la media de las métricas realizadas con cada conjunto de validación para observar el resultado final (Berrar, 2019). En este estudio se midió la precisión de cada modelo debido a que en el Trabajo de Fin de Grado de Carretero (2022) fue el parámetro más variable entre las pruebas de validación realizadas. Además, se llevo a cabo un barrido de los parámetros de cada modelo de clasificación utilizado para poder elegir el mejor modelo posible en cada caso. Los modelos utilizados en el proceso de validación cruzada fueron regresión logística, regresión logística con penalización L2, red neuronal basada en un perceptrón multicapa, SVM, árbol de decisión y *random forest*.

El primer modelo implementado fue regresión logística. Este modelo asigna un coeficiente a cada característica de forma independiente que mide su contribución individual sobre la predicción final (Boateng & Abaye, 2019). El modelo calcula probabilidades de acuerdo con la ecuación 4.8:

$$\ln \left(\frac{P(Y)}{1 - P(Y)} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (4.8)$$

donde Y es la variable dependiente, X_i es la característica de entrada i -ésima y β_i es el coeficiente asignado a cada X_i . Desarrollando la ecuación 4.8 se llega a 4.9.

$$P(Y) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}} \quad (4.9)$$

Mediante la ecuación 4.9 el modelo predice de forma directa la probabilidad de pertenencia a la clase Y a partir del conjunto de características. El objetivo final del modelo es la predicción del conjunto de coeficientes β_i que consiga el mejor modelo de clasificación posible haciendo uso del estimador de máxima verosimilitud (Boateng & Abaye, 2019). Para evitar el sobreajuste al asignar los coeficientes, se puede utilizar regularización. Este proceso consiste en favorecer los modelos más simples frente a los más complejos (Bühlmann & Van De Geer, 2011). En el TFM se creó un modelo cuya penalización fue de tipo $L2$, que añade una penalización igual al cuadrado de la

magnitud de los coeficientes penalizando en mayor medida a los coeficientes más altos (Bühlmann & Van De Geer, 2011).

A continuación, se implementó un modelo de red neuronal: perceptrón multicapa (*Multi-Layer Perceptron* o MLP). Su funcionamiento se basa en una estructura de tres capas: i) la capa de entrada, que recibe los datos de cada característica de entrada, ii) la capa oculta, que procesa la información de entrada asignando pesos a los valores de entrada, y iii) la capa de salida, que muestra el resultado de la clasificación (Renganathan, 2019). En este tipo de red neuronal, los valores de entrada se multiplican por sus respectivos pesos y se introducen en la capa oculta, la cual pondera nuevamente los datos de entrada multiplicándolos por sus pesos y transfiere el resultado a la capa de salida. Para encontrar los pesos, se utiliza el algoritmo de retropropagación. La función de activación empleada en este tipo de red neuronal es la función sigmoide (Renganathan, 2019).

El siguiente modelo implementado fue SVM. Este clasificador trata de encontrar el hiperplano que mejor divide los datos de entrenamiento entre las clases posibles. Para calcular el hiperplano se trata de tomar un margen de decisión \vec{w} lo más estrecho posible y libre de datos alrededor del umbral de decisión (Breyer et al., 2022). El algoritmo decide a qué clase pertenece un conjunto de datos en función de su posición respecto a la curva de decisión.

El siguiente modelo de clasificación implementado fue el árbol de decisión. Este tipo de clasificador funciona realizando pruebas sucesivas para clasificar, donde cada una de las características numéricas se compara con un umbral. La clasificación se lleva a cabo en forma de árbol: en cada nodo se procesan las características para clasificar en una categoría (Charbuty & Abdulazeez, 2021). Según el resultado de la comparación, se avanza a otro nodo hasta que se logra clasificar los datos de entrada en un grupo. Finalmente, se implementó una variante de este tipo de clasificador: el modelo *random forest*. Este tipo de clasificador utiliza una colección de árboles de decisión que únicamente contiene las características que mejor clasifican los datos de entrenamiento (Jackins et al., 2021). La relevancia de las características se cuantifica mediante la probabilidad de pertenecer a un grupo en función del resultado de una única característica, declarando como relevante aquellas características de probabilidad más alta. Este tipo de clasificador trabaja con un árbol de decisión por cada característica de entrada considerada como relevante (Jackins et al., 2021).

4.8.3. Métricas de clasificación

La implementación de los modelos descritos en la sección 4.8.2 tiene como objetivo llevar a cabo la clasificación de individuos en tres categorías distintas: controles, pacientes con DCL y pacientes con EA. Con el propósito de valorar la calidad de la clasificación, se emplean varios indicadores. El *coeficiente kappa de Cohen* (k) se utiliza para evaluar la concordancia entre los grupos predichos y los grupos observados, tomando en consideración la probabilidad de coincidencia por azar (Ruiz-Gómez et al., 2018). Siguiendo la metodología propuesta por (Carretero, 2022), se calcula la precisión, que refleja el porcentaje de sujetos clasificados correctamente, junto con la sensibilidad y especificidad, que indican el porcentaje de sujetos positivos y negativos clasificados adecuadamente, respectivamente. Estas medidas se calculan haciendo uso de un conjunto de validación no utilizado en el entrenamiento de los modelos ni en la selección de características, correspondiente al 25 % de los sujetos que componen cada base de datos. El resultado del diagnóstico se expone en una matriz de confusión, la cual se muestra en la tabla 4.2.

$$\text{Precisión} = \frac{VP + VN}{VP + VN + FP + FN}, \quad (4.10)$$

$$\mathbf{Sensibilidad} = \frac{VP}{VP + FN}, \tag{4.11}$$

$$\mathbf{Especificidad} = \frac{VN}{VN + FP}, \tag{4.12}$$

Donde VP representa la cantidad de verdaderos positivos, VN indica la cantidad de verdaderos negativos, FP representa los casos en los que se identificaron erróneamente como positivos, y FN son los casos en los que se identificaron erróneamente como negativos. La matriz de confusión, que se muestra en la tabla 4.3, proporciona una visualización de estas medidas de evaluación, enfocándose específicamente en la distinción entre el grupo de controles y el resto de grupos.

Observado\Estimado	Control	DCL	EA
Control	VN	FP en DCL	FP en EA
DCL	FN	VP en DCL	FP en EA
EA	FN	FP en DCL	VP en EA

Tabla 4.2. Estructura de la matriz de confusión. VP: verdadero positivo; FP: falso positivo; VN: verdadero negativo; FN: falso negativo

Observado\Estimado	Control	Pacientes
Control	VN	FP
Demencia	FN	VP

Tabla 4.3. Estructura de la matriz de confusión englobando DCL y EA bajo el grupo “Pacientes” . VP: verdadero positivo; FP: falso positivo; VN: verdadero negativo; FN: falso negativo

Capítulo 5

Resultados

5.1. Introducción	29
5.2. Optimización	29
5.3. Singularidad de meta-estado dominante	30
5.4. Área bajo el ICT	31
5.5. Diferencia de permanencia	32
5.6. Probabilidad de ocurrencia del meta-estado nulo	33
5.7. Percentil 5°	34
5.8. Grado de antagonismo	34
5.9. <i>Dwell time</i>	35
5.10. <i>Leap size</i>	36
5.11. Complejidad de TAS	37
5.12. Clasificación	38
5.12.1. K-mejores	39
5.12.2. FCBF	40
5.12.2.1. Clasificación con 3 clases	41
5.12.2.2. Clasificación con 2 clases	43
5.12.3. SFS	44
5.12.3.1. Clasificación con 3 clases	44
5.12.3.2. Clasificación con 2 clases	45
5.12.4. RFE	47
5.12.4.1. Clasificación con 3 clases	49
5.12.4.2. Clasificación con 2 clases	50

5.1. Introducción

En este capítulo se exponen los resultados derivados de las medidas y los algoritmos de clasificación que fueron introducidos en el capítulo anterior. Los resultados de las medidas se presentan mediante diagramas que permiten visualizar la distribución de los datos. La línea naranja representa el valor uno. El resultado de normalizar los datos mediante señales subrogadas implica que el valor uno se asocia con un comportamiento de origen no neural.

5.2. Optimización

El equipo utilizado en las pruebas cuenta con una CPU Intel(R) Core(TM) i9-9900K de 3,60 GHz, una GPU NVIDIA GeForce RTX 3080 Ti con 12 GB de memoria, y 64 GB de RAM. La optimización obtenida en cada fase del proceso de cálculo de las señales subrogadas a través de los métodos detallados en el capítulo previo se presenta en la tabla 5.1. En ninguna etapa del proceso se consigue reducir el tiempo de ejecución mediante el uso de ficheros ejecutables. En el caso del IAC, se consigue una reducción del 41,60 % de tiempo de cálculo mediante la implementación en C. La obtención de RPs se ve reducida un 61,35 % al utilizar ficheros en C y un 64,70 % mediante el uso de la GPU. Para finalizar, el algoritmo de detección de comunidades Louvain-GJA reduce el tiempo de ejecución un 10,12 % mediante ficheros en C y un 30,38 % mediante la utilización de GPU.

La Tabla 5.2 presenta los tiempos dedicados al cálculo de señales subrogadas para un individuo y para una base de datos completa, tanto antes como después de aplicar la optimización. Calcular las señales subrogadas para un individuo implica realizar 100 iteraciones del proceso. El tiempo necesario para calcular las señales de un único individuo sin aplicar la optimización es de 2,48 horas. Mediante la implementación de métodos de optimización, se reduce el tiempo de ejecución a 1,79 horas. Al aplicar la mejor técnica de optimización en cada etapa, se logra una reducción del 38,52 % en el tiempo total del proceso en comparación con la versión original del método de subrogación.

Reducción de tiempo de ejecución			
	Ficheros ejecutables	Ficheros en C/C++	GPU
IAC	-	41,60 %	-
RPs	-	61,35 %	64,70 %
Louvain-GJA	-	10,12 %	30,38 %

Tabla 5.1. Reducción del tiempo de ejecución conseguida en cada etapa del proceso de subrogación.

Tiempo total empleado		
	1 sujeto	160 sujetos
Antes de optimizar	2,48 horas	396,67 horas
Después de optimizar	1,79 horas	286,36 horas
Diferencia	0,69 horas	110,3 horas

Tabla 5.2. Tiempo empleado en el proceso de subrogación antes y después de aplicar los métodos de optimización.

5.3. Singularidad de meta-estado dominante

Esta métrica presenta resultados globales estadísticamente significativos en la banda alfa ($p = 0,0084$, Kruskal-Wallis, corrección FDR) en la BBDD del Hospital Universitario Río Hortega. En esta banda se identifican diferencias estadísticamente significativas entre el grupo de control y los pacientes con DCL ($p = 0,0094$, prueba U de Mann-Whitney con corrección FDR), así como entre el grupo de control y los pacientes con EA ($p = 0,0017$, prueba U de Mann-Whitney con corrección FDR). En el caso de la BBDD del proyecto POCTEP, los resultados globales son significativos en la banda zeta ($p = 0,03$, Kruskal-Wallis, corrección FDR), presentando diferencias estadísticamente significativas entre el grupo de control y los pacientes con EA ($p = 0,0017$, prueba U de Mann-Whitney con corrección FDR). En ambos casos, se observa un valor superior en los sujetos de control en comparación con los pacientes con EA. Además, en ambas BBDD, los sujetos de control exhiben una gama de valores más distante de 1 en comparación con los pacientes. Esto denota un patrón más evidente de conectividad dinámica funcional, en virtud de la actividad neuronal genuina. Los resultados se muestran en la figura 5.1.

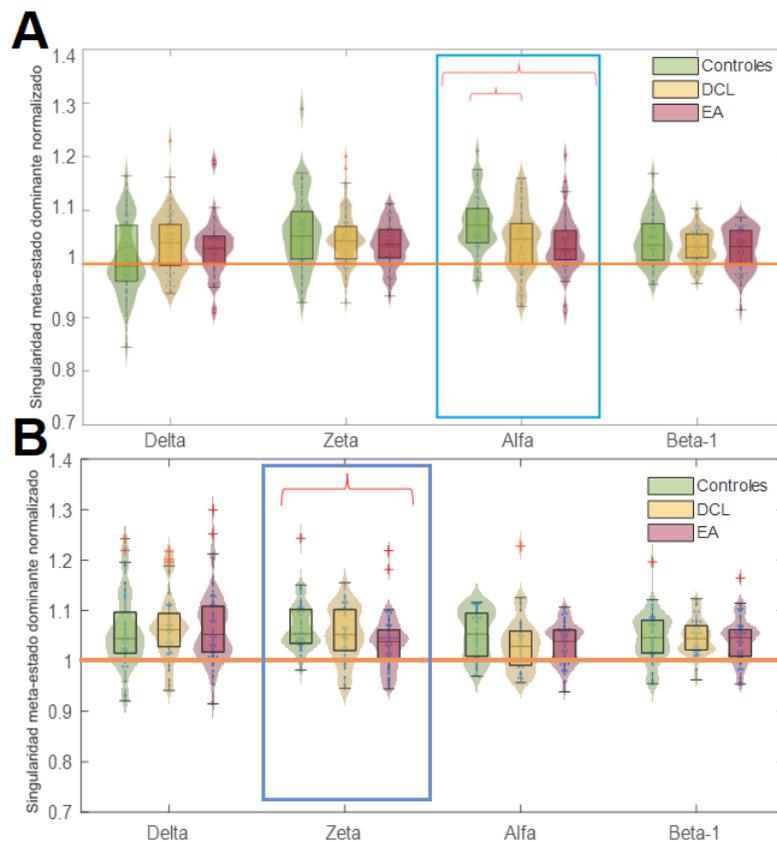


Figura 5.1. Resultado de la medida “singularidad de meta-estado dominante”. A: Base de datos del Hospital Universitario Río Hortega B: Base de datos del proyecto POCTEP. El cuadro azul indica diferencias estadísticamente significativas globales ($p < 0,05$, Kruskal-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).

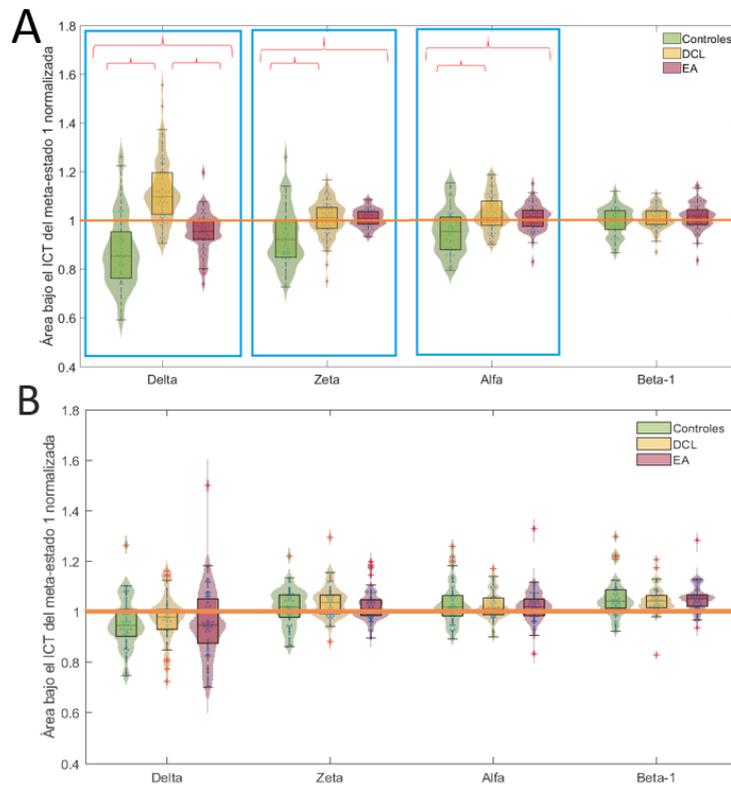


Figura 5.2. Resultado de la medida “área bajo el ICT” en el meta-estado dominante. A: Base de datos del Hospital Universitario Río Hortega B: Base de datos del proyecto POCTEP. Los cuadros azules indican diferencias globales estadísticamente significativas ($p < 0,05$, Kruskal-Wallis, corrección FDR). Las llaves rojas indican diferencias entre pares de grupos estadísticamente significativas ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).

5.4. Área bajo el ICT

El resultado de aplicar esta medida sobre el meta-estado dominante en la BBDD del Hospital Universitario Río Hortega da lugar a diferencias estadísticamente significativas en las bandas bandas delta ($p = 4,1329 \cdot 10^{-16}$, Kruskal-Wallis, corrección FDR), zeta ($p = 1,0181 \cdot 10^{-4}$, Kruskal-Wallis, corrección FDR) y alfa ($p = 1,0181 \cdot 10^{-4}$, Kruskal-Wallis, corrección FDR). En relación a las diferencias entre pares de grupos, en la banda delta se evidencian diferencias estadísticamente significativas entre los controles y los pacientes con DCL ($p = 5,58 \cdot 10^{-12}$, prueba U de Mann-Whitney, corrección FDR), así como entre los pacientes con DCL y aquellos con EA ($p = 1,0562 \cdot 10^{-11}$, prueba U de Mann-Whitney, corrección FDR), y también entre los controles y los pacientes con EA ($p = 4,7495 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR). En la banda zeta, estas diferencias se observan entre los sujetos de control y los pacientes con EA ($p = 2,0318 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR), así como entre los controles y los pacientes con DCL ($p = 2,0318 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR). Finalmente, en la banda alfa, se encuentran diferencias estadísticamente significativas entre los controles y los pacientes con DCL ($p = 1,1541 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR), así como entre los controles y los pacientes con EA ($p = 5,983 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR). Sin embargo, el resultado de aplicar esta medida en la BBDD del proyecto POCTEP no es significativo en ninguna de las bandas de frecuencia. Los resultados se muestran en la figura 5.2

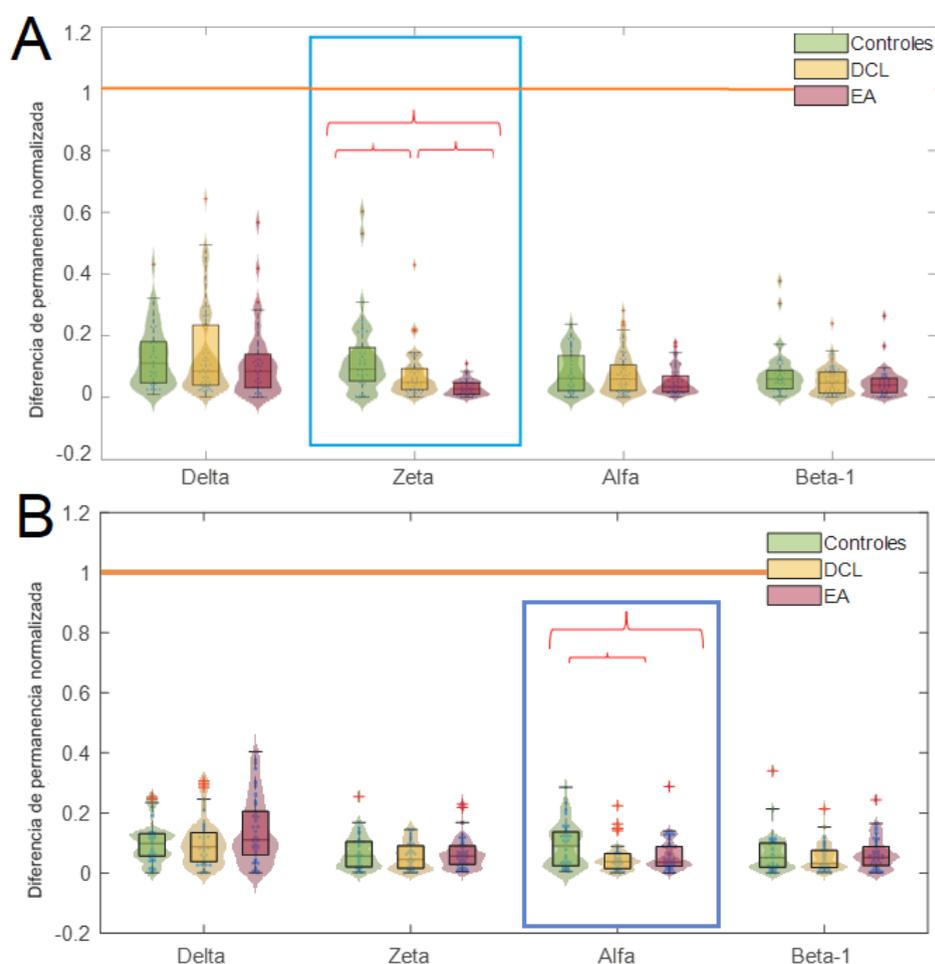


Figura 5.3. Resultado de la medida “diferencia de permanencia”. A: Base de datos del Hospital Universitario Río Hortega B: Base de datos del proyecto POCTEP. El cuadro azul indica diferencias globales estadísticamente significativas ($p < 0,05$, Kruskal-Wallis, corrección FDR). Las llaves rojas indican diferencias entre pares de grupos estadísticamente significativas ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).

5.5. Diferencia de permanencia

El resultado de aplicar la medida en la BBDD del Hospital Universitario Río Hortega da lugar a diferencias estadísticamente significativas globales en la banda zeta ($p = 5,2948 \cdot 10^{-7}$, Kruskal-Wallis, corrección FDR). En este caso, las diferencias entre pares de grupos son estadísticamente significativas entre controles y pacientes con DCL ($p = 0,0028$, prueba U de Mann-Whitney, corrección FDR), entre pacientes con DCL y EA ($p = 4,1917 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR) y entre controles y pacientes con EA ($p = 5,1333 \cdot 10^{-7}$, prueba U de Mann-Whitney, corrección FDR). Los resultados de aplicar la medida en la BBDD del proyecto POCTEP resultan significativos de forma global en la banda alfa ($p = 0,03$, Kruskal-Wallis, corrección FDR). En este caso, los resultados presentan diferencias estadísticamente significativas entre controles y pacientes con DCL ($p = 0,0442$, prueba U de Mann-Whitney, corrección FDR), y entre

controles y pacientes con EA ($p = 0,0442$, prueba U de Mann-Whitney, corrección FDR). En ambas BBDD, los resultados presentan valores alejados del uno en todas las bandas de frecuencia, por tanto, se puede excluir la posibilidad de que el comportamiento sea resultado de oscilaciones aleatorias o ruido. Los resultados se muestran en la figura 5.3.

5.6. Probabilidad de ocurrencia del meta-estado nulo

El resultado de aplicar esta medida en ambas BBDD se muestra en la figura 5.4. En el caso de la base de datos del Hospital Universitario Río Hortega, no se encuentran diferencias globales estadísticamente significativas. Sin embargo, el resultado de aplicar la medida en la BBDD del proyecto POCTEP presenta diferencias globales estadísticamente significativas en la banda zeta ($p = 0,01$, Kruskal-Wallis, corrección FDR), donde existen diferencias estadísticamente significativas entre controles y pacientes con EA ($p = 0,0025$, prueba U de Mann-Whitney, corrección FDR). Además, los valores se encuentran aproximadamente centrados en torno a uno. Por tanto, el resultado de la medida no parece tener relación con las propiedades dinámicas de la conectividad funcional de origen neural.

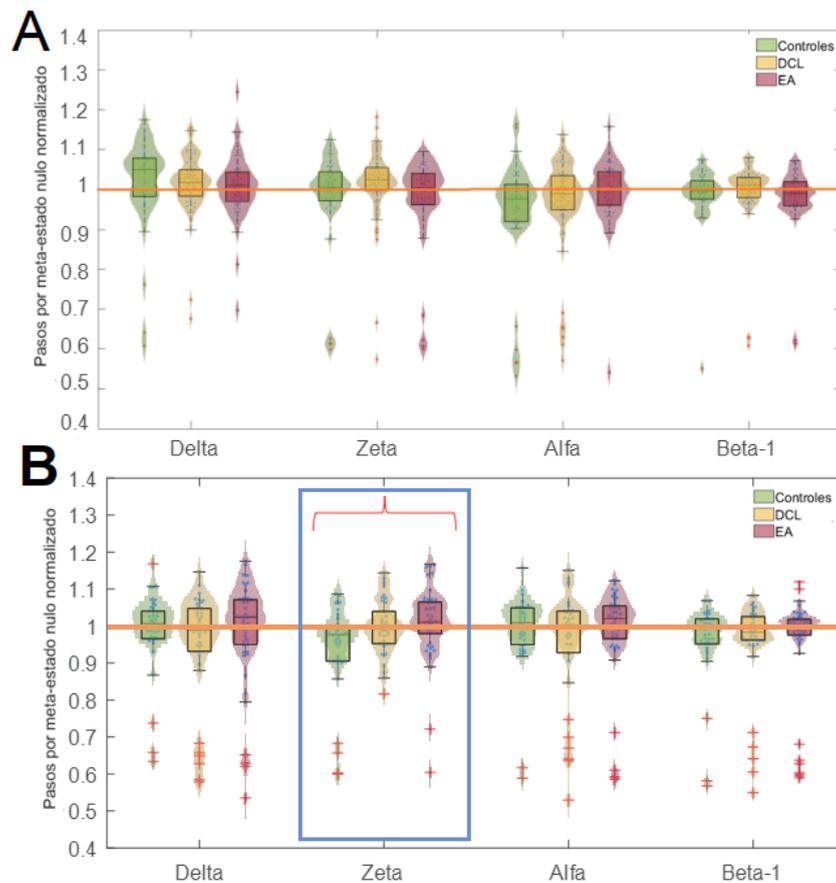


Figura 5.4. Resultado de la medida “probabilidad de ocurrencia del meta-estado nulo”. A: Base de datos del Hospital Universitario Río Hortega B: Base de datos del proyecto POCTEP. Los cuadros azules indican resultados globales significativos ($p < 0,05$, Kruskal-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).

5.7. Percentil 5°

En la figura 5.5 se muestra el resultado de aplicar esta medida en ambas BBDD. En el caso de la base de datos del Hospital Universitario Río Hortega, los resultados presentan diferencias globales estadísticamente significativas en las bandas delta ($p = 0,0174$, Kruskal-Wallis, corrección FDR), zeta ($p = 0,0174$, Kruskal-Wallis, corrección FDR), alfa ($p = 0,0174$, Kruskal-Wallis, corrección FDR) y beta-1 ($p = 0,0174$, Kruskal-Wallis, corrección FDR). Asimismo, las diferencias entre pares de grupos son estadísticamente significativas entre controles y pacientes con EA en las bandas delta ($p = 0,0071$, prueba U de Mann-Whitney, corrección FDR), zeta ($p = 0,0081$, prueba U de Mann-Whitney, corrección FDR), alfa ($p = 0,0081$, prueba U de Mann-Whitney, corrección FDR) y beta-1 ($p = 0,0087$, prueba U de Mann-Whitney, corrección FDR).

Los resultados de aplicar la medida en la BBDD del proyecto POCTEP presentan diferencias globales estadísticamente significativas en delta ($p = 8,55 \cdot 10^{-6}$, Kruskal-Wallis, corrección FDR), zeta ($p = 8,55 \cdot 10^{-6}$, Kruskal-Wallis, corrección FDR), alfa ($p = 8,55 \cdot 10^{-6}$, Kruskal-Wallis, corrección FDR) y beta-1 ($p = 8,55 \cdot 10^{-6}$, Kruskal-Wallis, corrección FDR). Las diferencias entre pares de grupos resultan estadísticamente significativas entre controles y pacientes con DCL ($p = 2,12 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR), entre controles y pacientes con EA ($p = 3,77 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR), entre pacientes con DCL y pacientes con EA ($p = 0,0011$, prueba U de Mann-Whitney, corrección FDR) en la banda delta; entre controles y pacientes con DCL ($p = 2,21 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR), entre controles y pacientes con EA ($p = 3,77 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR), entre pacientes con DCL y pacientes con EA ($p = 9,73 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR) en la banda zeta; entre controles y pacientes con DCL ($p = 2,12 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR), entre controles y pacientes con EA ($p = 3,78 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR), entre pacientes con DCL y pacientes con EA ($p = 0,001$, prueba U de Mann-Whitney, corrección FDR) en la banda alfa; y entre controles y pacientes con DCL ($p = 2,03 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR), entre controles y pacientes con EA ($p = 3,77 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR), entre pacientes con DCL y pacientes con EA ($p = 9,74 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR) en la banda beta1. En todos los casos, los valores del parámetro se alejan del valor 1, que indica la presencia de dFC de origen neural.

5.8. Grado de antagonismo

El resultado de aplicar esta medida en ambas BBDD se muestra en la figura 5.6. En el caso de la base de datos del Hospital Universitario Río Hortega, los valores presentan diferencias globales estadísticamente significativas en las bandas de frecuencia delta ($p = 1,5 \cdot 10^{-20}$, Kruskal-Wallis, corrección FDR), zeta ($p = 8,25 \cdot 10^{-20}$, Kruskal-Wallis, corrección FDR) y alfa ($p = 1,35 \cdot 10^{-19}$, Kruskal-Wallis, corrección FDR). Las diferencias entre pares de grupos resultan estadísticamente significativas entre controles y pacientes con DCL ($p = 3,4 \cdot 10^{-18}$, prueba U de Mann-Whitney, corrección FDR), y entre controles y pacientes con EA ($p = 1,8 \cdot 10^{-16}$, prueba U de Mann-Whitney, corrección FDR) en la banda delta; entre controles y pacientes con DCL ($p = 3,84 \cdot 10^{-17}$, prueba U de Mann-Whitney, corrección FDR), y entre controles y pacientes con EA ($p = 5,88 \cdot 10^{-16}$, prueba U de Mann-Whitney, corrección FDR) en la banda zeta; y entre controles y pacientes con DCL ($p = 8,48 \cdot 10^{-17}$, prueba U de Mann-Whitney, corrección FDR), y entre controles y pacientes con EA ($p = 1,25 \cdot 10^{-15}$, prueba U de Mann-Whitney, corrección FDR) en la banda alfa. Los resultados de aplicar la medida en la BBDD del proyecto POCTEP no resultan estadísticamente significativos.

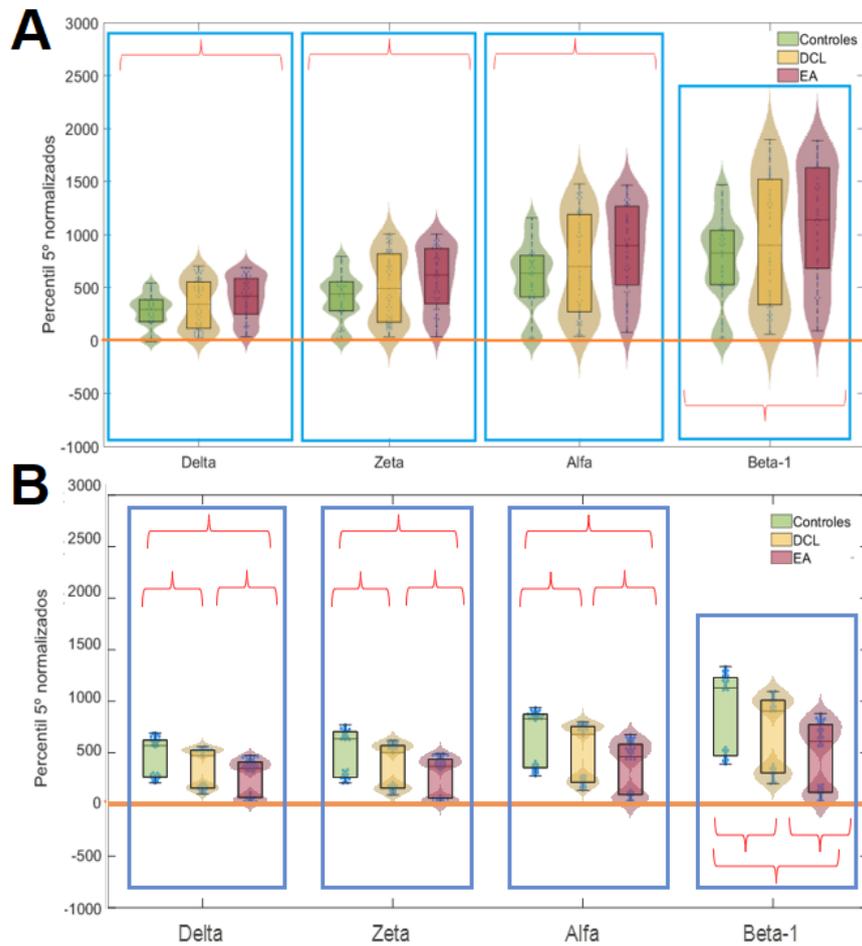


Figura 5.5. Resultado de la medida “percentil 5º”. Los cuadros azules indican resultados globales significativos ($p < 0,05$, Kruskal-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).

5.9. Dwell time

El resultado de aplicar la medida en la BBDD del Hospital Universitario Río Hortega presenta diferencias globales estadísticamente significativas en delta ($p = 0,02$, Kruskal-Wallis, corrección FDR), zeta ($p = 0,002$, Kruskal-Wallis, corrección FDR) y beta1 ($p = 0,02$, Kruskal-Wallis, corrección FDR). En este caso, se observan diferencias entre pares de grupos estadísticamente significativas entre controles y pacientes con DCL ($p = 0,01$, prueba U de Mann-Whitney, corrección FDR) en delta; entre controles y pacientes con DCL ($p = 0,0012$, prueba U de Mann-Whitney, corrección FDR), y entre controles y pacientes con EA ($p = 0,0012$, prueba U de Mann-Whitney, corrección FDR) en zeta; y entre controles y pacientes con DCL ($p = 0,032$, prueba U de Mann-Whitney, corrección FDR), y entre controles y pacientes con EA ($p = 0,011$, prueba U de Mann-Whitney, corrección FDR) en beta1. Los resultados de aplicar esta medida en la BBDD del proyecto POCTEP presentan diferencias globales estadísticamente significativas en zeta ($p = 0,04$, Kruskal-Wallis, corrección FDR) y en beta1 ($p = 0,04$, Kruskal-Wallis, corrección FDR). En este caso, se encuentran diferencias grupales estadísticamente significativas entre controles y pacientes con EA ($p = 0,001$, prueba U de Mann-Whitney, corrección FDR), y entre pacientes con DCL y pacientes con EA ($p = 0,04$, prueba U de Mann-Whitney, corrección FDR) en zeta; y entre contro-

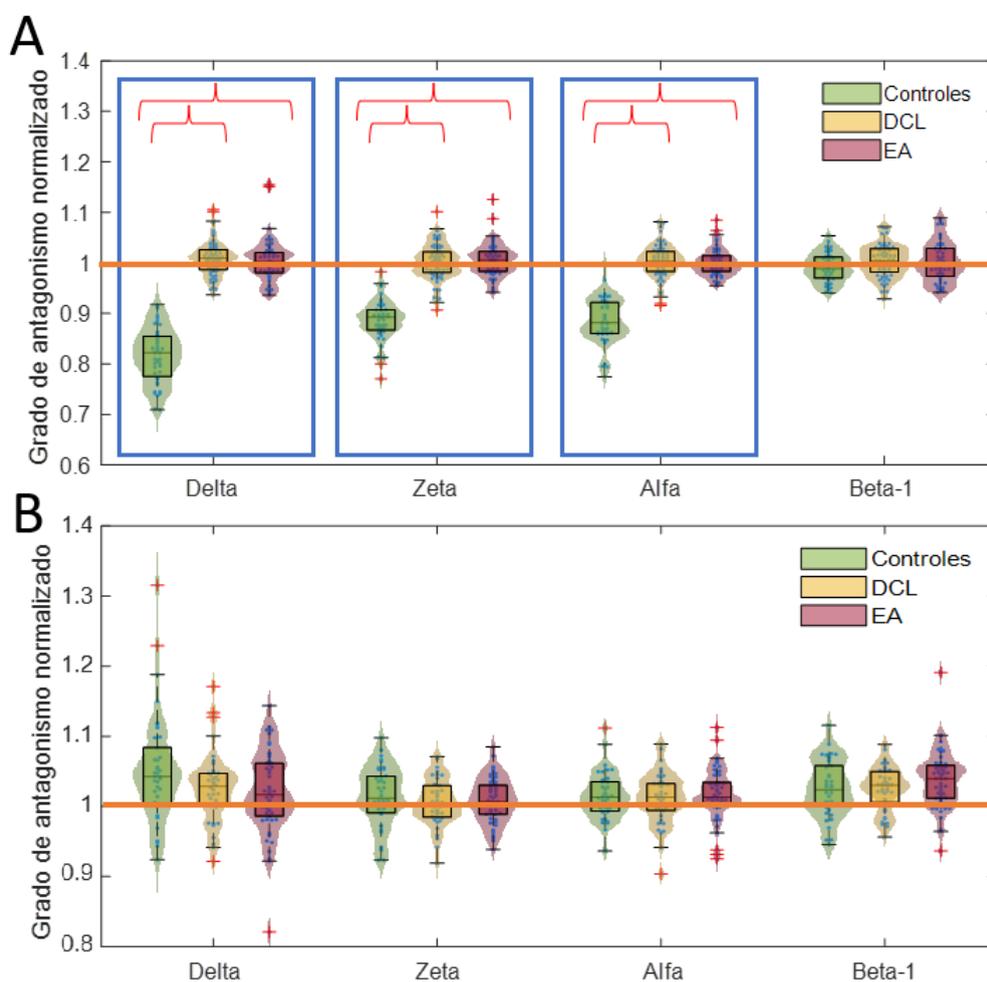


Figura 5.6. Resultado de la medida "grado de antagonismo". Los cuadros azules indican resultados globales significativos ($p < 0,05$, Kruskal-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).

les y pacientes con EA ($p = 0,03$, prueba U de Mann-Whitney, corrección FDR), y entre pacientes con DCL y pacientes con EA ($p = 0,03$, prueba U de Mann-Whitney, corrección FDR) en beta1. Los resultados se muestran en la figura 5.7.

5.10. *Leap size*

El resultado de aplicar esta medida en ambas BBDD se muestra en la figura 5.8. en la BBDD del Hospital Universitario Río Hortega, no existen diferencias estadísticamente significativas entre grupos. Sin embargo, en la BBDD del proyecto POCTEP se observan diferencias grupales estadísticamente significativas en delta ($p = 3,55 \cdot 10^{-10}$, Kruskal-Wallis, corrección FDR), zeta ($p = 1,8 \cdot 10^{-12}$, Kruskal-Wallis, corrección FDR), alfa ($p = 9,35 \cdot 10^{-14}$, Kruskal-Wallis, corrección FDR) y beta1 ($p = 1,38 \cdot 10^{-15}$, Kruskal-Wallis, corrección FDR). Las diferencias entre pares de grupos resultan estadísticamente significativas entre controles y pacientes con EA ($p = 2,46 \cdot 10^{-7}$, prueba U de Mann-Whitney, corrección FDR), y entre pacientes con DCL y pacientes con EA ($p = 1,2 \cdot 10^{-8}$, prueba U de Mann-Whitney, corrección FDR) en delta; entre controles y pacientes con EA ($p = 5,41 \cdot 10^{-9}$, prueba U de Mann-Whitney, corrección FDR), y entre pacientes con DCL

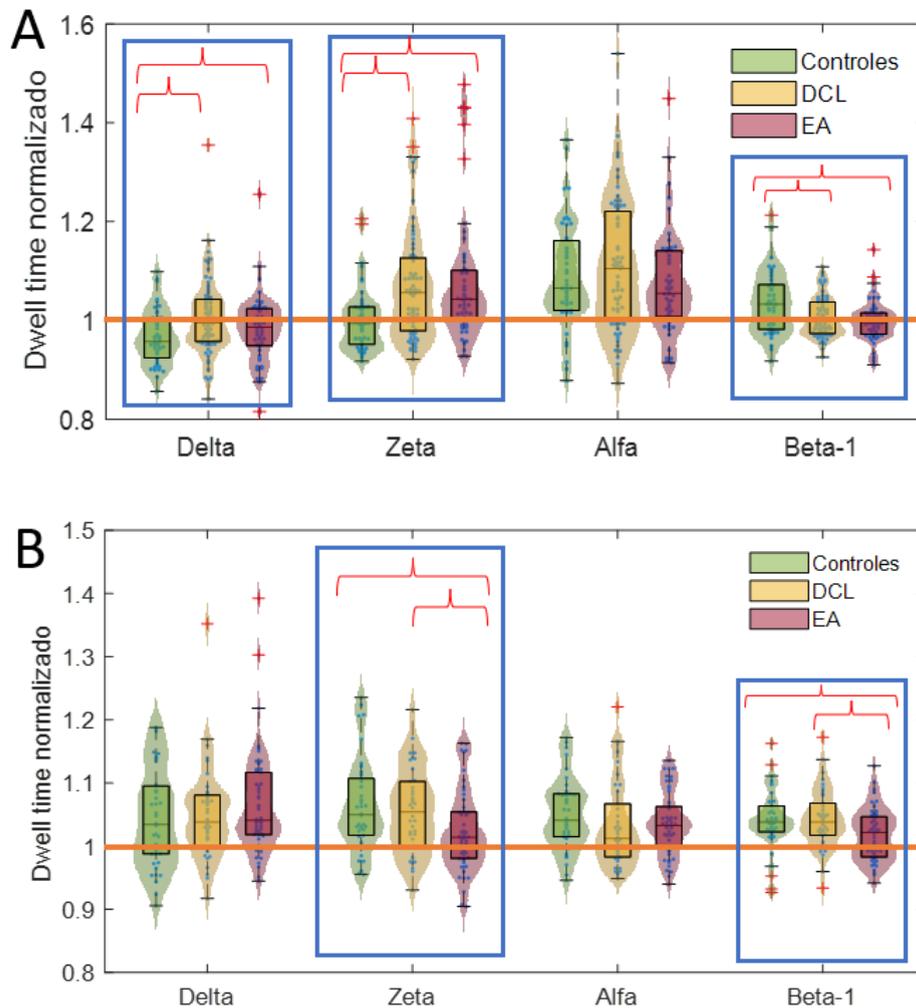


Figura 5.7. Resultado de la medida *dwell time*. Los cuadros azules indican resultados globales significativos ($p < 0,05$, Kruskal-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).

y pacientes con EA ($p = 2,04 \cdot 10^{-10}$, prueba U de Mann-Whitney, corrección FDR) en zeta; entre controles y pacientes con EA ($p = 6,58 \cdot 10^{-12}$, prueba U de Mann-Whitney, corrección FDR), y entre pacientes con DCL y pacientes con EA ($p = 2,372 \cdot 10^{-9}$, prueba U de Mann-Whitney, corrección FDR) en alfa; y entre controles y pacientes con EA ($p = 9,23 \cdot 10^{-13}$, prueba U de Mann-Whitney, corrección FDR), y entre pacientes con DCL y pacientes con EA ($p = 5,32 \cdot 10^{-12}$, prueba U de Mann-Whitney, corrección FDR) en beta1.

5.11. Complejidad de TAS

El resultado de aplicar esta medida en ambas BBDD se muestra en la figura 5.9. Los resultados en la BBDD del Hospital Universitario Río Hortega presentan diferencias grupales estadísticamente significativas en las bandas delta ($p = 2,04 \cdot 10^{-4}$, Kruskal-Wallis, corrección FDR) y beta1 ($p = 0,04$, Kruskal-Wallis, corrección FDR). Se encuentran diferencias entre pares de grupos significativas entre controles y pacientes con EA ($p = 4,19 \cdot 10^{-4}$, prueba U de Mann-Whitney, corrección FDR), y entre controles y pacientes con DCL ($p = 1,15 \cdot 10^{-4}$, prueba U de Mann-Whitney, corre-

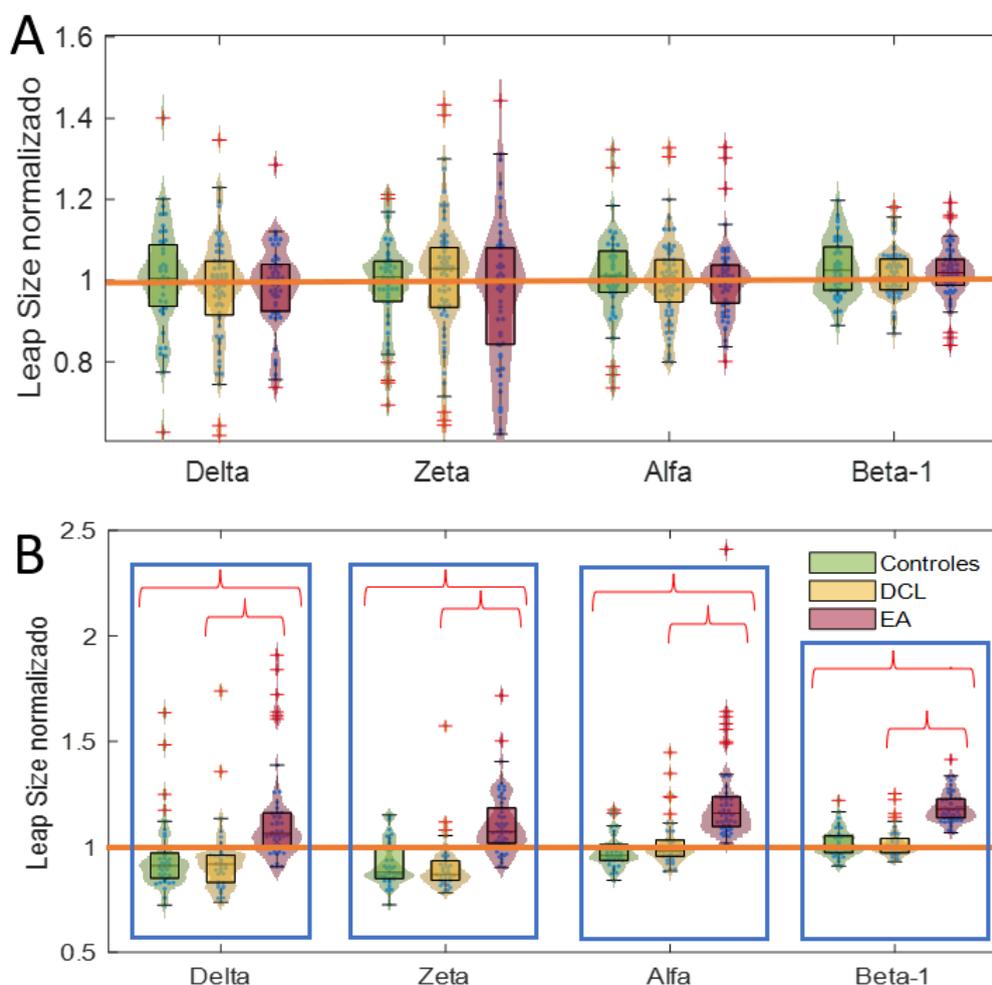


Figura 5.8. Resultado de la medida *leap size*. Los cuadros azules indican resultados globales significativos ($p < 0,05$, Kruskal-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).

ción FDR) en delta; y entre controles y pacientes con EA ($p = 0,046$, prueba U de Mann-Whitney, corrección FDR) en beta1. En el caso de la base de datos del proyecto POCTEP, se encuentran diferencias globales estadísticamente significativas en beta1 ($p = 0,0028$, Kruskal-Wallis, corrección FDR). En este caso, las diferencias entre pares de grupos resultan estadísticamente significativas entre controles y pacientes con EA ($p = 0,0024$, prueba U de Mann-Whitney, corrección FDR) y entre pacientes con DCL y pacientes con EA ($p = 0,0024$, prueba U de Mann-Whitney, corrección FDR). En ambas BBDD, se observa cómo los datos se concentran en torno al valor uno en el caso de pacientes con EA en las bandas de frecuencia donde se encontraron diferencias estadísticamente significativas. Por tanto, se aprecia la pérdida de dFC de origen neural conforme avanza la EA.

5.12. Clasificación

Después de obtener los resultados derivados de las diferentes métricas cronectómicas, se implementaron varios algoritmos de selección de características y modelos de clasificación. Específicamente, se emplearon los algoritmos de selección K-mejores, FCBF, SFS y RFE. A continuación, se presentan los resultados obtenidos al aplicar estos métodos a ambas BBDD. Para ilustrar la efi-

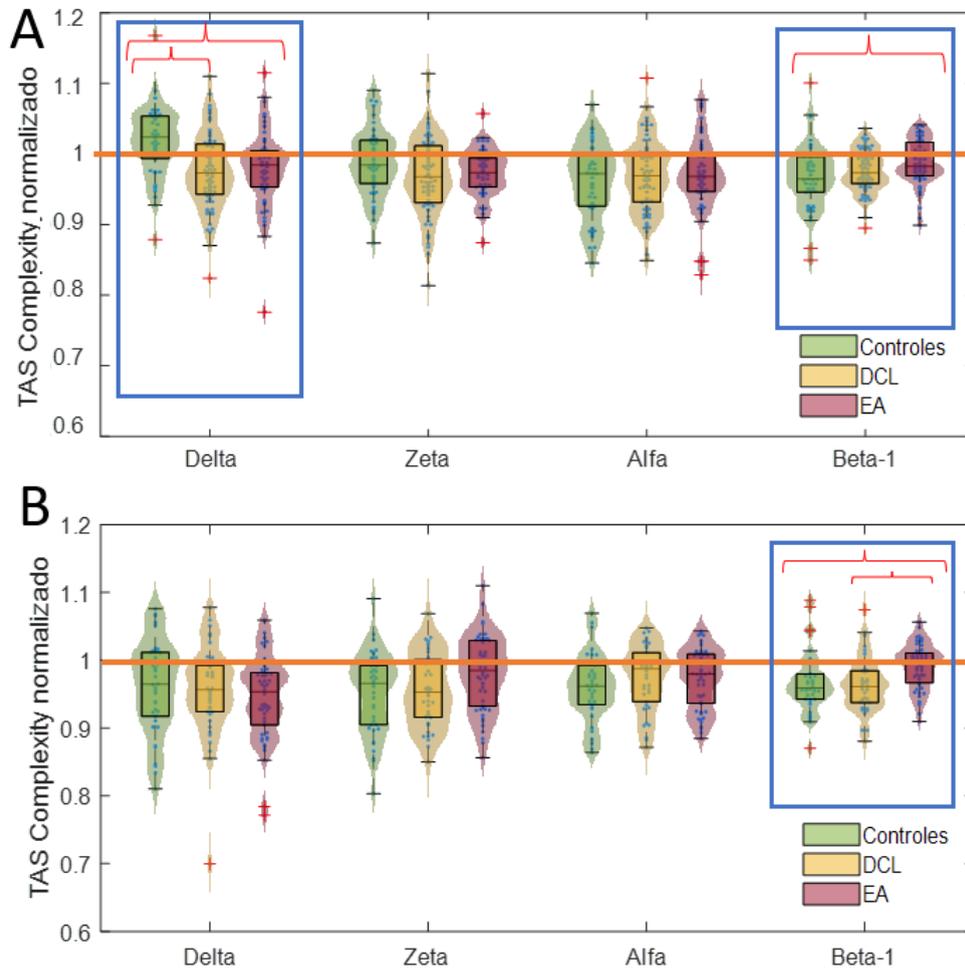


Figura 5.9. Resultado de la medida “complejidad de TAS”. Los cuadros azules indican resultados globales significativos ($p < 0,05$, Kruskal-Wallis, corrección FDR). Las llaves rojas indican diferencias estadísticamente significativas entre pares de grupos ($p < 0,05$, prueba U de Mann-Whitney, corrección FDR).

En la comparación de los modelos de clasificación, se exhibirán en cada caso la comparación de la precisión obtenida en cada modelo y la matriz de confusión y el valor de $kappa$ del mejor de los modelos expuestos. En la tabla 5.21 se muestra un resumen de los resultados obtenidos en cada caso.

5.12.1. K-mejores

El resultado de aplicar el algoritmo a las BBDD por separado se muestra en la figura 5.10. Cuanto más cercano sea el valor a uno, más importante se considera la medida. En el caso de la BBDD del Hospital Universitario Río Hortega, la característica más importante de forma individual es el grado de antagonismo en las bandas delta, zeta y alfa. Sin embargo, en la BBDD del proyecto POCTEP la característica más importante es el *leap size*. El resultado de aplicar el algoritmo a las dos bases de datos de forma conjunta se muestra en la figura 5.11. La importancia se mide como el logaritmo de los p-valores calculados a partir de la relación entre cada característica y la variable objetivo (Venkatesh & Anuradha, 2019). Un valor de importancia cercano a uno indicaría que una característica tiene buena capacidad de clasificación. En este caso, las medidas más importantes son el área del meta-estado dominante en la banda delta, el grado de antagonismo en delta, zeta y alfa, y el *leap size*. Este algoritmo de selección de características determina

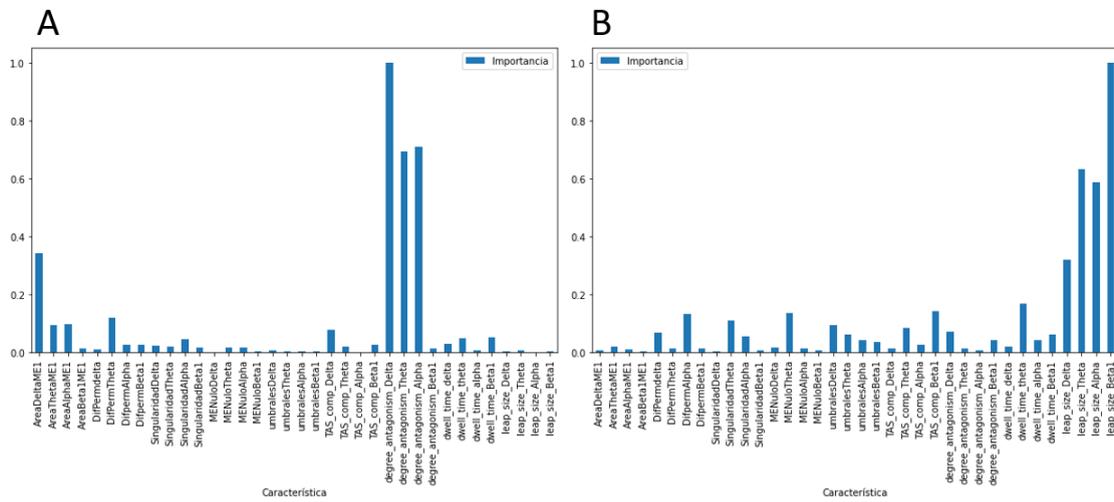


Figura 5.10. Importancia asignada a cada parámetro por el algoritmo de K-mejores en las BBDD del Hospital Universitario Río Hortega (A) y POCTEP (B).

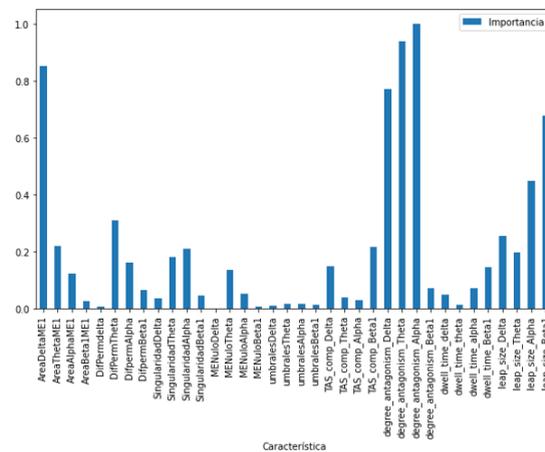


Figura 5.11. Importancia asignada a cada parámetro por el algoritmo de K-mejores en ambas BBDD.

la capacidad de cada característica de forma individual para clasificar (Jain & Saha, 2022). Al no indicar importancia en las mismas características para ambas BBDD no se utilizó este método para clasificación.

5.12.2. FCBF

En la figura 5.12 se muestra el resultado de aplicar el algoritmo a las BBDD por separado. En ambos casos se asigna un valor de correlación más bajo a la medida “percentil 5” obteniendo el resto de medidas un valor de correlación similar. En la figura 5.13 se muestra el resultado de aplicar el algoritmo a ambas BBDD de forma conjunta. Se observa que la medida “percentil 5” es la única que presenta valores de correlación más bajos.

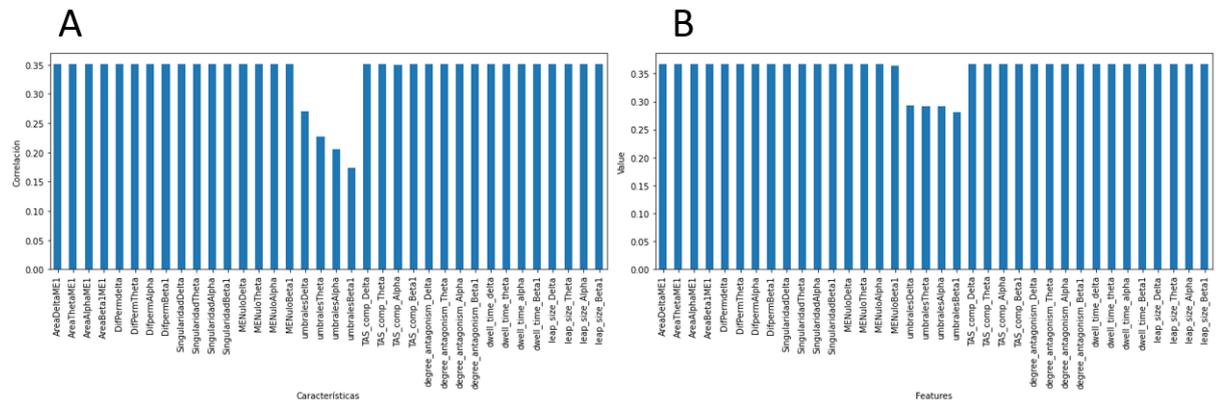


Figura 5.12. Correlación asignada a cada parámetro por el algoritmo FCBF en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B).

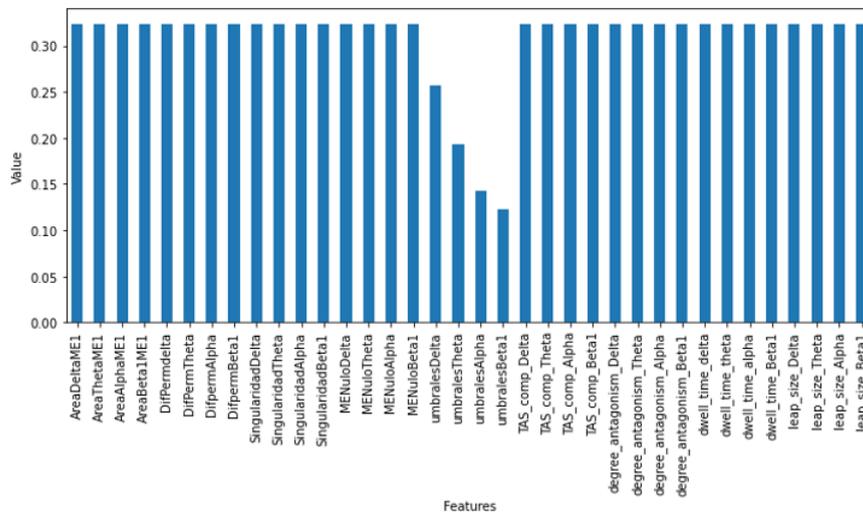


Figura 5.13. Correlación asignada a cada parámetro por el algoritmo FCBF en ambas BBDD.

5.12.2.1. Clasificación con 3 clases

La precisión de los modelos de clasificación con el uso de las características seleccionadas por FCBF para cada base de datos por separado se muestran en la figura 5.14. En el caso de la BBDD del Hospital Universitario Río Hortega el modelo de mejor rendimiento es el árbol de decisión. La matriz de confusión se muestra en la tabla 5.3. El valor *kappa* obtenido es 0,74, se obtiene una precisión del 82,5 %, una especificidad del 65 % y una sensibilidad del 100 %. En el caso de la BBDD del proyecto POCTEP el mejor es *Random Forest*. La matriz de confusión se muestra en la tabla 5.4. El valor *kappa* obtenido es 0,81, se obtiene una precisión del 87,5 %, una especificidad del 72,72 % y una sensibilidad del 95,23 %.

La precisión de los modelos de clasificación con el uso de las características seleccionadas por FCBF para la combinación de las BBDD se muestran en la figura 5.15. El modelo de mejor rendimiento es la red neuronal MLP. La matriz de confusión se muestra en la tabla 5.5. El valor *kappa* obtenido es 0,63, se obtiene una precisión del 75 %, una especificidad del 41,93 % y una sensibilidad del 100 %.

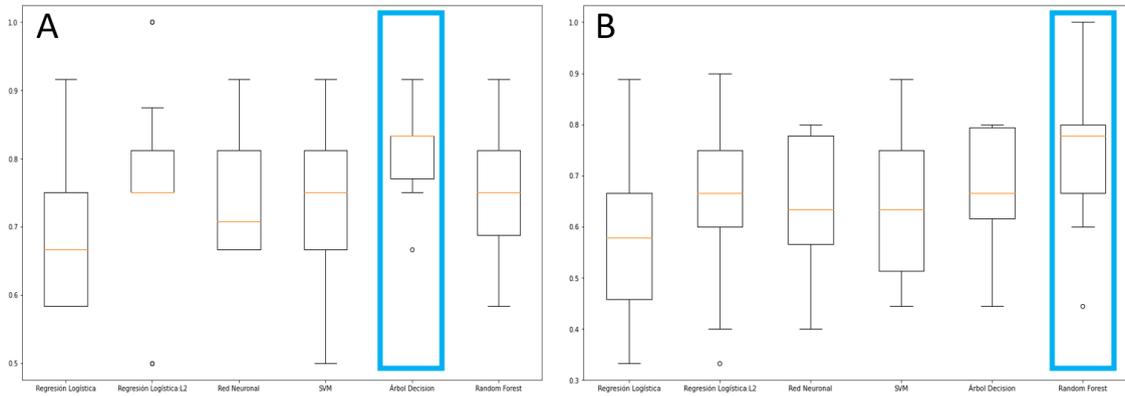


Figura 5.14. Precisión obtenida en la clasificación con tres clases (control, DCL y EA) en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B) con características seleccionadas por FCBF. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.

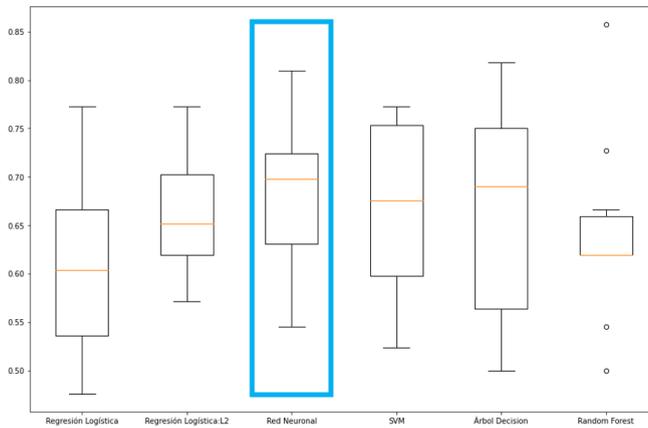


Figura 5.15. Precisión obtenida en la clasificación con tres clases (control, DCL y EA) en las BBDD de forma conjunta con características seleccionadas por FCBF. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.

Observado\Estimado	Control	DCL	EA
Control	13	0	0
DCL	0	10	7
EA	0	0	10

Tabla 5.3. Matriz de confusión utilizando el método de selección de características FCBF en la BBDD del Hospital Universitario Río Hortega.

Observado\Estimado	Control	DCL	EA
Control	8	3	0
DCL	1	9	0
EA	0	0	11

Tabla 5.4. Matriz de confusión utilizando el método de selección de características FCBF en la BBDD del proyecto POCTEP.

Observado\Estimado	Control	DCL	EA
Control	13	7	2
DCL	0	20	1
EA	0	8	21

Tabla 5.5. Matriz de confusión utilizando el método de selección de características FCBF en la combinación de ambas BBDD.

5.12.2.2. Clasificación con 2 clases

La precisión de los modelos de clasificación con el uso de las características seleccionadas por FCBF para cada base de datos por separado se muestran en la figura 5.16. En el caso de la BBDD del Hospital Universitario Río Hortega el modelo de mejor rendimiento es el árbol de decisión. La matriz de confusión se muestra en la tabla 5.6. El valor *kappa* obtenido es 1, se obtiene una precisión del 100 %, una especificidad del 100 % y una sensibilidad del 100 %. En el caso de la BBDD del proyecto POCTEP el mejor es el árbol de decisión. La matriz de confusión se muestra en la tabla 5.7. El valor *kappa* obtenido es 0,7, se obtiene una precisión del 87,5 %, una especificidad del 63,63 % y una sensibilidad del 100 %.

La precisión de los modelos de clasificación con el uso de las características seleccionadas por FCBF para la combinación de las BBDD se muestran en la figura 5.17. El modelo de mejor rendimiento es el árbol de decisión. La matriz de confusión se muestra en la tabla 5.8. El valor *kappa* obtenido es 0,44, se obtiene una precisión del 80,5 %, una especificidad del 36,36 % y una sensibilidad del 100 %.

Observado\Estimado	Control	Demencia
Control	13	0
Demencia	0	27

Tabla 5.6. Matriz de confusión utilizando el método de selección de características FCBF en la BBDD del Hospital Universitario Río Hortega.

Observado\Estimado	Control	Demencia
Control	7	4
Demencia	0	21

Tabla 5.7. Matriz de confusión utilizando el método de selección de características FCBF en la BBDD del proyecto POCTEP.

Observado\Estimado	Control	Demencia
Control	8	14
Demencia	0	50

Tabla 5.8. Matriz de confusión utilizando el método de selección de características FCBF en la combinación de ambas BBDD.

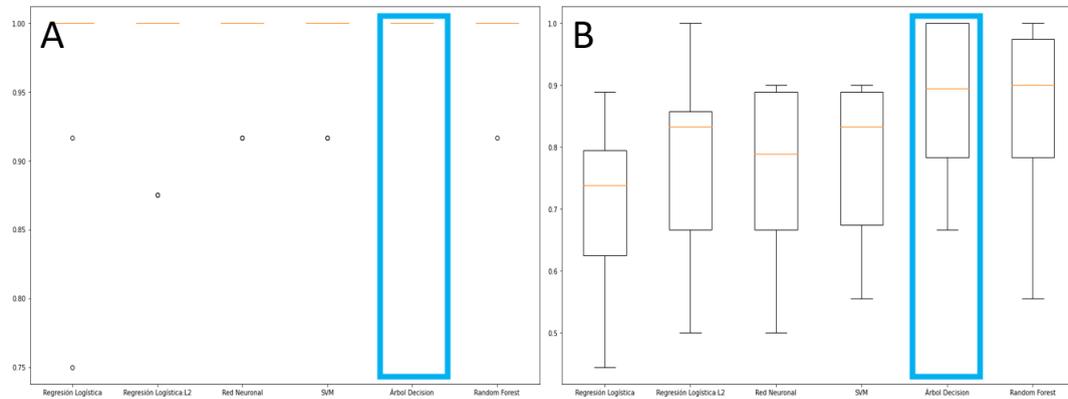


Figura 5.16. Precisión obtenida en la clasificación con dos clases (control y demencia) en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B) con características seleccionadas por FCBF. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.

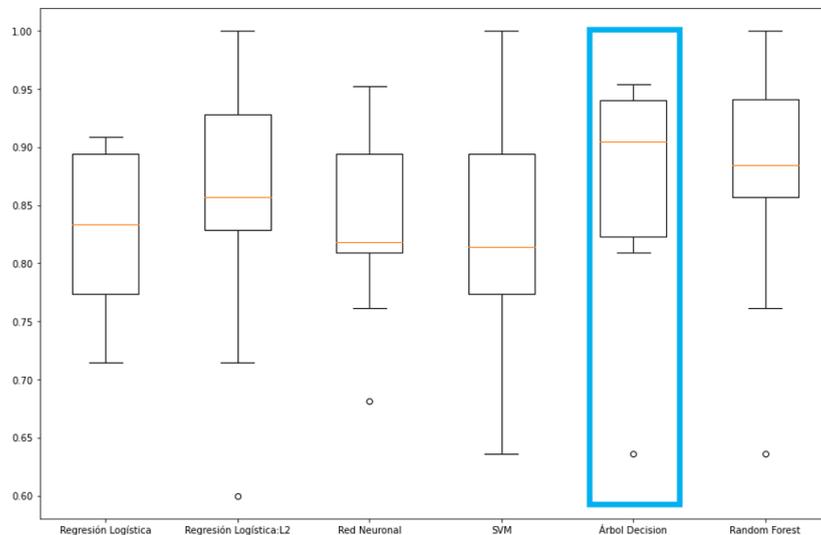


Figura 5.17. Precisión obtenida en la clasificación con dos clases (control y Demencia) en las BBDD de forma conjunta con características seleccionadas por FCBF. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.

5.12.3. SFS

En la figura 5.18 se muestran las características seleccionadas tras aplicar el algoritmo a las BBDD por separado. En ambas BBDD el número de características seleccionadas es 12, siendo el mejor conjunto de medidas distinto en cada base de datos. Sin embargo, el número de características óptimo en la combinación de ambas BBDD es 27 características. En la figura 5.19 se muestra el resultado de aplicar el algoritmo a ambas BBDD de forma conjunta.

5.12.3.1. Clasificación con 3 clases

La precisión de los modelos de clasificación con el uso de las características seleccionadas por SFS para cada base de datos por separado se muestran en la figura 5.20. En el caso de la BBDD del Hospital Universitario Río Hortega el modelo de mejor rendimiento es el árbol de decisión. La matriz de confusión se muestra en la tabla 5.9. El valor $kappa$ obtenido es 0,74, se obtiene una

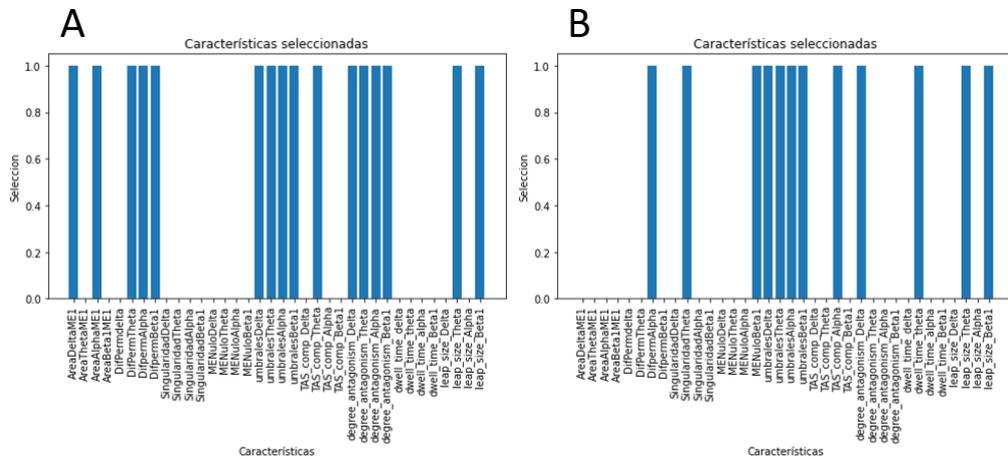


Figura 5.18. Características seleccionadas por el algoritmo SFS en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B).

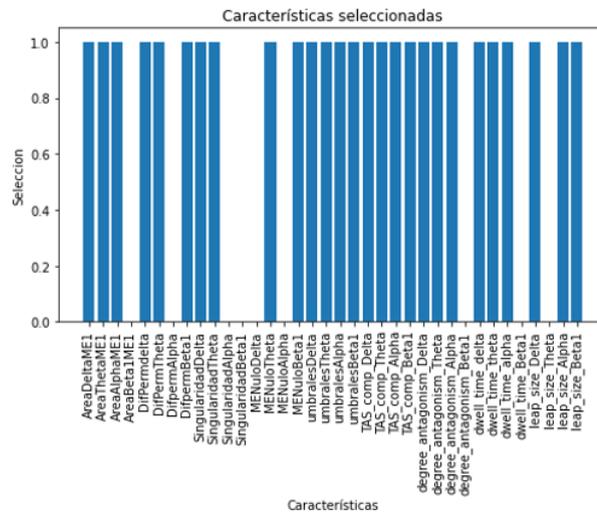


Figura 5.19. Características seleccionadas por el algoritmo SFS en ambas BBDD.

precisión del 82,5 %, una especificidad del 65 % y una sensibilidad del 100 %. En el caso de la BBDD del proyecto POCTEP el mejor es el *random forest*. La matriz de confusión se muestra en la tabla 5.10. El valor *kappa* obtenido es 0,86, se obtiene una precisión del 90,62 %, una especificidad del 75 % y una sensibilidad del 100 %.

La precisión de los modelos de clasificación con el uso de las características seleccionadas por SFS para la combinación de las BBDD se muestran en la figura 5.21. El modelo de mejor rendimiento es la regresión logística con penalización L2. La matriz de confusión se muestra en la tabla 5.11. El valor *kappa* obtenido es 0,61, se obtiene una precisión del 73,61 %, una especificidad del 51,61 % y una sensibilidad del 90,24 %.

5.12.3.2. Clasificación con 2 clases

La precisión de los modelos de clasificación con el uso de las características seleccionadas por SFS para cada base de datos por separado se muestran en la figura 5.22. En el caso de la BBDD

Observado\Estimado	Control	DCL	EA
Control	13	0	0
DCL	0	10	7
EA	0	0	10

Tabla 5.9. Matriz de confusión utilizando el método de selección de características SFS en la BBDD del Hospital Universitario Río Hortega.

Observado\Estimado	Control	DCL	EA
Control	9	2	0
DCL	0	9	1
EA	0	0	11

Tabla 5.10. Matriz de confusión utilizando el método de selección de características SFS en la BBDD del proyecto POCTEP.

Observado\Estimado	Control	DCL	EA
Control	16	4	2
DCL	2	18	1
EA	2	8	19

Tabla 5.11. Matriz de confusión utilizando el método de selección de características SFS en la combinación de ambas BBDD.

del Hospital Universitario Río Hortega el modelo de mejor rendimiento es el árbol de decisión. La matriz de confusión se muestra en la tabla 5.12. El valor *kappa* obtenido es 1, se obtiene una precisión del 100 %, una especificidad del 100 % y una sensibilidad del 100 %. En el caso de la BBDD del proyecto POCTEP el mejor es el árbol de decisión. La matriz de confusión se muestra en la tabla 5.13. El valor *kappa* obtenido es 0,7, se obtiene una precisión del 87,5 %, una especificidad del 63,63 % y una sensibilidad del 100 %.

La precisión de los modelos de clasificación con el uso de las características seleccionadas por SFS para la combinación de las BBDD se muestran en la figura 5.23. El modelo de mejor rendimiento es el árbol de decisión. La matriz de confusión se muestra en la tabla 5.14. El valor

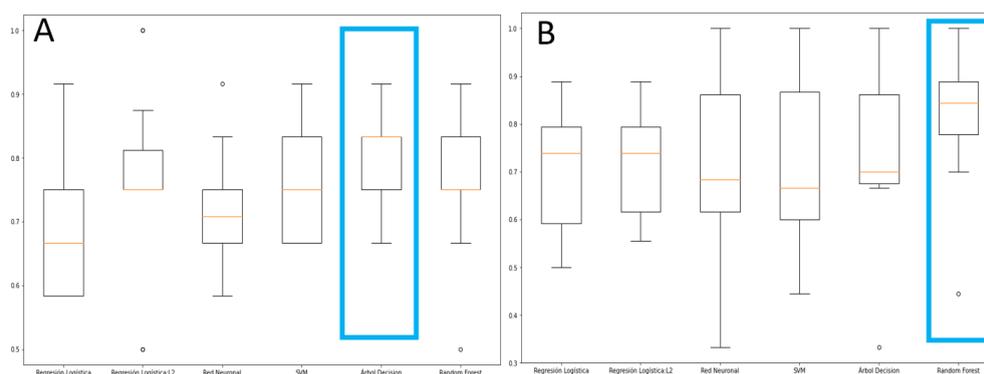


Figura 5.20. Precisión obtenida en la clasificación con tres clases (control, DCL y EA) en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B) con características seleccionadas por SFS. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.

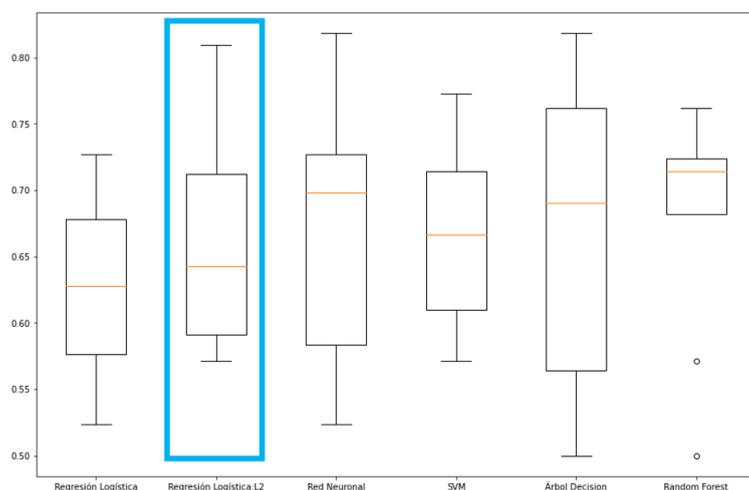


Figura 5.21. Precisión obtenida en la clasificación con tres clases (control, DCL y EA) en las BBDD de forma conjunta con características seleccionadas por SFS. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.

κ obtenido es 0,58, se obtiene una precisión del 84,72 %, una especificidad del 50 % y una sensibilidad del 100 %.

Observado\Estimado	Control	Demencia
Control	13	0
Demencia	0	27

Tabla 5.12. Matriz de confusión utilizando el método de selección de características SFS en la BBDD del Hospital Universitario Río Hortega.

Observado\Estimado	Control	Demencia
Control	7	4
Demencia	0	21

Tabla 5.13. Matriz de confusión utilizando el método de selección de características SFS en la BBDD del proyecto del proyecto POCTEP.

Observado\Estimado	Control	Demencia
Control	11	11
Demencia	0	50

Tabla 5.14. Matriz de confusión utilizando el método de selección de características SFS en la combinación de ambas BBDD.

5.12.4. RFE

Este algoritmo mantiene un número de características a seleccionar igual al expuesto en el método SFS. En la figura 5.24 se muestran las características seleccionadas en cada base de datos por separado. En la figura 5.25 se muestran las medidas seleccionadas cuando se combinan ambas BBDD.

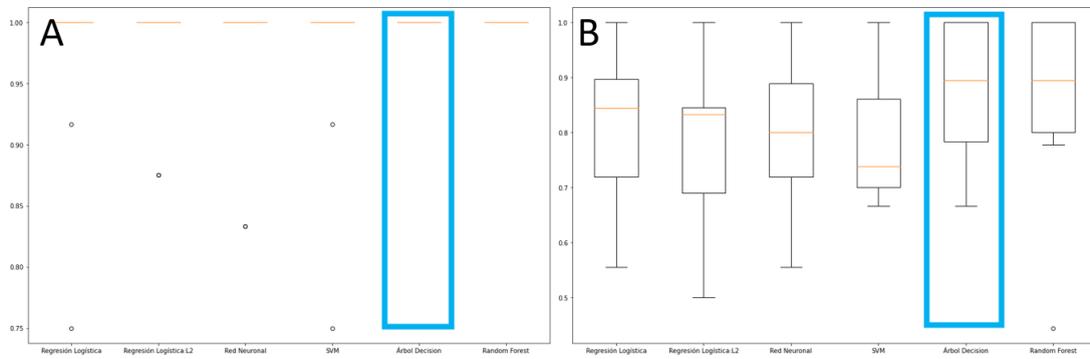


Figura 5.22. Precisión obtenida en la clasificación con dos clases (control y demencia) en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B) con características seleccionadas por SFS. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.

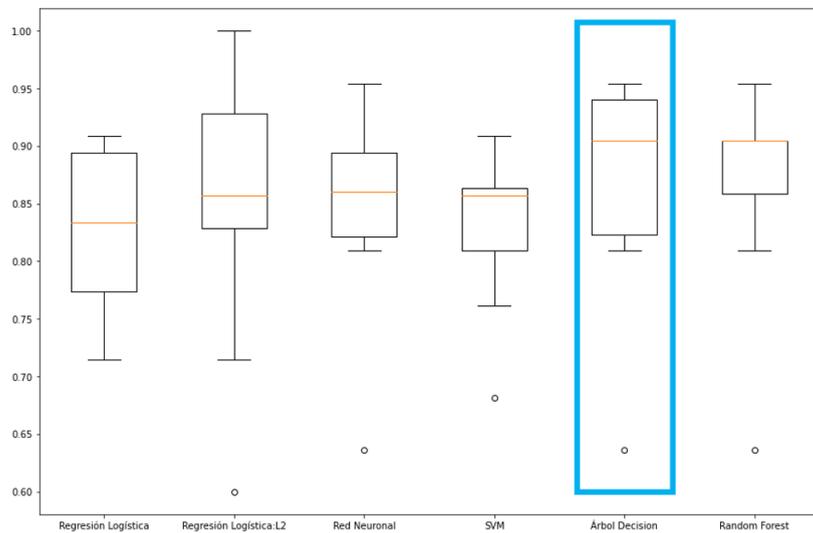


Figura 5.23. Precisión obtenida en la clasificación con dos clases (control y demencia) en las BBDD de forma conjunta con características seleccionadas por SFS. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.

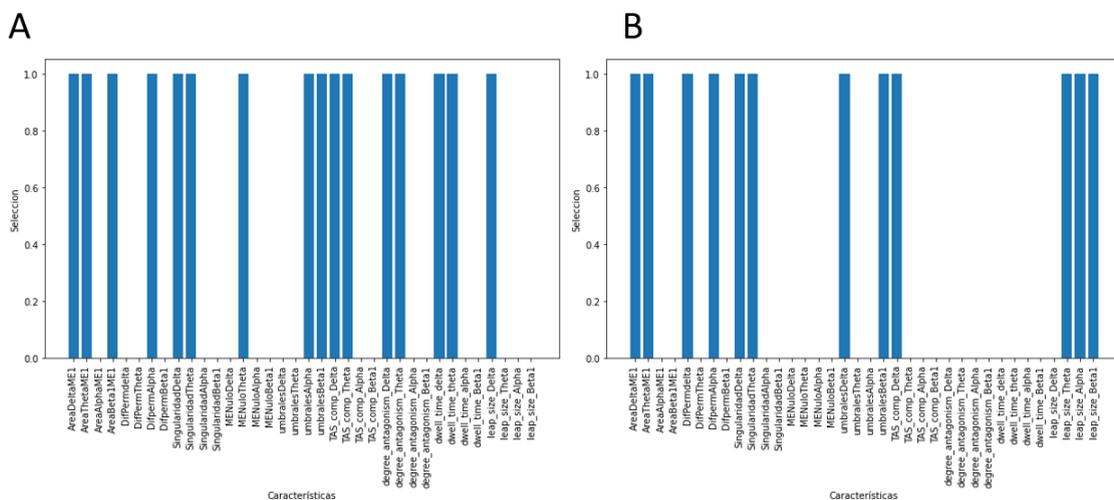


Figura 5.24. Características seleccionadas por el algoritmo RFE en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B).

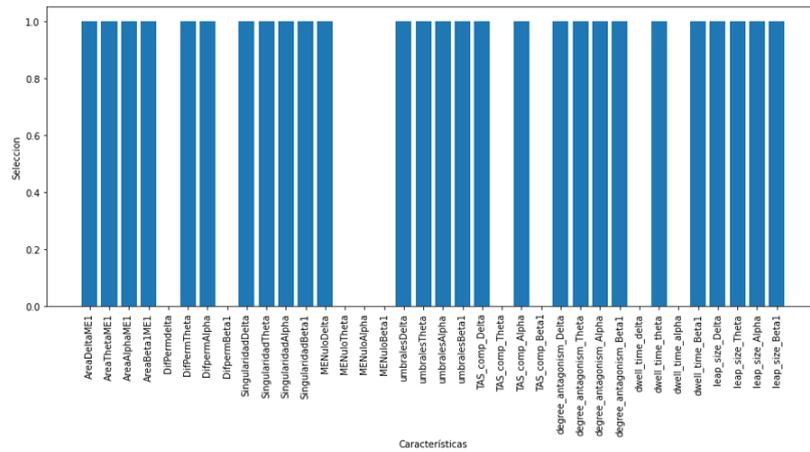


Figura 5.25. Características seleccionadas por el algoritmo RFE en ambas BBDD.

5.12.4.1. Clasificación con 3 clases

La precisión de los modelos de clasificación con el uso de las características seleccionadas por RFE para cada base de datos por separado se muestran en la figura 5.26. En el caso de la BBDD del Hospital Universitario Río Hortega el modelo de mejor rendimiento es el árbol de decisión. La matriz de confusión se muestra en la tabla 5.15. El valor *kappa* obtenido es 0,81, se obtiene una precisión del 87,5 %, una especificidad del 72,22 % y una sensibilidad del 100 %. En el caso de la BBDD del proyecto POCTEP el mejor es el *random forest*. La matriz de confusión se muestra en la tabla 5.16. El valor *kappa* obtenido es 0,81, se obtiene una precisión del 87,5 %, una especificidad del 81,81 % y una sensibilidad del 90,47 %.

La precisión de los modelos de clasificación con el uso de las características seleccionadas por RFE para la combinación de las BBDD se muestran en la figura 5.27. El modelo de mejor rendimiento es la red neuronal. La matriz de confusión se muestra en la tabla 5.17. El valor *kappa* obtenido es 0,6, se obtiene una precisión del 73,61 %, una especificidad del 44,82 % y una sensibilidad del 93,02 %.

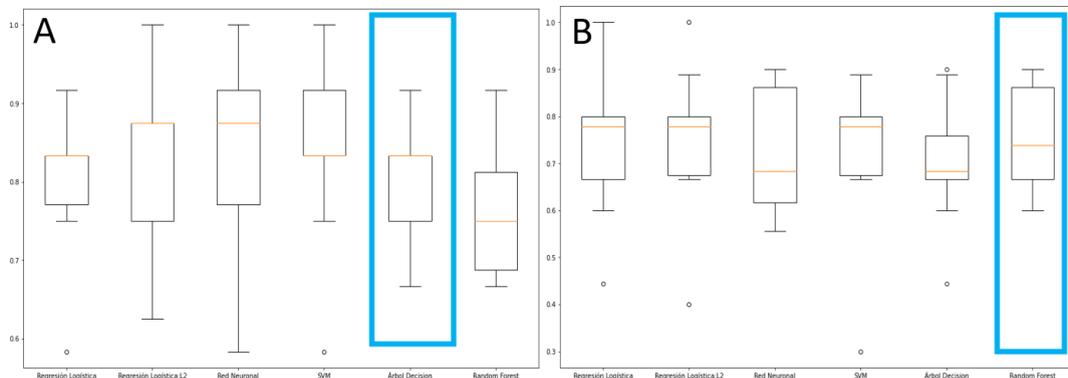


Figura 5.26. Precisión obtenida en la clasificación con tres clases (control, DCL y EA) en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B) con características seleccionadas por RFE. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.

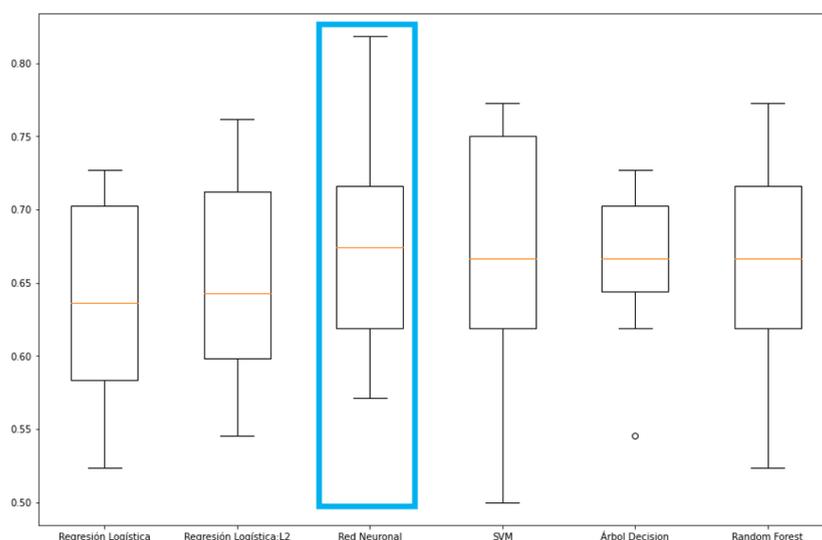


Figura 5.27. Precisión obtenida en la clasificación con tres clases (control, DCL y EA) en las BBDD de forma conjunta con características seleccionadas por RFE. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.

Observado\Estimado	Control	DCL	EA
Control	13	0	0
DCL	0	12	5
EA	0	0	10

Tabla 5.15. Matriz de confusión utilizando el método de selección de características RFE en la BBDD del Hospital Universitario Río Hortega.

Observado\Estimado	Control	DCL	EA
Control	9	2	0
DCL	2	8	0
EA	0	0	11

Tabla 5.16. Matriz de confusión utilizando el método de selección de características RFE en la BBDD del proyecto POCTEP.

Observado\Estimado	Control	DCL	EA
Control	13	6	3
DCL	2	18	1
EA	1	6	22

Tabla 5.17. Matriz de confusión utilizando el método de selección de características RFE en la combinación de ambas BBDD.

5.12.4.2. Clasificación con 2 clases

La precisión de los modelos de clasificación con el uso de las características seleccionadas por RFE para cada base de datos por separado se muestran en la figura 5.28. En el caso de la BBDD del Hospital Universitario Río Hortega el modelo de mejor rendimiento es el árbol de decisión. La matriz de confusión se muestra en la tabla 5.18. El valor $kappa$ obtenido es 1, se obtiene una

precisión del 100 %, una especificidad del 100 % y una sensibilidad del 100 %. En el caso de la BBDD del proyecto POCTEP el mejor es el árbol de decisión. La matriz de confusión se muestra en la tabla 5.19. El valor *kappa* obtenido es 0,69, se obtiene una precisión del 87,5 %, una especificidad del 63,63 % y una sensibilidad del 100 %.

La precisión de los modelos de clasificación con el uso de las características seleccionadas por RFE para la combinación de las BBDD se muestran en la figura 5.29. El modelo de mejor rendimiento es el árbol de decisión. La matriz de confusión se muestra en la tabla 5.20. El valor *kappa* obtenido es 0,68, se obtiene una precisión del 87,5 %, una especificidad del 68,18 % y una sensibilidad del 96 %.

Observado\Estimado	Control	Demencia
Control	13	0
Demencia	0	27

Tabla 5.18. Matriz de confusión utilizando el método de selección de características SFS en la BBDD del Hospital Universitario Río Hortega.

Observado\Estimado	Control	Demencia
Control	7	4
Demencia	0	21

Tabla 5.19. Matriz de confusión utilizando el método de selección de características SFS en la BBDD del proyecto del proyecto POCTEP.

Observado\Estimado	Control	Demencia
Control	15	7
Demencia	2	48

Tabla 5.20. Matriz de confusión utilizando el método de selección de características RFE en la combinación de ambas BBDD.

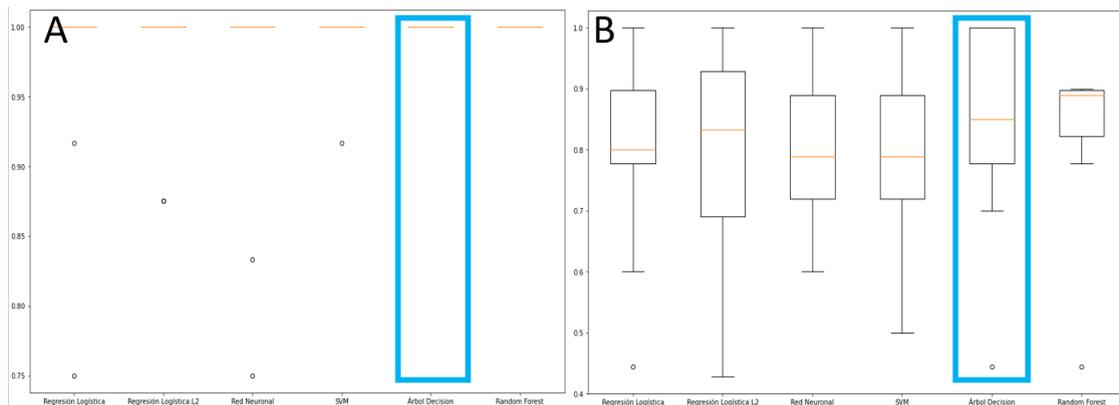


Figura 5.28. Precisión obtenida en la clasificación con dos clases (control y demencia) en las BBDD del Hospital Universitario Río Hortega (A) y del proyecto POCTEP (B) con características seleccionadas por RFE. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.

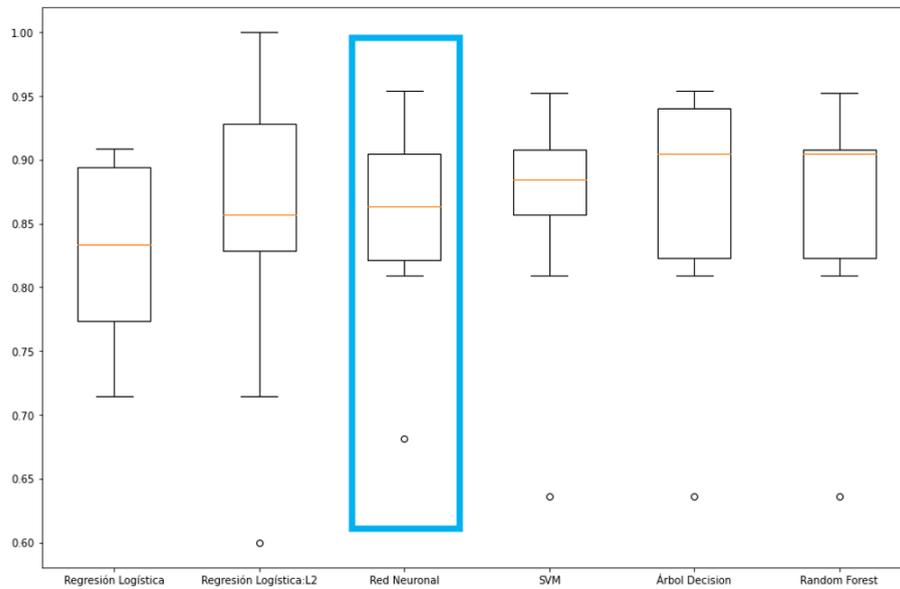


Figura 5.29. Precisión obtenida en la clasificación con dos clases (control y demencia) en las BBDD de forma conjunta con características seleccionadas por RFE. El rectángulo azul indica el modelo con mejor rendimiento en cada caso.

Valor $kappa$ obtenido						
Base de datos	3 clases			2 clases		
	FCBF	SFS	RFE	FCBF	SFS	RFE
Río Hortega	0.74 (AD)	0.74 (AD)	0.81 (AD)	1 (AD)	1 (AD)	1 (AD)
POCTEP	0.81 (RF)	0.81 (RF)	0.81 (RF)	0.7 (AD)	0.7 (AD)	0.69 (AD)
Conjunto	0.63 (RN)	0.61 (RLL2)	0.6 (RN)	0.44 (AD)	0.58 (AD)	0.68 (RN)

Tabla 5.21. Valor $kappa$ obtenido en la clasificación para diferentes métodos de selección de características y diferentes combinaciones de las BBDD. AD: Árbol de Decisión; RF: *Random Forest*; RN: Red Neuronal; RLL2: Regresión Logística con penalización $L2$.

Capítulo 6

Discusión

6.1. Introducción	54
6.2. Optimización	54
6.3. Medidas cronectómicas	54
6.4. Clasificación	57
6.5. Limitaciones y líneas futuras	59

6.1. Introducción

En este TFM se optimizó el proceso de subrogación utilizado para descartar que las medidas cronectómicas implementadas fueran sensibles a fluctuaciones de origen no neural. Tras calcular las medidas, se elaboró una comparación entre diversos métodos de selección de características y modelos de clasificación. El mejor modelo de clasificación se utilizó en el conjunto de validación para medir la calidad en la clasificación.

6.2. Optimización

El primero de los métodos propuestos para optimizar el proceso es el uso de ficheros ejecutables. En ninguna etapa del proceso se aprecia una reducción del tiempo de ejecución. Para la ejecución de ficheros ejecutables del sistema es necesario codificar los datos procedentes de MATLAB en un formato compatible con el sistema operativo (MathWorks-Inc., 2023). Esto implica una comunicación con el sistema operativo para el intercambio de información, manejo de la memoria y cambios de contexto que aumentan de forma significativa el tiempo de ejecución del programa. El uso de ficheros en C/C++ reduce el tiempo de ejecución en todos los casos apoyando la hipótesis de Dmitrović (2023), que demuestra la superioridad de la eficiencia en lenguajes compilados respecto a lenguajes interpretados. Sin embargo, la ejecución en GPU no reduce el tiempo de procesado en todos los casos. Los resultados de optimización sugieren que la GPU no resulta eficiente cuando es necesaria una comunicación frecuente entre CPU y GPU de acuerdo con el estudio realizado por Kang et al. (2020), que indica que el tamaño de matriz que una GPU puede manejar de forma eficiente está limitado por la memoria relativamente pequeña de la GPU en comparación con la de la CPU. Así, el procesado de matrices de gran tamaño almacenadas en CPU conlleva un exceso de comunicación (conocido como *overhead*) de datos hacia la GPU dispositivo (Kang et al., 2020).

6.3. Medidas cronectómicas

En el TFM se implementaron diversas medidas cronectómicas: singularidad de meta-estado dominante, área bajo el ICT, diferencia de permanencia, probabilidad de ocurrencia del meta-estado nulo, percentil 5, grado de antagonismo, *dwell time*, *leap size* y complejidad de TAS. Los resultados de aplicar la medida “singularidad de meta-estado dominante” sugieren alteraciones en la definición de los meta-estados consecuencia del avance de la EA. Un valor alto de esta medida implica la presencia del meta-estado dominante de forma única, siendo la topología observada notablemente la de este meta-estado. En la figura 5.1 se observa como, cuando los resultados son significativos, los controles presentan valores de singularidad más altos que en pacientes con EA. Esta hipótesis apoya los hallazgos de Miraglia et al. (2017), que indica una pérdida de eficiencia de comunicación entre regiones cerebrales en la banda alfa conforme avanza la EA. Los valores bajos de los pacientes en comparación con los controles sugieren alteraciones en la definición de los meta-estados en las bandas zeta y alfa, lo que concuerda con los resultados de Núñez et al. (2021). En el caso de los pacientes de EA, se observa como los valores son cercanos a uno indicando, por tanto, una pérdida de estabilidad y estructura temporal en estos sujetos (Núñez et al., 2021). En relación a la definición de los meta-estados también se calculó la medida “área bajo el ICT”. Los resultados se muestran en la figura 5.2. Los valores son diferentes en ambas BBDD, mostrando diferencias significativas en las bandas delta, zeta y alfa en la BBDD del Hospital Universitario Río Hortega, pero no en la del proyecto POCTEP. Un valor alto de esta medida refleja que la topología observada coincide con la del meta-estado dominante durante mayor tiempo que

el resto de meta-estados. En el caso de la BBDD del Hospital Universitario Río Hortega, en las bandas zeta y alfa los valores de los pacientes con DCL y con EA se concentran en torno al valor uno. Por tanto, se aprecia una menor variación en la persistencia del meta-estado dominante a medida que se desarrolla la EA. Estos resultados apoyan los resultados de Núñez et al. (2021), que indica la pérdida en la definición de los meta-estados a medida que se desarrolla la EA. Además, los pacientes presentan un rango de valores en torno al uno indicando un comportamiento ajeno a dFC de origen neural. En la BBDD del Hospital Universitario Río Hortega, en la banda delta se observa como los pacientes con DCL presentan un rango de valores alejado del valor uno y con valores más altos que los controles. Este resultado refleja la hipótesis de de Haan et al. (2012), que menciona un mecanismo de compensación neuronal en estos pacientes.

La medida “diferencia de permanencia” refleja la diferencia de permanencia temporal entre el meta-estado más persistente y el segundo en términos de persistencia (Carretero, 2022). Un valor alto de la medida indicaría una presencia mayor del meta-estado más persistente mientras que un valor bajo demostraría que no existe un meta-estado con mayor persistencia que el resto. Los resultados de aplicar esta medida se muestran en la figura 5.3. Se puede observar como, en los casos en los que las diferencias son significativas, el valor en los controles es superior que en los pacientes con DCL y con EA. Por tanto, los resultados sugieren una pérdida de estructura y aumento de la aleatoriedad a medida que se desarrolla la EA; además, se observa una menor permanencia del meta-estado dominante en pacientes que en controles, apoyando los resultados de Núñez et al. (2021). Esto apoya la idea de que los sujetos cognitivamente sanos mantienen el mismo meta-estado dominante durante más tiempo, con menos transiciones, pero más significativas. De nuevo al encontrarse diferencias significativas en las bandas zeta y alfa, los resultados apoyan la hipótesis de Miraglia et al. (2017), sugiriendo pérdidas de eficiencia en la comunicación entre regiones cerebrales en los pacientes. Además, en todos los casos los valores se encuentran alejados de uno, lo que indica un comportamiento asociado a dFC de origen neural.

La medida “probabilidad de ocurrencia del meta-estado nulo” cuantifica la precisión del ICT al resumir el dominio de un meta-estado. Esta medida trata de cuantificar la ocurrencia simultánea de un valor de correlación cercano a cero en los tres meta-estados. Esta situación se denomina “meta-estado nulo” (Carretero, 2022). En la figura 5.4 se observa como los resultados se encuentran centrados en torno al valor uno indicando origen no neuronal, salvo en la banda zeta en la BBDD del proyecto POCTEP, donde los pacientes con EA presentan valores superiores al resto de sujetos sugiriendo alteraciones en la estructura temporal conforme avanza la EA (Núñez et al., 2021). A partir de esta medida surge la métrica “percentil 5º”, que caracteriza el valor de correlación utilizado para determinar la ocurrencia del meta-estado nulo (Carretero, 2022). Un valor alto de esta medida indica que el valor de correlación correspondiente al percentil 5º se concentra en torno a la media de la distribución de los valores, y por tanto, también lo hace la distribución de la persistencia de los meta-estados. Como se observa en la figura 5.5, existen diferencias significativas entre pacientes con EA y controles en la BBDD del Hospital Universitario Río Hortega, y entre todos los grupos en la BBDD del proyecto POCTEP en todas las bandas de frecuencia. Se puede observar como los valores correspondientes a pacientes con EA y DCL se alejan del resultado obtenido en el caso de los controles. En la BBDD del Hospital Universitario Río Hortega los valores muestran una tendencia hacia valores más altos en el caso de los pacientes, mientras que en la BBDD del proyecto POCTEP los pacientes presentan valores más bajos. Los cambios en la tendencia podrían deberse a que el meta-estado dominante no es idéntico en las dos bases de datos. Por tanto, el percentil 5 de permanencia podría desviarse debido a una mínima diferencia en la topología del meta-estado dominante. Esta discrepancia podría deberse a que se utilizan meta-

estados grupales, obteniendo diferencias grupales que pueden ser eliminadas individualizando el análisis si se utilizan meta-estados individuales. Serían necesarios nuevos análisis para confirmar estas hipótesis. En ambos casos los resultados sugieren una alteración en la persistencia temporal de los meta-estados a medida que avanza la EA, apoyando los resultados de Núñez et al. (2021). En todos los casos, los resultados se encuentran alejados del valor uno indicando un comportamiento asociado a dFC de origen neural (Núñez et al., 2021).

La medida “grado de antagonismo” se encuentra relacionada con la definición de los meta-estados. Esta medida refleja la diferencia de fuerza ejercida entre la correlación más atractiva del ICT, reflejada en la correlación positiva de mayor valor, y la más repulsiva (Bautista-Salinero, 2021). En la figura 5.6 se muestran los resultados de aplicar esta medida. Se puede observar que en la BBDD del proyecto POCTEP no existen diferencias significativas, mientras que en la BBDD del Hospital Universitario Río Hortegam se puede observar como los controles presentan un rango de valores inferior a los pacientes con DCL y con EA. Los resultados presentes en la BBDD del Hospital Universitario Río Hortega, sugieren que los pacientes mantienen un comportamiento ajeno a dFC genuina al mantener valores cercanos al uno. Estos cambios sugieren que existen alteraciones en la separación entre meta-estados a medida que se desarrolla la EA de acuerdo con el estudio de Bautista-Salinero (2021). De nuevo, apoya los resultados de Núñez et al. (2021) sugiriendo una pérdida de la estructura temporal subyacente en pacientes.

En relación al tiempo promedio que dominan los meta-estados surge la medida “*dwell time*”. Cuanto mayor es el valor de esta medida, mayor tiempo permanecen los meta-estados como dominantes. En la figura 5.7 se muestra el resultado de aplicar esta medida a ambas BBDD. En ambos casos, en la banda beta-1, se observa como los controles presentan valores más altos y alejados del uno que los pacientes con EA. Este resultado sugiere alteraciones en la transición de meta-estados de acuerdo con el trabajo de Sitnikova et al. (2018), que observó una disminución en la duración de los estados de actividad de las redes neuronales asociada al avance de la EA. Sin embargo, en la banda zeta se observan distintas tendencias en cada base de datos: en la del Hospital Universitario Río Hortega son los controles los que mantienen un valor cercano al uno, mientras que en la del proyecto POCTEP son los pacientes con EA los que se comportan de esta forma. Estos resultados, a pesar de las diferentes tendencias, de nuevo apoyan el estudio de Núñez et al. (2021), que indica una alteración en la estructura temporal subyacente en la transición de meta-estados. En relación a la transición entre meta-estados, surge la medida “*leap size*”. Esta medida cuantifica el coste de transición entre meta-estados. Un valor alto implica un coste metabólico alto en la transición entre meta-estados. Como se observa en la figura 5.8, en la BBDD del Hospital Universitario Río Hortega no se aprecian diferencias significativas, mientras que en la BBDD del proyecto POCTEP los pacientes con EA presentan valores más altos que el resto de sujetos en todas las bandas. Este resultado indica que existe un mayor coste en la transición entre meta-estados a medida que avanza la EA, apoyando los resultados de Jones et al. (2012) que observa alteraciones en la dinámica de transiciones entre meta-estados a medida que se desarrolla la enfermedad que se asocia a la inestabilidad cognitiva.

Finalmente, en relación a la estructura subyacente en la secuencia temporal de meta-estados surge la medida “complejidad de TAS”. Un valor alto de la medida indica una complejidad mayor de la secuencia de activación en los meta-estados. En la figura 5.9 se observa que en ambas BBDD en la banda beta-1 existen diferencias significativas entre controles y pacientes con EA. Se observa como la distribución de los valores en los controles se concentra en torno a valores inferiores que los resultados asociados a pacientes con EA. Además, la distribución en el caso de

pacientes con EA se encuentra más próxima al valor uno. Un valor menor de complejidad sugiere que la transición entre meta-estados no es completamente aleatoria de acuerdo con Vidaurre et al. (2017). Además el aumento de la complejidad sugiere una mayor aleatoriedad y pérdida de estructura asociado al avance de la EA en concordancia con los resultados de Ramirez-Mahaluf et al. (2020).

Cabe resaltar que diversas medidas, como “meta-estado nulo”, “*dwel time*”, “*leap size*” y “complejidad de TAS” pueden presentar tendencias distintas en las BBDD utilizadas. Estas medidas se definen mediante más de un meta-estado. Como se puede observar en la figura 4.1, las BBDD únicamente presentan el meta-estado 1 de forma común. La variación observada en los resultados de las medidas podría deberse a que los meta-estados son diferentes en ambas bases de datos. Además, utilizar meta-estados individuales podría eliminar las discrepancias entre bases de datos al realizar un análisis individualizado y no grupal. Serían necesarios nuevos análisis para confirmar esta hipótesis. Además, los cambios en la tendencia podrían deberse a diferencias contextuales, ya que cada base de datos se generó en hospitales distintos y a que el tamaño de muestra es reducido en ambos casos (Ma et al., 2020).

6.4. Clasificación

El primer método de selección de características utilizado fue el método univariante de *k*-mejores. Los resultados se pueden observar en la figura 5.11. Se puede observar como la medida con mayor importancia en la BBDD del Hospital Universitario Río Hortega es el grado de antagonismo, mientras que en la BBDD del proyecto POCTEP es el “*leap size*”. En ambos casos se seleccionan las bandas zeta y alfa. Estos resultados sugieren que la mayor de las diferencias entre grupos se encuentran en la arquitectura temporal y en las transiciones entre meta-estados. Al dar una importancia alta en ambas características en zeta y alfa, los resultados apoyan el estudio de Núñez et al. (2021) reflejando alteraciones en la arquitectura temporal y en la transición entre meta-estados en estas bandas de frecuencia a medida que se desarrolla la EA. El método FCBF asigna a todas las características un valor de correlación similar. Respecto a los algoritmos RFE y SFS, el número óptimo de características se acerca al 10 % del tamaño de la base de datos en cada caso. Estos métodos seleccionan el conjunto de medidas que mejor clasifican de forma conjunta. Comparando los diagramas de barras de las figuras 5.18, 5.19, 5.24 y 5.25, se observa como en ambos casos se seleccionan de forma más común medidas en las bandas zeta y alfa, siendo SFS el método que más características selecciona en la banda alfa. En la comparación de los valores *kappa* representada en la tabla 5.21, se observa como en la mayoría de casos SFS presenta mejores resultados que RFE. Esto sugiere que la banda alfa permite efectuar una mejor distinción entre grupos que el resto de frecuencias.

Respecto a los resultados de clasificación, en las matrices de confusión se puede observar como en la BBDD del Hospital Universitario Río Hortega los pacientes con DCL mal clasificados se etiquetan como pacientes de EA, mientras que en la BBDD del proyecto POCTEP los pacientes con DCL mal clasificados se confunden con sujetos cognitivamente sanos. Por tanto, los resultados están en la línea de los obtenidos por Anderson (2019), indicando que el DCL es una etapa intermedia que mantiene un comportamiento similar al de los controles en algunos aspectos, pero que puede distinguirse del envejecimiento normal. Respecto a los resultados de clasificación cuando se combinan ambas BBDD, se puede observar como el rendimiento de los modelos es inferior al obtenido cuando se evalúan las BBDD por separado como ocurrió con el estudio de Ma et al. (2020), que encontró diferencias entre tres bases de datos con características demográficas similares. Esta

diferencia de comportamiento podría tener como origen las diferencias observadas en las medidas cronectómicas que resultan diferentes en cada base de datos y cuyo motivo se discute previamente.

Respecto al rendimiento de los distintos modelos, se puede observar como las redes neuronales, los árboles de decisión y los *random forest* presentan mejor rendimiento que las SVM y los modelos de regresión logística. Tanto los modelos de regresión logística como las SVMs realizan aproximaciones lineales. En ambos casos se utilizan aproximaciones de carácter paramétrico. De esta forma, estos modelos no hacen usos de soluciones complejas (como *splines* o polinomios de alto orden) para aproximar un problema (Westreich et al., 2010). Sin embargo, los árboles de decisión y las redes neuronales son más eficientes aproximando soluciones complejas. Por tanto, los resultados sugieren que la solución del problema de clasificación es de carácter no lineal.

Los mejores resultados de clasificación obtenidos para distinguir entre tres clases son un valor *kappa* de 0,81, una precisión del 87,5 %, una especificidad del 72,22 % y una sensibilidad del 100 % en el caso de la BBDD del Hospital Universitario Río Hortega mediante RFE. En la BBDD del proyecto POCTEP los resultados son iguales en los tres métodos de selección de características; se obtiene un valor *kappa* de 0,81, una precisión del 87,5 %, una especificidad del 81,81 % y una sensibilidad del 100 %. Cuando se combinan ambas BBDD el mejor método es FCBF, donde se obtiene un valor *kappa* de 0,63, una precisión del 75 %, una especificidad del 41,93 % y una sensibilidad del 100 %. En el caso de la clasificación entre dos clases (controles y pacientes con demencia, agrupando pacientes con DCL y con EA) en la BBDD del Hospital Universitario Río Hortega los tres métodos obtienen un valor *kappa* igual a 1, una precisión del 100 %, una especificidad del 100 % y una sensibilidad del 100 %. En la BBDD del proyecto POCTEP, tanto FCBF como SFS obtienen un valor *kappa* igual a 0.7, una precisión del 87,5 %, una especificidad del 63,63 % y una sensibilidad del 100 %. Finalmente, cuando se combinan ambas BBDD, se obtiene un valor *kappa* de 0.68 cuando se aplica RFE. Se obtiene una precisión del 87,5 %, una especificidad del 68,18 % y una sensibilidad del 96 %.

Los resultados de clasificación obtenidos pueden ser comparados con los obtenidos en otros estudios previos. Por ejemplo, Pirrone et al. (2022) logró una precisión del 73 % en la clasificación entre sujetos control y pacientes con EA (incluyendo DCL y demencia debido a EA), al analizar la potencia en diferentes bandas de frecuencia en cada sensor, mediante análisis de discriminación lineal (LDA) y validación cruzada en K grupos con K igual a 10. Por otro lado, Durongbhan et al. (2019) obtuvo una sensibilidad del 93 %, especificidad del 89 % y precisión del 90,26 % en la clasificación entre sujetos control y pacientes con EA, a través del análisis de bandas de frecuencia y canales de EEG de forma individual, mediante un algoritmo *K-Nearest Neighbor* (KNN) con $K=2$ y validación cruzada con 10 grupos. Shan et al. (2022) aplicó una red neuronal convolucional espacio-temporal dinámica (ST-GCN) a matrices de adyacencia de conectividad funcional y dinámica de las señales EEG; y mediante validación cruzada con 10 grupos, logrando una precisión del 92,3 %. Por su parte, Ruiz-Gómez et al. (2018) alcanzó una precisión del 87,10 %, especificidad del 79,41 % y sensibilidad del 82,35 % en la distinción entre sujetos control y pacientes con DCL y EA conjuntamente, mediante el análisis espectral y no lineal de las señales EEG. Aplicó LDA, análisis de discriminante cuadrático (QDA) y una red convolucional basada en grafos espacio-temporales (*Spatial Temporal Graph Convolutional Networks* o ST-GCN) y validación cruzada con 10 grupos. Vecchio et al. (2020) aplicó teoría de grafos sobre redes no dirigidas y ponderadas obtenidas mediante la coherencia lineal retardada evaluada por *exact Low Resolution Brain Electromagnetic Tomography* (eLORETA), obtuvo una precisión del 95 %, especificidad del 96 % y sensibilidad del 95 % mediante la aplicación de una SVM y validación cruzada con 20

grupos. Huggins et al. (2021) logró una precisión del 98,9 % mediante la comparación de mapas topográficos generados a partir de gráficos tiempo-frecuencia, tras el procesamiento de señales de EEG. Utilizó una red neuronal basada en aprendizaje profundo (DLNN) y validación cruzada dejando un sujeto fuera (*Leave One Out Cross Validation* o LOOCV). Meghdadi et al. (2021) utilizó distribuciones de potencia espectral derivadas empíricamente en el dominio del tiempo para distinguir entre sujetos control y pacientes con EA, mediante la aplicación de LDA y LOOCV consiguió un AUC de 0,85, y para separar a los sujetos control de los pacientes con DCL con un AUC de 0,60. Cejnek et al. (2021) obtuvo una especificidad del 90,73 % y una sensibilidad del 85,51 % en la clasificación entre sujetos control y pacientes con EA, al emplear la “novedad” (medida de entropía) de las señales EEG medida a partir de parámetros de filtrado adaptativo y considerada como una medida de entropía. Aplicó un modelo *linear neural unit* (LNU) y validación cruzada dejando tres sujetos fuera. Yu et al. (2020) alcanzó una precisión del 97,12 % en la clasificación entre sujetos control y pacientes con EA, mediante la aplicación del algoritmo de visibilidad ponderada y la transformación de canales de EEG en redes. El modelo utilizado fue *Takagi-Sugeno-Kang* (TSK) y utilizó una validación cruzada con 10 grupos. You et al. (2020) combinó señales de movimiento y fragmentos de señales EEG para la clasificación entre sujetos control y pacientes con demencia, obteniendo valores de precisión del 93,09 % en la comparación entre controles y pacientes con demencia (DCL y con EA agrupados), 98,63 % en la clasificación entre controles y pacientes con DCL, y 91,07 % en la clasificación entre controles, pacientes con DCL y pacientes con EA. En su estudio, You et al. (2020) aplicó redes convolucionales gráficas espaciotemporales basadas en la atención (*Attention-Based Spatial Temporal Graph Convolutional Networks* o AST-GCN) sin validación cruzada. Fiscon et al. (2018) obtuvo precisión del 83 % en la clasificación entre sujetos control y pacientes con EA, del 92 % en la distinción entre sujetos control y pacientes con DCL, y del 79 % para la clasificación entre pacientes con DCL y EA. En su estudio, utilizó el análisis en tiempo-frecuencia de señales EEG aplicando las transformadas de Fourier y Wavelet basado en árboles de decisión y validación cruzada con 10 grupos. Fouladi et al. (2022) obtuvo un 70 % de precisión en la clasificación entre controles, pacientes con DCL y pacientes con EA mediante la aplicación de redes neuronales convolucionales de aprendizaje profundo (*Deep Learning Convolutional Neural Network* o DLCNN) y validación cruzada con 5 grupos, sobre señales de EEG aplicando las transformadas de Fourier, Wavelet y Ricker Wavelet. Finalmente, el estudio de Carretero (2022) logró un valor *kappa* de 0,74 en la clasificación entre sujetos control, pacientes con DCL y pacientes con EA, y una precisión del 98,75 %, especificidad del 95,35 % y sensibilidad del 100 % en la clasificación binaria, utilizando la base de datos del Hospital Universitario Rio Hortega mediante la aplicación de técnicas de conectómica aplicadas a LDA y validación LOOCV. Como se evidencia en la tabla de comparaciones, los resultados de este TFM superan los valores obtenidos en los estudios mencionados. Esto sugiere que el análisis de señales EEG y la extracción de medidas conectómicas pueden ser herramientas prometedoras para el apoyo al diagnóstico de la demencia debida a EA.

6.5. Limitaciones y líneas futuras

Durante la elaboración de este TFM surgieron diversas limitaciones que podrían ser abordadas en investigaciones futuras. En relación al origen de los datos, una restricción importante se encuentra en la adquisición, específicamente en la localización de fuentes que involucró 19 electrodos para identificar 68 regiones de interés. El uso de un mayor número de electrodos aportaría una mayor resolución espacial. En relación a las BBDD, convendría considerar una ampliación en la cantidad de sujetos que las forman, asegurándose de que cada grupo cuente con el mismo tamaño.

Estudio	Modelo	Sujetos	Resultados	Parámetros utilizados
(Cejnek et al., 2021)	LNU	102 C, 7 DCL, 59 EA	e= 90,73 %; s= 85,51 %	Novedad de las señales de EEG.
(Dorongbhan et al., 2019)	KNN, k=2	80 C, 56 EA	p= 90,26 %; e= 89 %; s= 93 %	Bandas de frecuencia y canales de EEG de forma individual.
(Fiscon et al., 2018)	Árbol de Decisión	14 C, 37 DCL, 49 EA	p= 83 % CvsEA; p= 92 % CvsDCL; p= 79 % DCLvsEA	Análisis en tiempo-frecuencia de señales de EEG.
(Huggins et al., 2021)	DLNN	52 C, 37 DCL, 52 EA	p= 98,9 %	Mapas topográficos creados mediante gráficos tiempo-frecuencia resultado del procesamiento de señales de EEG.
(Meghdadi et al., 2021)	LDA	246 C, 53 DCL, 26 EA	AUC= 0,85 CvsEA; AUC= 0,60 CvsDCL	Distribuciones de potencia espectral derivadas empíricamente en el dominio del tiempo.
(Pirrone et al., 2022)	LDA	23 C, 37 DCL, 49 EA	p= 73 %	Potencia en cada banda de frecuencia de cada sensor.
(Ruiz-Gómez et al., 2018)	LDA,QDA, ST-GCN	37 C, 37 DCL, 37 EA	p= 87,10 %; e= 79,41 %; s= 82,35 %	Análisis espectral y no lineal de las señales EEG.
(Shan et al., 2022)	ST-GCN	20 C, 19 EA	p = 92,3 %	Matrices de adyacencia de conectividad funcional y dinámica de las señales EEG.
(Vecchio et al., 2020)	SVM	120 C, 175 EA	p= 95 %; e= 96 %; s= 95 %	Teoría de grafos sobre redes no dirigidas y ponderadas obtenidas mediante la coherencia lineal retardada evaluada por eLORETA.
(You et al., 2020)	AST-GCN	35 C, 35 DCL, 17 EA	p= 93,09 % CvsDCL/EA; p= 98,63 % DCLvsEA; p=91,07 % CvsDCLvsEA	Señales de movimiento y fragmentos de señales de EEG.
(Yu et al., 2020)		30 C, 30 EA	p= 97,12 %	Parámetros topológicos de la red obtenida mediante wavelet.
(Fouladi et al., 2022)	DLCNN	61 C, 56 DCL, 63 EA	p= 70 % CvsDCLvsEA	Análisis tiempo-frecuencia de señales EEG
(Carretero, 2022)	LDA	45 C, 67 DCL, 50 EA	p= 98,75 % CvsDCL/EA; e= 95,35 % CvsDCL/EA; s= 100 % CvsDCL/EA; k= 0,7439 CvsDCLvsEA	Métodos de cronectómica.
Estudio realizado en el TFM-Rio Hortega	Árbol de Decisión	45 C, 67 DCL, 50 EA	p= 100 % CvsDCL/EA; e= 100 % CvsDCL/EA; s= 100 % CvsDCL/EA; k= 0,81 CvsDCLvsEA	Métodos de cronectómica.
Estudio realizado en el TFM-POCTEP	Árbol de Decisión	38 C, 38 DCL, 50 EA	p= 87,5 % CvsDCL/EA; e= 63,63 % CvsDCL/EA; s= 100 % CvsDCL/EA; k= 0,81 CvsDCLvsEA	Métodos de cronectómica.

Tabla 6.1. Resultados obtenidos en otros estudios de EEG para identificar controles frente a pacientes con DCL y EA. p: precisión; e: especificidad; s: sensibilidad; k: valor *kappa*; AUC: área bajo la curva ROC; C: controles; DCL: deterioro cognitivo leve; EA: enfermedad de Alzheimer.

Entre los sujetos de la base de datos, los pacientes con DCL tienden a dividirse en dos subgrupos basados en si desarrollan o no demencia a lo largo del tiempo de forma similar a la realización de otros estudios. No obstante, en este estudio, no se ha separado a los pacientes con DCL debido a la ausencia de estos datos para todos los sujetos en la investigación. Además, sería beneficioso incorporar otras bases de datos para aplicar las medidas implementadas, de manera que los resultados puedan ser replicados y respaldados de manera más sólida. En relación a las bandas de frecuencia evaluadas, se han considerado únicamente las bandas delta, zeta, alfa y beta-1. La obtención de los meta-estados ha resultado en un coste computacional elevado, lo que ha limitado la inclusión de resultados correspondientes a bandas de frecuencia superiores a beta-1 en el TFM. Por tanto, sería interesante calcular las medidas presentadas en el TFM en estas bandas de alta frecuencia, con el fin de analizar las alteraciones temporales que la EA induce en su estructura. Los meta-estados fueron obtenidos para cada base de datos de forma individual en este TFM. En el caso de la BBDD del proyecto POCTEP, no se incluyeron los pacientes con EA severo, sin embargo sí que fueron utilizados para la obtención de meta-estados. Debido al alto coste computacional, no se calcularon los meta-estados sin estos sujetos. Además, sería interesante obtener los meta-estados haciendo uso de la combinación de ambas BBDD sin incluir pacientes con EA severo con el fin de que los meta-estados sean iguales en las dos. Otra posibilidad sería individualizar los análisis calculando meta-estados individuales en lugar de grupales, lo cual permitiría dejar de lado las diferencias entre meta-estados grupales existentes entre BBDD observadas en el TFM. En último término, sería interesante investigar nuevas métricas sobre el TAS y el ICT desde diversas perspectivas, permitiendo una comprensión más detallada de las alteraciones que produce la enfermedad en la información encapsulada bajo estas secuencias.

Capítulo 7

Conclusiones

7.1. Introducción	63
7.2. Consecución de los objetivos del trabajo	63
7.3. Conclusiones	63

7.1. Introducción

Para concluir, este capítulo aborda el logro de los objetivos establecidos en la introducción, las conclusiones derivadas de la implementación y aplicación de medidas cronectómicas innovadoras, y finalmente, las líneas futuras que podrían ampliar el alcance del presente trabajo.

7.2. Consecución de los objetivos del trabajo

A lo largo de la ejecución de este TFM, se han estudiado y aplicado algoritmos de optimización, medidas cronectómicas y modelos de clasificación con el propósito de analizar la actividad cerebral en estado de reposo y sus alteraciones a lo largo de la progresión de la EA. La exploración y aplicación mencionadas permitieron la consecución de los objetivos fundamentales establecidos para este TFM, los cuales se detallaron en el primer capítulo:

- i. Se ha realizado la búsqueda y revisión de artículos científicos y documentación con el fin de adquirir conocimientos sobre la EA y las medidas cronectómicas implementadas.
- ii. Se han examinado diferentes fuentes de datos con el propósito de comprender el procesado de la señal y el método de subrogación.
- iii. Se ha llevado a cabo la búsqueda y análisis de artículos sobre optimización de algoritmos de procesado de señal. Se utilizaron ficheros en C/C++ y procesado en GPU para optimizar el proceso de subrogación.
- iv. Se ha utilizado la herramienta MATLAB para la ejecución del procesamiento de señales, implementación de medidas cronectómicas y la evaluación de los resultados obtenidos.
- v. Se ha utilizado la herramienta Pyhon para la implementación de modelos de clasificación a partir de los resultados de las medidas cronectómicas.
- vi. Se han empleado pruebas estadísticas en los resultados obtenidos con el propósito de analizar la presencia de discrepancias a nivel poblacional entre los resultados de los diferentes conjuntos de individuos.
- vii. Se han aplicado métodos de selección de características y aplicado diversos modelos de clasificación. El modelo de mejor rendimiento en la mayoría de casos fue el árbol de decisión.
- viii. Se han analizado y comparado los resultados de este TFM con los resultados presentes en trabajos previamente publicados, los cuales se enfocan en áreas temáticas afines a la presente investigación.
- ix. En última instancia, se han extraído conclusiones, se han identificado limitaciones y posibles direcciones para investigaciones futuras a partir del trabajo efectuado.

7.3. Conclusiones

Después del análisis presentado en el capítulo anterior, ante los resultados de los métodos de optimización, las medidas cronectómicas y los algoritmos de clasificación, es posible plantear las siguientes conclusiones:

1. La máxima mejora en términos de optimización se logra mediante una combinación de la utilización de ficheros en C/C++ junto con la ejecución en GPU. Este enfoque resulta más efectivo al aplicar los archivos C/C++ a los métodos que involucran el procesamiento de múltiples matrices de gran tamaño. Por otro lado, la ejecución en GPU es más adecuada para métodos que no requieren la transferencia de una gran cantidad de datos desde la CPU a la GPU. En el contexto del proceso examinado, se ha observado que la optimización más significativa se logra al optimizar el cálculo del IAC a través de archivos en C reduciendo el tiempo de ejecución un 41,6 %, mientras que la obtención de RPs y la aplicación del algoritmo de detección de comunidades se benefician de la ejecución en GPU reduciendo el tiempo de ejecución en un 61,35 % y un 30,38 %, respectivamente.
2. Los resultados de aplicar las medidas cronectómicas a ambas BBDD permiten observar alteraciones en la definición de los meta-estados, alteraciones en la transición de meta-estados y aumento de la aleatoriedad conforme progresa la enfermedad.
3. En último término, la aplicación de las medidas implementadas en este TFM da lugar a resultados de clasificación que superan a los obtenidos en investigaciones anteriores. Se obtuvo una precisión del 87,5 % en las BBDD del Hospital Universitario Río Hortega y del proyecto POCTEP, y una precisión del 75 % para la combinación de ambas BBDD en la clasificación entre controles, pacientes con DCL y pacientes con EA. En la clasificación entre controles y el conjunto de pacientes con DCL y con EA, se obtuvo una precisión de 100 % en la BBDD del Hospital Universitario Río Hortega, de 87,5 % en la BBDD del proyecto POCTEP y de 87,5 % para la combinación de ambas BBDD. Estos resultados son prometedores y sugieren que las medidas derivadas de la cronectómica podrían ayudar al diagnóstico y la comprensión de la EA.

Glosario de siglas y acrónimos

Aβ:	Proteína Beta-Amiloide
BBDD:	Bases de Datos
DLCNN:	<i>Deep Learning Convolutional Neural Network</i>
DCL:	Deterioro Cognitivo Leve
dFC:	Conectividad Funcional Dinámica
EA:	Enfermedad de Alzheimer
EEG:	Electroencefalografía
eLORETA:	exact Low Resolution Brain Electromagnetic Tomography
EPSP:	Potencial Postsináptico Excitatorio
FCBF:	<i>Fast-Correlation-Based-Filter</i>
FDR:	<i>False Discovery Rate</i>
IAC:	<i>Instantaneous Amplitude Correlation</i>
ICT:	<i>Instantaneous Correlation Tensor</i>
LDA:	<i>Linear Discriminant Analysis</i>
LOOCV:	<i>Leave One Out Cross Validation</i>
LZC:	<i>Lempel-Ziv Complexity</i>
MEG:	Magnetoencefalografía
MLP:	<i>Multi-Layer Perceptron</i>
MMSE:	<i>Mini-Mental State Examination</i>
NIA-AA:	<i>National Institute on Aging - Alzheimer's Association</i>
QDA:	<i>Quadratic Discriminant Analysis</i>
RFE:	<i>Recursive Feature Elimination</i>
ROC:	<i>Region of Convergence</i>
RP:	<i>Recurrence Plots</i>
sFC:	Conectividad Funcional Estática
SFS:	<i>Sequential Forward Selection</i>
sLORETA:	<i>Standarized Low Resolution Brain Electromagnetic Tomography</i>
SVM:	<i>Support Vector Machine</i>
ST-GNN:	<i>Spatial Temporal Graph Convolutional Networks</i>
TAS:	<i>Temporal Activation Sequence</i>
TFG:	Trabajo Fin de Grado
TFM:	Trabajo Fin de Máster

Bibliografía

- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3):270–279.
- Alhenawi, E., Al-Sayyed, R., Hudaib, A., & Mirjalili, S. (2022). Feature selection methods on gene expression microarray data for cancer classification: A systematic review. *Computers in Biology and Medicine*, 140:105051.
- Alzheimer's Association (2023). 2023 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 19(4):1598–1695.
- Anderson, N. D. (2019). State of the science on mild cognitive impairment (mci). *CNS spectrums*, 24(1):78–87.
- Asaduzzaman, A., Trent, A., Osborne, S., Aldershof, C., & Sibai, F. N. (2021). Impact of CUDA and OpenCL on Parallel and Distributed Computing. In 2021 8th International Conference on Electrical and Electronics Engineering (ICEEE), 238-242.
- Bautista-Salineró, M. (2021). Estudio de la arquitectura temporal de la actividad neuronal en la demencia por enfermedad de Alzheimer mediante técnicas de conectividad funcional dinámica. *Trabajo de Fin de Grado*. Universidad Rey Juan Carlos.
- Berrar, D. (2019). Cross-Validation. In Encyclopedia of Bioinformatics and Computational Biology. *Academic Press*, 1(1):542–545.
- Boateng, E. Y. & Abaye, D. A. (2019). A review of the logistic regression model with emphasis on medical research. *Journal of data analysis and information processing*, 7(4):190–207.
- Breyer, M., Van Craen, A., & Pflüger, D. (2022). A comparison of SYCL, OpenCL, CUDA, and OpenMP for massively parallel support vector machine classification on multi-vendor hardware. In *International Workshop on OpenCL*, 1-12.
- Bühlmann, P. & Van De Geer, S. (2011). *Statistics for high-dimensional data: methods, theory and applications*. Springer Science & Business Media.
- Calhoun, V., Miller, R., Pearlson, G., & Adalı, T. (2014). The chronnectome: Time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron*, 84(2):262–274.
- Cancino, M. & Rehbein, L. (2016). Factores de riesgo y precursores del Deterioro Cognitivo Leve (DCL): Una mirada sinóptica. *Terapia Psicológica*, 34(3):183–189.

- Carr, J. J. & Brown, J. M. J. M. (2001). *Introduction to biomedical equipment technology*. Ed. Pearson.
- Carretero, P. (2022). Aplicación de métodos de cronectómica para caracterizar las alteraciones en la arquitectura temporal de la actividad neuronal en estado de reposo provocadas por la enfermedad de Alzheimer *Trabajo de Fin de Grado*. Universidad de Valladolid.
- Cassani, R., Estarellas, M., San-Martin, R., Fraga, F. J., & Falk, T. H. (2018). Systematic review on resting-state EEG for alzheimer's disease diagnosis and progression assessment. *Disease Markers*, 2018(1):5174815.
- Cejnek, M., Vysata, O., Valis, M., & Bukovsky, I. (2021). Novelty detection-based approach for Alzheimer's disease and mild cognitive impairment diagnosis from EEG. *Medical & Biological Engineering & Computing*, 59(11-12):2287–2296.
- Charbuty, B. & Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*, 2(01):20–28.
- Cohen, M. X. (2017). Where does EEG come from and what does it mean? *Trends in Neurosciences*, 40(4):208–218.
- Dauwels, J., Vialatte, F., & Cichocki, A. (2010). Diagnosis of Alzheimers disease from EEG Signals: Where Are We Standing? *Current Alzheimer research*, 7:487–505.
- de Haan, W., Mott, K., van Straaten, E. C. W., Scheltens, P., & Stam, C. J. (2012). Activity dependent degeneration explains hub vulnerability in Alzheimer's disease. *PLOS Computational Biology*, 8(8):1–14.
- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *NeuroImage*, 31(3):968–980.
- Dmitrović, S. (2023). What is c++? In *Modern C++ for Absolute Beginners: A Friendly Introduction to the C++ Programming Language and C++ 11 to C++ 23 Standards*. Springer.
- Durongbhan, P., Zhao, Y., Chen, L., Zis, P., De Marco, M., Unwin, Z. C., Venneri, A., He, X., Li, S., Zhao, Y., Blackburn, D. J., & Sarrigiannis, P. G. (2019). A dementia classification framework using frequency and time-frequency features based on EEG signals. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(5):826–835.
- Figuroa, R. L., Zeng-Treitler, Q., Kandula, S., & Ngo, L. H. (2012). Predicting sample size required for classification performance. *BMC Medical Informatics and Decision Making*, 12:1–10.
- Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., Papademetris, X., & Constable, R. T. (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature Neuroscience*, 18(11):1664–1671.
- Fiscon, G., Weitschek, E., Cialini, A., Felici, G., Bertolazzi, P., De Salvo, S., Bramanti, A., Bramanti, P., & De Cola, M. C. (2018). Combining EEG signal processing with supervised methods for Alzheimer's patients classification. *BMC Medical Informatics and Decision Making*, 18(1):35.

- Fouladi, S., Safaei, A. A., Mammone, N., Ghaderi, F., & Ebadi, M. J. (2022). Efficient Deep Neural Networks for Classification of Alzheimer's Disease and Mild Cognitive Impairment from Scalp EEG Recordings. *Cognitive Computation*, 14(4):1247–1268.
- García Cobo, P. (2022). Enfermedad de Alzheimer y salud mental. *Trabajo de Fin de Grado*. Universidad de Cantabria.
- Huggins, C. J., Escudero, J., Parra, M. A., Scally, B., Anghinah, R., Vitória Lacerda De Araújo, A., Basile, L. F., & Abasolo, D. (2021). Deep learning of resting-state electroencephalogram signals for three-class classification of Alzheimer's disease, mild cognitive impairment and healthy ageing. *Journal of Neural Engineering*, 18(4):046087.
- Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., Della Penna, S., Duyn, J. H., Glover, G. H., Gonzalez-Castillo, J., Handwerker, D. A., Keilholz, S., Kiviniemi, V., Leopold, D. A., de Pasquale, F., Sporns, O., Walter, M., & Chang, C. (2013). Dynamic functional connectivity: Promise, issues, and interpretations. *NeuroImage*, 80:360–378. Mapping the Connectome.
- Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J. L., Montine, T., Phelps, C., Rankin, K. P., Rowe, C. C., Scheltens, P., Siemers, E., Snyder, H. M., Sperling, R., Elliott, C., Masliah, E., Ryan, L., & Silverberg, N. (2018). NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, 14:535–562.
- Jackins, V., Vimal, S., Kaliappan, M., & Lee, M. Y. (2021). Ai-based smart prediction of clinical disease using random forest classifier and Naive Bayes. *The Journal of Supercomputing*, 77:5198–5219.
- Jafari, Z., Kolb, B. E., & Mohajerani, M. H. (2020). Neural oscillations and brain stimulation in Alzheimer's disease. *Progress in Neurobiology*, 194:1018782.
- Jain, S. & Saha, A. (2022). Rank-based univariate feature selection methods on machine learning classifiers for code smell detection. *Evolutionary Intelligence*, 15(1):609–638.
- Jones, D. T., Vemuri, P., Murphy, M. C., Gunter, J. L., Senjem, M. L., Machulda, M. M., Przybelski, S. A., Gregg, B. E., Kantarci, K., Knopman, D. S., Boeve, B. F., Petersen, R. C., & Jack, Jr, C. R. (2012). Non-stationarity in the “resting brain's” modular architecture. *PLOS ONE*, 7(6):1–15.
- Kang, H., Kwon, H. C., & Kim, D. (2020). Hpmax: heterogeneous parallel matrix multiplication using CPUs and GPUs. *Computing*, 102(12):2607–2631.
- Kumari, S., Singh, K., Khan, T., Ariffin, M. M., Mohan, S. K., Baleanu, D., & Ahmadian, A. (2023). A novel approach for continuous authentication of mobile users using Reduce Feature Elimination (RFE): A machine learning approach. *Mobile Networks and Applications*.
- Kwame, A. E., Martey, E. M., & Chris, A. G. (2017). Qualitative assessment of compiled, interpreted and hybrid programming languages. *Communications*, 7(7):8–13.
- Lempel, A. & Ziv, J. (1976). On the complexity of finite sequences. *IEEE Transactions on information theory*, 22(1):75–81.

- Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne, C., Burns, A., Cohen-Mansfield, J., Cooper, C., Costafreda, S. G., Dias, A., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Kivimäki, M., Larson, E. B., Ogunniyi, A., Orgeta, V., Ritchie, K., Rockwood, K., Sampson, E. L., Samus, Q., Schneider, L. S., Selbæk, G., Teri, L., & Mukadam, N. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet*, 396:413–446.
- Lopes da Silva, F. (2013). EEG & MEG: Relevance to neuroscience. *Neuron*, 80(5):1112–1128.
- Ma, X., Long, L., Moon, S., Adamson, B. J., & Baxi, S. S. (2020). Comparison of population characteristics in real-world clinical oncology databases in the us: Flatiron health, seer, and nprc. *Medrxiv*, pages 2020–03.
- MathWorks-Inc. (2023). Operating System Commands (r2023a). *The MathWorks Inc.*
- Maturana-Candelas, A., Gómez, C., Poza, J., Ruiz-Gómez, S. J., & Hornero, R. (2020). Inter-band Bispectral Analysis of EEG Background Activity to Characterize Alzheimer’s Disease Continuum. *Frontiers in Computational Neuroscience*, 14:70.
- Matusiak, M. (2021). An optimized MATLAB tool for efficient evaluation of fractional-order differential equations of time-varying orders. In *2021 American Control Conference (ACC)*.
- Meghdadi, A. H., Stevanović Karić, M., McConnell, M., Rupp, G., Richard, C., Hamilton, J., Salat, D., & Berka, C. (2021). Resting state EEG biomarkers of cognitive decline associated with Alzheimer’s disease and mild cognitive impairment. *PLoS One*, 16(2):e0244180.
- Merelo-Guervós, J. J., Blancas-Alvarez, I., Castillo, P. A., Romero, G., García-Sánchez, P., Rivas, V. M., García-Valdez, M., Hernández-Águila, A., & Román, M. (2016). Ranking the Performance of Compiled and Interpreted Languages in Genetic Algorithms. In *Proceedings of the International Conference on Evolutionary Computation Theory and Applications, Porto, Portugal*, volume 11.
- Miraglia, F., Vecchio, F., & Rossini, P. M. (2017). Searching for signs of aging and dementia in EEG through network analysis. *Behavioural Brain Research*, 317:292–300.
- Müller-Putz, G. R. (2020). Chapter 18 - Electroencephalography. Handbook of clinical neurology. En Ramsey, N. F. & del R. Millán J. editores. Brain-Computer Interfaces. *Handbook of Clinical Neurology*. Elsevier, 168:249–262.
- Niedermeyer, E., Schomer, D., & da Silva, F. (2011). *Niedermeyer’s Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Niedermeyer’s Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Wolters Kluwer/Lippincott Williams & Wilkins Health.
- Noback, C. R., Strominger, N. L., & Demarest, R. J. (1993). *El sistema nervioso: introducción y repaso*. Interamericana McGraw-Hill, México. (Traducción de la 4ª edición original en inglés).
- Novo-Olivas, C. (2010). Aproximaciones al estudio de la neurociencia del comportamiento. En: *Mapeo Electroencefalográfico y Neurofeedback*. Universidad Autónoma de Guerrero, 371-412, isbn = 978-970-764-911-8.
- Núñez, P., Poza, J., Gómez, C., Rodríguez-González, V., Hillebrand, A., Tewarie, P., Ángel Tola-Arribas, M., Cano, M., & Hornero, R. (2021). Abnormal meta-state activation of dynamic brain networks across the Alzheimer spectrum. *NeuroImage*, 232:117898.

- Pirrone, D., Weitschek, E., Di Paolo, P., De Salvo, S., & De Cola, M. C. (2022). EEG signal processing and supervised machine learning to early diagnose Alzheimer's disease. *Applied Sciences*, 12(11):5413.
- Poza, J. (2008). Tiempo-frecuencia la actividad magnetoencefalográfica espontánea en la enfermedad de Alzheimer. *Tesis Doctoral*. Universidad de Valladolid.
- Rackauckas, C. (2023). A comparison between differential equation solver suites in matlab, r, julia, python, c, mathematica, maple, and fortran. *Authorea Preprints*.
- Ramirez-Mahaluf, J. P., Medel, V., Ángeles Tepper, Allende, L. M., Sato, J. R., Ossandon, T., & Crossley, N. A. (2020). Transitions between human functional brain networks reveal complex, cost-efficient and behaviorally-relevant temporal paths. *NeuroImage*, 219:117027.
- Real Academia Española (2023). Diccionario de la lengua española, 23ª ed.
- Renganathan, V. (2019). Overview of artificial neural network models in the biomedical domain. *Bratislava Medical Journal/Bratislavské Lekárske Listy*, (7).
- Rossini, P. M., Di Iorio, R., Vecchio, F., Anfossi, M., Babiloni, C., Bozzali, M., Bruni, A. C., Cappa, S. F., Escudero, J., & Fraga, F. J. (2020). Early diagnosis of Alzheimer's disease: the role of biomarkers including advanced EEG signal analysis. report from the IFCN-sponsored panel of experts. *Clinical Neurophysiology*, 131(6):1287–1310.
- Ruiz-Gómez, S. J., Gómez, C., Poza, J., Gutiérrez-Tobal, G. C., Tola-Arribas, M. A., Cano, M., & Hornero, R. (2018). Automated multiclass classification of spontaneous EEG activity in Alzheimer's disease and mild cognitive impairment. *Entropy*, 20(1):35.
- Sanei, S. & Chambers, J. (2013). *EEG Signal Processing*. Wiley.
- Scott, L. R., Clark, T., & Bagheri, B. (2021). *Scientific parallel computing*. Princeton University Press.
- Shafiee, S., Lied, L. M., Burud, I., Dieseth, J. A., Alsheikh, M., & Lillemo, M. (2021). Sequential forward selection and support vector regression in comparison to lasso regression for spring wheat yield prediction based on uav imagery. *Computers and Electronics in Agriculture*, 183:106036.
- Shan, X., Cao, J., Huo, S., Chen, L., Sarrigiannis, P. G., & Zhao, Y. (2022). Spatial-temporal graph convolutional network for Alzheimer classification based on brain functional connectivity imaging of electroencephalogram. *Human Brain Mapping*, 1:1–16.
- Sitnikova, T. A., Hughes, J. W., Ahlfors, S. P., Woolrich, M. W., & Salat, D. H. (2018). Short timescale abnormalities in the states of spontaneous synchrony in the functional neural networks in alzheimer's disease. *NeuroImage: Clinical*, 20:128–152.
- Siuly, S., Li, Y., & Zhang, Y. (2016). Electroencephalogram (EEG) and Its Background. En: *EEG Signal Analysis and Classification*. Health Information Science. Springer, Cham, 3-21.
- Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLOS Computational Biology*, 1(4).

- Tahir, M. A., Farhat, Z., Khan, M. U. R., Gaber, J., Anwar, M., & Ovuoraye, P. E. (2023). Integration of Machine Learning Algorithms Classifiers and Sequential Forward Selection Features in Diabetes Prediction. In *2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)*.
- Tewarie, P., Abeysuriya, R., Áine Byrne, O'Neill, G. C., Sotiropoulos, S. N., Brookes, M. J., & Coombes, S. (2019). How do spatially distinct frequency specific meg networks emerge from one underlying structural connectome? the role of the structural eigenmodes. *NeuroImage*, 186:211–220.
- Uhlhaas, P. J., Haenschel, C., Nikolić, D., & Singer, W. (2008). The Role of Oscillations and Synchrony in Cortical Networks and Their Putative Relevance for the Pathophysiology of Schizophrenia. *Schizophrenia Bulletin*, 34(5):927–943.
- Vecchio, F., Miraglia, F., Alù, F., Menna, M., Judica, E., Cotelli, M., & Rossini, P. M. (2020). Classification of Alzheimer's disease with respect to physiological aging with innovative EEG biomarkers in a machine learning implementation. *Journal of Alzheimer's Disease*, 75(4):1253–1261.
- Venkatesh, B. & Anuradha, J. (2019). A hybrid feature selection approach for handling a high-dimensional data. In *Innovations in Computer Science and Engineering: Proceedings of the Sixth ICICSE 2018*, pages 365–373. Springer.
- Vidaurre, D., Smith, S. M., & Woolrich, M. W. (2017). Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences*, 114(48):12827–12832.
- Vohryzek, J., Deco, G., Cessac, B., Kringelbach, M. L., & Cabral, J. (2020). Ghost Attractors in Spontaneous Brain Activity: Recurrent Excursions Into Functionally-Relevant BOLD Phase-Locking States. *Frontiers in Systems Neuroscience*, 14.
- Westreich, D., Lessler, J., & Funk, M. J. (2010). Propensity score estimation: neural networks, support vector machines, decision trees (cart), and meta-classifiers as alternatives to logistic regression. *Journal of Clinical Epidemiology*, 63(8):826–833.
- You, Z., Zeng, R., Lan, X., Ren, H., You, Z., Shi, X., Zhao, S., Guo, Y., Jiang, X., & Hu, X. (2020). Alzheimer's disease classification with a cascade neural network. *Frontiers in Public Health*, 8:584387.
- Yu, H., Zhu, L., Cai, L., Wang, J., Liu, J., Wang, R., & Zhang, Z. (2020). Identification of Alzheimer's EEG with a WVG network-based fuzzy learning approach. *Frontiers in Neuroscience*, 14:641.
- Yu, L. & Liu, H. (2004). Efficient feature selection via analysis of relevance and redundancy. *The Journal of Machine Learning Research*, 5:1205–1224.
- Zhang, X. & Liu, C.-A. (2023). Model averaging prediction by K-fold cross-validation. *Journal of Econometrics*, 235(1):280–301.