

**UNIVERSIDAD FRANCISCO DE VITORIA**

**Escuela Internacional de Doctorado**



Universidad  
Francisco de Vitoria  
**UFV** Madrid  
*Escuela Internacional  
de Doctorado*

---

**Análisis de electroencefalogramas  
como marcador de la enfermedad de  
Parkinson mediante técnicas de  
Deep Learning**

---

**TESIS DOCTORAL DE**

**Ana María Maitín López**

**DIRECTORES**

**Dr. Juan Pablo Romero Muñoz**

**Dr. Álvaro José García Tejedor**

**Programa de Doctorado en Biotecnología,  
Medicina y Ciencias Biosanitarias**

**2023**



# Análisis de electroencefalogramas como marcador de la enfermedad de Parkinson mediante técnicas de Deep Learning

*Tesis doctoral presentada por*

**Ana María Maitín López**

*para la obtención del grado de*

**Doctor en Biotecnología, Medicina y Ciencias Biosanitarias**

*Directores:*

**Dr. Juan Pablo Romero Muñoz**

**Dr. Álvaro José García Tejedor**

**Programa de Doctorado en Biotecnología,  
Medicina y Ciencias Biosanitarias**

**Escuela Internacional de Doctorado  
Universidad Francisco de Vitoria**

**2023**



*A Sergio con todo mi amor.  
A mi hija Celeste,  
los límites están donde tú los pongas.*



# Agradecimientos

Quisiera comenzar expresando mi agradecimiento a mis directores, Juan Pablo Romero Muñoz y Álvaro García Tejedor, que han hecho posible este trabajo. Por su confianza, dedicación y apoyo durante este periodo, por guiarme en el transcurso del aprendizaje, por su profesionalidad y rigor, que me han enriquecido a nivel profesional y personal. Gracias por haberlo hecho posible.

También me gustaría mostrar mi gratitud hacia todas aquellas personas que han nutrido mi camino durante estos años.

A Aida Arroyo, por tu amistad, con quien empecé este viaje y he compartido tantos días buenos y malos. Gracias por escucharme y por esas cervezas en el momento justo.

A mis compañeros del CEIEC. A Alberto Nogales, por tu apoyo y amistad, y a Iván Barcia, siempre dispuesto a ayudar y del que he aprendido tanto de informática.

A todos los becarios que han pasado por el CEIEC con quienes he tenido el placer de trabajar. Especialmente, a Antonio Morales, Pedro Chazarra, Silvia Romero, Enrique Ruiz e Inés Álvarez de Toledo, por su colaboración durante el transcurso de la tesis.

A mis compañeros del GINDAT. A Fran Sánchez, Yeray González y Laura Valenzuela, por su importante papel en el registro de los datos, y a Marcos Moreno, por estar disponible y dispuesto a ayudarme cuando lo he necesitado.

A Diego Herráez y Ramiro Perezzan, por ser los primeros en confiar en mí y buscar los medios para abrir las puertas de este proyecto. También, gracias por vuestro apoyo durante el periodo de análisis de señales y por compartir vuestro conocimiento.

A Ignacio Serrano y María Dolores del Castillo, por instruirme en el proceso de limpieza de artefactos del EEG.

Por último, quisiera agradecer a mi familia todo su apoyo en cada uno de los pasos

que he dado y me han conducido hasta aquí. Por darme ánimos cuando los míos fallaban, por darme su tiempo y ayudarme en todo lo que necesitaba. Pero sobre todo gracias a mi marido Sergio por confiar siempre en mí, por ser mi soporte, por su paciencia, su ayuda en los momentos difíciles, por hacerme mejor persona, por nutrir mis sueños y animarme a alcanzarlos. A todos ellos, gracias por todo.



# Abstract

Parkinson's disease (PD), characterized by the progressive degeneration of dopaminergic neurons, is the second most common neurodegenerative disorder after Alzheimer's disease, affecting 7-10 million people worldwide. This is equivalent to 0,3% of the world's population, being one of the leading causes of permanent physical disability.

The diagnosis of PD is based on the physical examination of the patient and the recognition of certain symptoms. The specialist's experience in detecting the disease is crucial, so this criterion may vary among different professionals. The diagnosis is made when the patient has motor symptoms (bradykinesia, muscle rigidity and resting tremor). These symptoms appear when 50%-70% of the dopaminergic neurons are already affected, so delaying the diagnosis of this disease may mean losing the opportunity to treat the patient in early stages of the disease. Although the treatment of PD is symptomatic, the main therapy is based on the administration of levodopa. It should be noted that the detection of this disease at an early stage may favor the search for course-modifying therapies or interventions that, when applied early, can improve the quality of life of patients.

Currently there are no specific diagnostic tests for PD. To complement the clinical diagnosis, alternative tests can be performed to rule out other disorders and favor an accurate diagnosis. These tests, usually imaging tests, are not always available to patients and, moreover, are costly for the health system. In this thesis we propose the analysis with AI techniques of electroencephalograms (EEGs) as the test to support the clinical diagnosis of PD. The electroencephalography has numerous advantages among which are availability, low cost, and being a non-invasive test with high temporal resolution.

However, non-invasive monitoring of brain activity through the scalp adds a high noise content to the recorded signals, increasing their complexity and making them difficult to analyze. Consequently, noise minimization and artifact removal preprocessing is required for further analysis. Due to the nature of the data, as a first approach to carry out the identification of patients with PD, time series analysis techniques were used. EEGs were analyzed by means of different connectivity measures defined in their temporal and spectral domain. The results of this analysis showed an increase in connectivity in the PD

patients with respect to the controls, in addition to indicating that the effects of medication brought the connectivity values closer to the state of the healthy participants. However, the results obtained were influenced by the EEG cleaning process, and given that this process is not standardized, it can eliminate components intrinsic to PD, making comparison between studies difficult. These limitations led us to think that in order to develop an objective and accurate marker of the disease, we should look for other types of data analysis techniques, robust with respect to signal noise. We found these techniques in the field of Artificial Intelligence and more specifically in the field of Deep Learning (DL).

DL techniques have proven their efficiency in the analysis of EEGs in other neurological diseases. In order to define a starting point for our research, a literature review of Machine Learning techniques used for EEG analysis in PD was performed. Most of the articles considered covered the diagnostic problem, predominantly using Support Vector Machine (SVM) models and neural models based on Artificial Neural Networks (ANN). Technical aspects related to the architecture of the models and methods of validation of the results were evaluated. An analysis focused on the typology of the study within the field of neuroscience indicated the lack of studies related to functional connectivity from EEGs. From a computational approach, the absence of Natural Language Processing (NLP) models for the analysis EEG signals was particularly noteworthy, since these models have shown great potential in fields unrelated to text processing. Therefore, in this thesis we have proposed the development of a disease marker by means of these two approaches.

To develop a marker based on functional connectivity from EEGs, we first modeled the EEGs by means of graphs where the nodes are the electrodes and the edges represent the values of a given measure of connectivity among the 4 that have been evaluated. To carry out this process optimally, we developed a library, EEGraph, which allows to obtain these graphs from the records collected by the different electrodes distributed on the patient's scalp. Furthermore, this library is being used by researchers in different projects in the field of neuroscience. Secondly, we used a Graph Convolutional Neural Network (GCNN) to analyze the differences in connectivity between PD patients in different dopaminergic states and controls. This analysis showed a significant increase in brain connectivity in PD patients compared to controls, indicating greater independence in the functions of healthy brains. Spectral measures showed a greater ability than temporal measures to generalize the results to particular cases. No significant differences in functional connectivity produced by medication were observed. The results obtained for this biomarker were  $93,08 \pm 0,95\%$  for accuracy and a value of  $95,29 \pm 1,19\%$  for sensitivity in the test set for controls versus PD patients before the intake of medication. The values for these metrics for controls vs. PD patients after taking medication were

92,04  $\pm$  2.75 % and 93,53  $\pm$  4,29 % for accuracy and sensitivity, respectively, demonstrating that this test is not affected by connectivity differences produced by levodopa-derived dopaminergic activation.

Regarding PLN techniques, we developed a marker based on a BERT model to identify PD patients. In order to apply these techniques, we performed an equivalence between EEG and text, given the sequential nature of both types of data. Different experiments and various training strategies were carried out depending on the electroencephalographic tests recorded and the number of channels. The model results indicated that resting tests did not improve model performance, but did decrease network overtraining, indicating the need for a larger amount of data to perform training and generalize the results. The values of the metrics obtained for this biomarker were, for the blind test set, 86,41  $\pm$  7,08 % for accuracy and 81,63  $\pm$  11,58 % for sensitivity in the case of motor tests with 64 electrodes.

Both through the functional connectivity study using graphs and through the signal analysis using PLN techniques, EEG-based PD biomarkers have shown good performance on a blind set with a high sensitivity in the disease detection. Therefore, we conclude that electroencephalography has shown to be able to record the disease state in the patient and that, by means of DL techniques, disease-specific tests can be developed to support clinical diagnosis and favor its detection in prodromic states.



# Resumen

La enfermedad de Parkinson (EP), caracterizada por la degeneración progresiva de las neuronas dopaminérgicas, es el segundo trastorno neurodegenerativo más común después de la enfermedad de Alzheimer afectando a 7-10 millones de personas en el mundo. Este valor equivale al 0,3 % de la población mundial, siendo una de las principales causas de discapacidad física permanente.

El diagnóstico de la EP se basa en el examen físico del paciente y el reconocimiento de determinados síntomas. La experiencia del especialista en la detección de la enfermedad es crucial, por lo que este criterio puede variar entre distintos profesionales. El diagnóstico se realiza cuando el paciente presenta síntomas motores (bradicinesia, rigidez muscular y temblor en reposo). Dichos síntomas aparecen cuando entre el 50 % y el 70 % de las neuronas dopaminérgicas están ya afectadas, por lo que retrasar el diagnóstico de esta enfermedad puede significar perder la oportunidad de tratar al paciente en fases precoces de la enfermedad. Aunque el tratamiento de la EP es sintomático, la terapia principal se basa en la administración de levodopa. Se debe tener en cuenta que la detección de esta enfermedad en un estadio precoz puede favorecer la búsqueda de terapias modificadoras del curso de la enfermedad o intervenciones que al aplicarse precozmente puedan mejorar la calidad de vida de los pacientes.

Actualmente no existen pruebas diagnósticas específicas de la EP. Para complementar el diagnóstico clínico, se pueden realizar pruebas alternativas que permiten descartar otros trastornos y favorecer un diagnóstico preciso. Estas pruebas, generalmente de imagen, no siempre están disponibles para los pacientes y, además, son costosas para el sistema sanitario. En esta tesis proponemos el análisis con técnicas de IA de los electroencefalogramas (EEGs) como la prueba de apoyo al diagnóstico clínico de la EP. La electroencefalografía cuenta con numerosas ventajas entre las cuales se encuentran una alta disponibilidad, bajo coste, además de ser una prueba no invasiva con una gran resolución temporal.

Sin embargo, la monitorización no invasiva de la actividad cerebral a través del cuero cabelludo añade un alto contenido de ruido a las señales registradas, aumentando su

complejidad y dificultando su análisis. Consecuentemente, se requiere de un tratamiento previo de minimización de ruido y eliminación de artefactos para su posterior análisis. Debido a la naturaleza de los datos, como primera aproximación para llevar a cabo la identificación de pacientes con EP, se utilizaron las técnicas empleadas en el análisis de series temporales. Los EEGs se analizaron por medio de diferentes medidas de conectividad definidas en su dominio temporal y espectral. Los resultados de este análisis reflejaron un aumento de la conectividad en los pacientes enfermos con respecto a los controles, además de indicar que los efectos de la medicación acercaban los valores de conectividad hacia el estado de los participantes sanos. Sin embargo, los resultados obtenidos se vieron influenciados por el proceso de limpieza de los EEGs, y dado que este proceso no está estandarizado, puede eliminar componentes intrínsecas a la EP dificultando la comparación entre estudios. Estas limitaciones nos llevaron a pensar que para desarrollar un marcador objetivo y preciso de la enfermedad, debíamos buscar otro tipo de técnicas de análisis de datos que fuesen robustas respecto al ruido de las señales. Encontramos estas técnicas en el campo de la Inteligencia Artificial y más concretamente en el campo del Deep Learning (DL).

Las técnicas de DL han demostrado su eficiencia en el análisis de EEGs en otras enfermedades neurológicas. Para definir un punto de partida en nuestra investigación se realizó una revisión de la literatura acerca de las técnicas de Machine Learning utilizadas para el análisis de EEGs en la EP. La mayoría de los artículos considerados abarcaron el problema de diagnóstico, predominando el uso de modelos Support Vector Machine (SVM) y de modelos neuronales, basados en Redes Neuronales Artificiales (ANN). Se evaluaron aspectos técnicos, relacionados con la arquitectura de los modelos y métodos de validación de los resultados. Un análisis enfocado en la tipología del estudio dentro del campo de la neurociencia, indicó la falta de estudios relacionados con la conectividad funcional a partir de EEGs. Desde un enfoque computacional, destacó especialmente la ausencia de modelos de Procesamiento del Lenguaje Natural (PLN) para analizar las señales del EEG, ya que estos modelos han mostrado un gran potencial en campos no relacionados con el procesamiento de texto. Por lo tanto, en esta tesis se ha propuesto el desarrollo de un marcador de la enfermedad por medio de estos dos enfoques.

Para desarrollar un marcador basado en la conectividad funcional a partir de EEGs, en primer lugar modelamos los EEGs por medio de grafos donde los nodos son los electrodos y las aristas representan los valores de una determinada medida de conectividad de entre las 4 que se han evaluado. Para llevar a cabo este proceso de forma óptima, desarrollamos una librería, EEGraph, que permite obtener dichos grafos a partir de los registros recogidos por los distintos electrodos distribuidos sobre la cabeza del paciente. Además, esta librería está siendo utilizada por investigadores en diferentes proyectos

en el campo de la neurociencia. En segundo lugar, utilizamos una Graph Convolutional Neural Network (GCNN) para analizar las diferencias de conectividad entre pacientes con EP en diferentes estados dopaminérgicos y controles. Este análisis mostró un aumento significativo en la conectividad cerebral de los pacientes con EP en comparación con los controles, indicando una mayor independencia en las funciones de los cerebros sanos. Las medidas espectrales mostraron una mayor capacidad para generalizar los resultados a casos particulares que las medidas temporales. No se observaron diferencias significativas en la conectividad funcional producidas por la medicación. Los resultados obtenidos por este biomarcador fueron de  $93,08 \pm 0,95\%$  para la tasa de aciertos y un valor de  $95,29 \pm 1,19\%$  para la sensibilidad en el conjunto de prueba para controles frente a pacientes con EP antes de la ingesta de la medicación. Los valores de estas métricas en el caso de controles frente a pacientes con EP tras la toma de la medicación fueron de  $92,04 \pm 2,75\%$  y  $93,53 \pm 4,29\%$  para la tasa de aciertos y sensibilidad respectivamente, demostrando que esta prueba no se ve afectada por las diferencias en la conectividad producidas por la activación dopaminérgica derivada de la levodopa.

Con respecto a las técnicas de PLN, desarrollamos un marcador basado en un modelo BERT para identificar a los pacientes con EP. Para poder aplicar estas técnicas, realizamos una equivalencia entre EEG y texto, dada la naturaleza secuencial de ambos tipos de datos. Se llevaron a cabo distintos experimentos y diversas estrategias de entrenamiento en función de las pruebas electroencefalográficas registradas y número de canales. Los resultados del modelo indicaron que las pruebas de reposo no mejoraron el rendimiento del modelo, pero sí disminuyeron el sobre-entrenamiento de la red, indicando la necesidad de una mayor cantidad de datos para realizar el entrenamiento y generalizar los resultados. Los valores de las métricas que se obtuvieron para este biomarcador fueron, para el conjunto ciego de prueba, un  $86,41 \pm 7,08\%$  para la tasa de aciertos y un  $81,63 \pm 11,58\%$  para la sensibilidad considerando las pruebas motoras con 64 electrodos.

En ambos casos, mediante el estudio de conectividad funcional por medio de grafos y mediante el análisis de señales con técnicas de PLN, los biomarcadores de la EP basados en EEG han mostrado un buen rendimiento sobre un conjunto ciego con una alta sensibilidad en la detección de la enfermedad. Por ello, concluimos que la electroencefalografía ha demostrado ser capaz de registrar el estado de la enfermedad en el paciente y por medio de las técnicas de DL pueden desarrollarse pruebas específicas de la enfermedad que apoyen el diagnóstico clínico y favorezcan su detección en estados prodrómicos.





# Índice de Contenidos

<b>Agradecimientos</b>	<b>I</b>
<b>Abstract</b>	<b>III</b>
<b>Resumen</b>	<b>VII</b>
<b>Acrónimos</b>	<b>XXIII</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Marco clínico de la enfermedad de Parkinson	1
1.1.1. Epidemiología	1
1.1.2. Fisiopatología	1
1.1.3. Criterios diagnósticos	4
1.1.4. Evolución de la EP	6
1.1.5. Tratamiento de la enfermedad de Parkinson	8
1.1.6. Pruebas Complementarias	9
1.2. Electroencefalografía	13
1.2.1. Orígenes del electroencefalograma	13
1.2.2. Análisis e interpretación convencional del EEG	16
1.2.3. Análisis cuantitativo de EEG	18
1.2.4. Aplicaciones del EEG cuantitativo en trastornos neurológicos	20
1.2.5. Electroencefalografía y Parkinson	22
1.2.6. Ventajas y desventajas de la electroencefalografía	23
1.3. Inteligencia Artificial	26
1.3.1. Definición de Inteligencia Artificial	26
1.3.2. Machine Learning	28
1.3.3. Redes Neuronales Artificiales	31
1.3.4. Deep Learning	36
	<b>XI</b>

1.3.5. Aplicaciones de Inteligencia Artificial en el campo de la medicina . . .	43
1.4. Inteligencia Artificial en el diagnóstico de la EP . . . . .	44
<b>2. Objetivos e hipótesis</b>	<b>47</b>
2.1. Objetivo general . . . . .	47
2.2. Hipótesis del estudio . . . . .	47
2.3. Objetivos específicos . . . . .	48
<b>3. Diseño experimental, Pacientes, Materiales y Métodos</b>	<b>51</b>
3.1. Diseño experimental y cálculo de la muestra . . . . .	51
3.1.1. Diseño experimental . . . . .	51
3.1.2. Cálculo de la muestra . . . . .	51
3.2. Pacientes . . . . .	53
3.2.1. Reclutamiento . . . . .	53
3.2.2. Criterios de inclusión . . . . .	53
3.2.3. Criterios de exclusión . . . . .	54
3.2.4. Aspectos éticos . . . . .	55
3.3. Materiales . . . . .	55
3.3.1. Adquisición de electroencefalogramas . . . . .	55
3.3.2. Procesamiento de señales biomédicas . . . . .	56
3.3.3. Revisión sistemática de la literatura . . . . .	58
3.3.4. Desarrollo de técnicas de Deep Learning . . . . .	58
3.4. Métodos . . . . .	58
3.4.1. Adquisición de electroencefalogramas . . . . .	58
3.4.2. Procesamiento de señales biomédicas . . . . .	60
3.4.3. Revisión sistemática de la literatura . . . . .	70
3.4.4. Modelado de EEGs como grafos para el análisis de conectividad por medio GCNN . . . . .	75
3.4.5. Adaptación de técnicas de PLN para el análisis de los EEGs . . . . .	79
3.4.6. Métodos de validación de resultados . . . . .	81
<b>4. Resultados</b>	<b>87</b>
4.1. Análisis de series temporales aplicado a EEGs . . . . .	87
4.1.1. Aplicación de las medidas de análisis de series temporales . . . . .	89
4.1.2. Resultados de las medidas de conectividad . . . . .	91

4.2. Revisión bibliográfica de las técnicas de ML aplicadas a los EEGs de pacientes con EP . . . . .	96
4.2.1. Revisión sistemática desde un punto de vista clínico . . . . .	96
4.2.2. Revisión sistemática desde el punto de vista de las ciencias computacionales . . . . .	105
4.2.3. Resultados que contribuyen al desarrollo de esta tesis . . . . .	114
4.3. Conectividad funcional en pacientes con EP . . . . .	116
4.3.1. Diseño y arquitectura de la GCNN . . . . .	121
4.3.2. Resultados de los estudios de conectividad relacionados con controles . . . . .	124
4.3.3. Resultados del estudio de conectividad de los estados dopaminérgicos . . . . .	132
4.3.4. Generalización de los resultados . . . . .	134
4.4. Aplicación de las técnicas de PLN para el diagnóstico de la EP . . . . .	138
4.4.1. Clasificador de la EP basado en BERT . . . . .	140
4.4.2. Resultados en el diagnóstico de la EP . . . . .	141
4.5. Contribuciones derivadas de esta tesis . . . . .	143
4.5.1. Artículos científicos . . . . .	143
4.5.2. Conferencias y congresos . . . . .	144
<b>5. Discusión</b>	<b>147</b>
5.1. Estudio mediante análisis de series temporales . . . . .	147
5.1.1. Estudio comparativo entre pacientes con EP y controles . . . . .	148
5.1.2. Afectación de la levodopa . . . . .	149
5.2. Análisis bibliográfico . . . . .	150
5.3. Conectividad funcional mediante teoría de grafos . . . . .	151
5.3.1. Estudio de conectividad asociado con controles . . . . .	151
5.3.2. Estudio de conectividad para los estados dopaminérgicos . . . . .	153
5.3.3. Generalización de los resultados . . . . .	153
5.4. Modelo BERT en el diagnóstico de la EP . . . . .	154
5.4.1. Experimentos con 28 canales . . . . .	154
5.4.2. Experimentos con 64 canales . . . . .	156
5.5. Discusión global . . . . .	157
5.6. Trabajo Futuro . . . . .	161
<b>6. Conclusiones</b>	<b>163</b>

---

<b>A. Tablas auxiliares MDS</b>	<b>167</b>
<b>B. Protocolo de Validación de EEGs</b>	<b>169</b>
<b>C. Protocolo de Limpieza de EEGs</b>	<b>173</b>
<b>D. Tablas de la Revisión Sistemática desde el Punto de Vista Médico</b>	<b>181</b>
<b>E. Tablas de la Revisión Sistemática desde el Punto de Vista de las Ciencias Computacionales</b>	<b>193</b>
<b>F. Detalles del software de EEGraph</b>	<b>225</b>
F.1. Paquetes requeridos para su instalación . . . . .	225
F.2. Datos de entrada . . . . .	226
F.3. Construcción de nodos . . . . .	226
F.4. Construcción de aristas . . . . .	227
F.5. Medidas de conectividad . . . . .	227
F.6. Datos de salida . . . . .	228
F.7. Guía de usuario de EEGraph . . . . .	230
<b>Financiación</b>	<b>233</b>
<b>Bibliografía</b>	<b>235</b>

# Índice de figuras

1.1. Prevalencia estandarizada por edad de la EP por cada 100.000 habitantes, por lugar de residencia en 2016. Los países que se muestran con colores cálidos tienen una mayor tasa de prevalencia estandarizada por edad que los países mostrados con colores fríos. . . . .	2
1.2. Proyecciones de dopamina de los circuitos córtico-estriales y frontal-estrial.	3
1.3. Se muestra el sistema de estadificación de Braak de la enfermedad de Parkinson, indicando los lugares de inicio en el bulbo olfatorio y la médula oblonga, hasta la posterior infiltración de la patología de Lewy en las regiones corticales. . . . .	4
1.4. A. Muestra un esquema del cambio de polaridad en la membrana del axón de una neurona dando lugar a la propagación del potencial de acción. B. Registro real de un potencial de acción, que muestra los procesos de polarización, despolarización y valor umbral a partir del cual se propaga el impulso. . . . .	14
1.5. Registro de uno de los primeros electroencefalogramas grabados por Hans Berger. . . . .	15
1.6. Esquema de un dispositivo de electroencefalografía junto con un ejemplo de electroencefalograma registrado. . . . .	16
1.7. Ejemplo de EEG de un paciente de 19 años con epilepsia. . . . .	17
1.8. Placa conmemorativa de la celebración de las conferencias <i>The Dartmouth Summer Research Project of Artificial Intelligence</i> en 1956. . . . .	27
1.9. Esquema de la estructura jerárquica de los subcampos de la IA y su fecha fundacional. . . . .	28
1.10. Esquema con algunos modelos de ML etiquetados por el tipo de procesamiento y tipo de aprendizaje. . . . .	30
1.11. Esquema del funcionamiento de una neurona, indicando los pasos de suma ponderada y función de activación. . . . .	32

1.12. Ejemplo de funciones de activación discontinua. Se muestra la función escalón y la función signo. . . . .	32
1.13. Ejemplo de funciones de activación continua. Se muestran las funciones lineal, sigmoide, tangente hiperbólica y ReLU. . . . .	33
1.14. Ejemplo de arquitectura de una red neuronal, indicando la capa de entrada a la red, la cada de salida y las capas ocultas. . . . .	34
1.15. Ejecución de un filtro sobre una imagen representada por una matriz. . . . .	38
1.16. Ejemplo de una arquitectura CNN. . . . .	39
1.17. Grafo obtenido por medio de la librería EEGraph de un paciente con epilepsia para la medida de la correlación de Pearson con el umbral especificado por defecto. . . . .	40
1.18. Arquitectura de un modelo Transformer. . . . .	41
3.1. Ejecución del software G*Power para el cálculo del tamaño de la muestra. . . . .	52
3.2. Representación esquemática de la distribución de los electrodos de un casco de electroencefalografía de 64 canales. . . . .	56
3.3. Representación visual del grafo obtenida mediante EEGraph, para el ejemplo de un EEG aportado por la librería. . . . .	78
3.4. Representación matricial del grafo anterior obtenida mediante EEGraph. . . . .	79
3.5. Tamaño de ventana y paso empleados en el modelo BERT para realizar la equivalencia texto-EEG. . . . .	80
3.6. Métricas consideradas para la evaluación de los resultados de los modelos junto con su definición. . . . .	82
3.7. Ejemplo que muestra cómo se realiza la selección de subconjuntos durante la ejecución del método $k$ -fold CV. . . . .	85
4.1. Representación esquemática del montaje de EEG de 64 canales. El círculo interior representa la región de interés en nuestras medidas. Los electrodos amarillos corresponden al hemisferio izquierdo (lado dominante) y los electrodos naranjas corresponden al hemisferio derecho (no dominante). Se han elegido los electrodos centrales (grises) como referencia en las medidas de asimetría. Los electrodos azules se han eliminado sistemáticamente del análisis. . . . .	89

4.2. Distribuciones de los coeficientes de correlación por pares de la entropía espectral en el intervalo (1 – 30) Hz, para los controles (verde) y pacientes antes de la ingesta de su medicación (rojo) en los tres registros de electroencefalográficos. El área amarilla representa los valores con una correlación positiva. *** $p < 0,001$ . . . . .	91
4.3. Coeficiente de correlación en el dominio temporal (paneles de la izquierda). Mediana y std de la coherencia espectral (paneles de la derecha). Los resultados se muestran para cada prueba registrada. Las líneas discontinuas en los paneles de la derecha indican la línea de base en estado de reposo, y las flechas reflejan el cambio producido por la activación motora. *** $p < 0,001$ . . . . .	92
4.4. Distribución de las divergencias de Kullback-Leibler de la entropía espectral entre los electrodos de ambos hemisferios (izquierdo y derecho). Los resultados se muestran para las tres pruebas electroencefalográficas. *** $p < 0,001$ . . . . .	94
4.5. Comparación del efecto de la levodopa sobre: Coeficiente de Correlación de la Entropía Espectral (panel izquierdo), Divergencia de Kullback-Leibler (panel central) y Mediana y std de la coherencia de la señal (panel derecho), para el estado de reposo, la pruebas de FT. * $p < 0,05$ , ** $p < 0,01$ , *** $p < 0,001$ . . . . .	95
4.6. Diagrama PRISMA de la revisión bibliográfica realizada desde el punto de vista clínico. . . . .	97
4.7. Número de artículos seleccionados que satisfacen cada uno de los ítems de la lista de verificación presentada en la Tabla D.1. . . . .	98
4.8. Diagrama circular que indica los tipos de modelos usados. En rojo se muestran aquellos con procesamiento simbólico y en azul los modelos con procesamiento subsimbólico. . . . .	101
4.9. Forest-plot con la Diferencia de Medias Estandarizada de la exactitud o accuracy en ambos modelos, SVM and KNN, calculada usando la $g$ de Hedges ponderada según la inversa de la varianza. . . . .	102
4.10. Diagrama PRISMA de la revisión bibliográfica desde el punto de vista de las ciencias computacionales. . . . .	106
4.11.A: Diagrama de barras acerca del número de artículos publicado por año teniendo en cuenta la última fecha de publicación. B: Gráfico circular con la distribución de los artículos seleccionados según el país y continente asociado a la primera filiación del primer autor. . . . .	107

---

4.12.A: Modelos considerados en los artículos que componen la revisión junto con el número de apariciones indicando su pertenencia al grupo de procesamiento simbólico o subsimbólico. B: Diagrama circular del tipo de modelo utilizado, especificando las técnicas de ML no-neuronales y aquellas que pertenecen al subgrupo de ANN. . . . .	113
4.13. Se muestran las posibles combinaciones entre los conjuntos de datos incluidos en el estudio, estos son, pacientes en estado ON (post LD), pacientes en estado OFF (pre LD) y controles. . . . .	118
4.14. Representa la arquitectura del modelo basado en BERT para el diagnóstico de la EP. Para cada capa se muestra la dimensión del tensor de entrada y salida indicando la dirección en la que se producen los procesos. . . . .	141
A.1. Tabla con criterios de diagnóstico clínicos de la EP. . . . .	167
F.1. Ejemplo de salida de un grafo por medio del método “visualize” con la opción de visualización. . . . .	230
F.2. Muestra el código empleado para cargar, modelar y visualizar un EEG mediante la librería EEGraph para la correlación de Pearson, con un threshold de 0,8 y con un tamaño de ventana de 2 segundos. . . . .	231



# Índice de Tablas

1.1. Descripción de los ítems que componen la escala HY en cada uno de sus estados. . . . .	7
3.1. Resumen de las características del dataset especificadas para los EEGs de pacientes con EP y los EEGs de controles. . . . .	57
3.2. Criterios de exclusión e inclusión definidos en la revisión literaria desde el punto de vista de la medicina. . . . .	71
3.3. Criterios de exclusión junto con la razón de la exclusión, definidos en la revisión literaria desde el punto de vista de las ciencias computacionales. . . . .	74
3.4. Descripción de los ítems considerados para el análisis de los artículos. . . . .	75
4.1. Datos demográficos y clínicos de los participantes incluidos en el estudio basado en el análisis de series temporales. Se muestra el valor medio y la desviación estándar de los datos demográficos y clínicos de los participantes. . . . .	88
4.2. Resumen de los resultados, características introducidas a los modelos y filtrado de las señales basados en las Tablas D.2 y D.3 del Apéndice D de los artículos seleccionados. Se incluye el año de publicación de cada estudio. . . . .	104
4.3. Datos demográficos y clínicos de los participantes incluidos en el estudio de conectividad cerebral basado en el modelado de EEGs como grafos. Se muestra el valor medio y la desviación estándar de los datos demográficos y clínicos de los participantes. . . . .	117
4.4. Se presenta el objetivo de cada uno de los modelos especificados que contienen los EEGs de los controles. . . . .	119
4.5. Se presenta el objetivo del modelo que utiliza los EEGs de pacientes con EP en los diferentes estados de activación dopaminérgica. . . . .	119

4.6. Se presenta, para cada medida de conectividad, la cantidad de procesos que se llevarán a cabo en la experimentación. . . . .	120
4.7. Se muestra el valor medio y la desviación estándar de los datos demográficos y clínicos de los pacientes con EP incluidos en el conjunto de prueba para el estudio de conectividad basado en el modelado de EEGs como grafos. También se muestra el p valor obtenido por medio de una distribución t-Student con respecto a los participantes considerados para el conjunto de entrenamiento. . . . .	121
4.8. Se muestra el valor medio y la desviación estándar del proceso de validación, para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 1.a, 2.a y 3.a. . . . .	124
4.9. Resultados de las métricas accuracy, sensitivity, specificity y precision para la medida de la correlación de Pearson en el proceso de prueba de los modelos 1.a, 2.a y 3.a. Los resultados muestran el valor medio y la desviación estándar durante el proceso de validación cruzada. . . . .	125
4.10. Se muestra el valor medio y la desviación estándar para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 1.b, 2.b y 3.b. . . . .	125
4.11. Resultados de las métricas accuracy, sensitivity, specificity y precision para la medida de la entropía de Shannon en el proceso de prueba para los modelos 1.b, 2.b y 3.b. Se muestra el valor medio y la desviación estándar durante el proceso de validación cruzada. . . . .	126
4.12. Se muestra el valor medio y la desviación estándar del proceso de validación, para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 1.c y 1.d. . . . .	127
4.13. Resultados de las métricas accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el experimento de controles vs. pacientes pre LD. Los resultados muestran el valor medio y la desviación estándar durante el proceso de validación cruzada para cada métrica del conjunto de prueba. . . . .	127
4.14. Se muestra el valor medio y la desviación estándar del proceso de validación, para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 2.c y 2.d. . . . .	128

4.15. Resultados de las métricas accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el experimento de controles vs. pacientes post LD. Los resultados muestran el valor medio y la desviación estándar durante el proceso de validación cruzada para cada métrica en el conjunto de prueba. . . . .	129
4.16. Se muestra el valor medio y la desviación estándar de cada una de las folds del CV, para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 3.c y 3.d. . . . .	130
4.17. Resultados de cada uno de los experimentos realizados para las métricas de accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el experimento de controles vs. pacientes en los estados ON y OFF. Los resultados muestran el valor medio y la desviación estándar de cada una de las folds del CV para cada métrica del conjunto de prueba. . . . .	131
4.18. Se muestra el valor medio y la desviación estándar para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 4.a y 4.b. . . . .	132
4.19. Resultados de las métricas accuracy, sensitivity, specificity y precision del proceso de prueba para los modelos 4.a y 4.b. Se muestra el valor medio y la desviación estándar obtenidas durante el proceso de validación. . . . .	132
4.20. Se muestra el valor medio y la desviación estándar del proceso de validación cruzada, para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 4.c y 4.d. . . . .	133
4.21. Resultados de las métricas accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el experimento de pacientes en estado ON vs. pacientes en estado OFF. Los resultados muestran el valor medio y la desviación estándar durante el proceso de validación cruzada para el conjunto de prueba. . . . .	134
4.22. Resultados de las métricas accuracy, sensitivity, specificity y precision para el conjunto de test externo para los modelos 1.a, 2.a y 3.a. Se muestra el valor medio y la desviación estándar durante el proceso de validación cruzada. . . . .	135
4.23. Resultados de las métricas accuracy, sensitivity, specificity y precision para el conjunto de test externo para los modelos 1.b, 2.b y 3.b. Se muestra el valor medio y la desviación estándar durante el proceso de validación cruzada. . . . .	135

4.24. Resultados de las métricas accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el conjunto de test externo para los modelos 1.c y 1.d. Los resultados muestran el valor medio y la desviación.	136
4.25. Resultados de las métricas accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el conjunto de test externo para los modelos 2.c y 2.d. Los resultados muestran el valor medio y la desviación.	137
4.26. Resultados de las métricas accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el conjunto de test externo para los modelos 3.c y 3.d. Los resultados muestran el valor medio y la desviación durante el proceso de validación cruzada. . . . .	137
4.27. Datos demográficos y clínicos de los participantes incluidos en el estudio de diagnóstico de la EP mediante un modelo BERT. Se muestra el valor medio y la desviación estándar de cada una de las medidas consideradas.	139
4.28. Evaluación de los resultados de los modelos de 28 canales con ambas estrategias de entrenamiento, esto es, por un lado, con todas las pruebas electroencefalográficas y, por otro lado, únicamente con las pruebas asociadas a la activación motora. . . . .	142
4.29. Evaluación de los resultados de los modelos de 64 canales con ambas estrategias de entrenamiento, esto es, por un lado, con todas las pruebas electroencefalográficas y, por otro lado, únicamente con las pruebas asociadas a la activación motora. . . . .	143
D.1. Ítems a incluir al informar acerca de modelos predictivos en investigación biomédica. . . . .	183
D.2. Summary of the clinical variables, such as objectives, subjects, EEG recording protocol, EEG cleaning protocol and dataset pre-processing. . . . .	186
D.3. Summary of the model parameters, such as features extracted, models used, model architecture, training and validation methods, results and metric used for the articles included in this review. The columns with the references and objectives of the selected articles are included again for the sake of readability. . . . .	192
E.1. Summary of the objectives, participants, state of PD, EEG pre-processing, features, models used, model parameters, training and validation methods and best results for each article included in this review. Here n stands for the number of patients. . . . .	223

# Acrónimos

ABREVIATURA	SIGNIFICADO
Adam	Adaptative Moment Estimation
AMS	Atrofia Multisistémica
ANN	Artificial Neural Network
AOB	Auditory Oddball
ASR	Artifact Subspace Reconstruction
AUC	Area Under the Curve
BERT	Bidirectional Encoder Representations from Transformers
BN	Bayes Net
BNA	Brain Network Analytucs
BNN	Bayesian Neural Networks
BP-NN	Back Propagation Neural Network
CFS	Correlation based Feature Selector
CNN	Convolutional Neural Network
CNV	Contingent Negative Variation
CSP	Common Spatial Patterns
CV	Cross-Validaction
DBS	Deep Brain Stimulation
DFA	Discriminant Function Analysis
DGHNet	Dynamical System Generated HNet
DL	Deep Learning
DT	Decision Tree
DTF	Directed Transfer Function
EAC	Enfermedad de las Arterias Coronarias
ECG	Electrocardiogram
ECP	Estimulación Cerebral Profunda
EEG	Electroencefalograma
ELM	Extreme Learning Machine

ABREVIATURA	SIGNIFICADO
EMG	Electromiograma
EOG	Electrooculogram
EP	Enfermedad de Parkinson
ERP	Event Related Potential
ESN	Echo State Network
EW	Effective Walking
FAR	False Alarm Rate
FC	agglomerative Feature Clustering
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FKNN	Fuzzy K-Nearest Neighbors
fMRI	Functional Magnetic Resonance Imaging
FN	False Negative
FOG	Freezing Of Gait
FP	False Positive
FPR	False Positive Rate
FRR	False Reassurance Rate
FT	Finger Tapping
FTD	Finger Tapping Derecho
FTI	Finger Tapping Izquierdo
GA	Genetic Algorithms
GAN	Generative Adversarial Network
GCNN	Graph Convolutional Neural Network
GELU	Generative Pre-trained Transformer
GIF	Gait Initiation Failure
GRU	Gated Recurrent Unit
HC	Healthy Controls
H-COG	High Cognition
HNet	Hybrid Network
HOS	Higher Order Spectrum
HY	Hoehn-Yahr
IA	Inteligencia Artificial
IC	Intervalo de Confianza
ICA	Independent Component Analysis
ICA-ENM	Independent Component Analysis by Entropy bound Minimization

ABREVIATURA	SIGNIFICADO
inter-COG	Intermediated Cognition
IPSFAM	Integrated PSFAM
KLD	Divergencia de Kullback-Leibler
KNN	K-Nearest Neighbors
KW	Kruskal-Wallis
LASSO	Least Absolute Shrinkage and Selection Operator
L-COG	Low Cognition
LD	Levodopa
LOO	Leave One Out
LR	Logistic Regression
LSSVM	Least Square SVM
LSTM	Long-Short Term Memory
MAE	Mean Average Error
MARA	Multiple Artifact Rejection Algorithm
MCC	Matthews Correlation Coefficient
MCI	Mild Cognitive Impairment
MDS	Movement Disorders Society
MEG	Magnetoencefalografía
mELM	Morphological ELM
MIBG	Metayodobenzilguanidina
ML	Machine Learning
MLP	Multilayer Perceptron
MRI	Magnetic Resonance Imaging
NB	Naïve Bayes
NC	Normal Cognition
NMT	Neural Machine Translation
NPV	Negative Predictive Value
PCA	Principal Component Analysis
PD	Parkinson's Disease
PET	Positron Emission Tomography
PLI	Phase Lag Index
PLN	Procesamiento del Lenguaje Natural
PLV	Phase Locking Value
PNN	Probabilistic Neural Network
PPG	Photoplethysmography
PPV	Positive Predictive Value

ABREVIATURA	SIGNIFICADO
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PS	Power Spectrum
PSD	Power Spectrum Density
PSFAM	Probabilistic SFAM
PSP	Parálisis Supranuclear Progresiva
QEEG	Quantitative EEG
RA	Respiratory
RAE	Real Academia Española
RBD	REM Behavior Disorder
RBF	Radial Basis Function
ReLU	Rectified Linear Unit
REM	Rapid Eye Movement
RF	Random Forest
RMSE	Root Mean Square Error
RMSProp	Root Mean Square Propagation
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
ROI	Regions Of Interest
RT	Random Tree
SFAM	Simplified Fuzzy ARTMAP
SGD	Steepest Gradient Descent
SL	Simple Logistic
SM	Statistical Models
SMD	Standardized Mean Difference
SPET	Single-Photon Emission Computed Tomography
SPWVD	Smoothed Pseudo-Wigner Ville Distribution
std	Standard Deviation
SVD	Singular Value Decomposition
SVM	Support Vector Machine
TE	Tremor Esencial
TFR	Time-Frequency Representation
TN	True Negative
TP	True Positive
TRODAP	$99m^Tc - TRODAT - 1$
UPDRS	Unified Parkinson's Disease Rating Scale



---

ABREVIATURA	SIGNIFICADO
WPLI	Weighted Phase Lag Index

---



# Capítulo 1

## Introducción

### 1.1. Marco clínico de la enfermedad de Parkinson

#### 1.1.1. Epidemiología

Los trastornos neurológicos son la principal fuente de discapacidad a nivel mundial y, junto con el envejecimiento de la población, hace que las enfermedades neurodegenerativas tengan una mayor prevalencia [1]. Concretamente, la Enfermedad de Parkinson (EP) es el segundo trastorno neurodegenerativo más común después de la enfermedad de Alzheimer, que afecta entre 2-3 % de la población mayor de 65 años [2], y a un 0,3 % de la población total, siendo esta una de las causas más frecuentes de discapacidad física [3].

Como se puede ver en la Fig. 1.1, el envejecimiento de la población es una tendencia global, por lo que la incidencia de la EP tiene una progresión creciente y se estima que en el año 2040 el número de personas diagnosticadas con esta enfermedad llegue hasta los aproximadamente 14 millones de pacientes [4].

Los estudios epidemiológicos, centrados especialmente en evaluar el progreso y la gravedad de la enfermedad, siguen siendo escasos a nivel global, sobre todo en aquellos países con baja incidencia, donde estos estudios son necesarios para comprender la evolución real y orientar los esfuerzos para reducir la carga de la enfermedad.

#### 1.1.2. Fisiopatología

La EP, también conocida como parkinsonismo idiopático, es un trastorno neurodegenerativo que se caracteriza por la pérdida progresiva de las neuronas dopaminérgicas [5]. Las neuronas dopaminérgicas, localizadas en la sustancia negra y otras localizaciones del sistema nervioso central, son las encargadas de producir el neurotransmisor

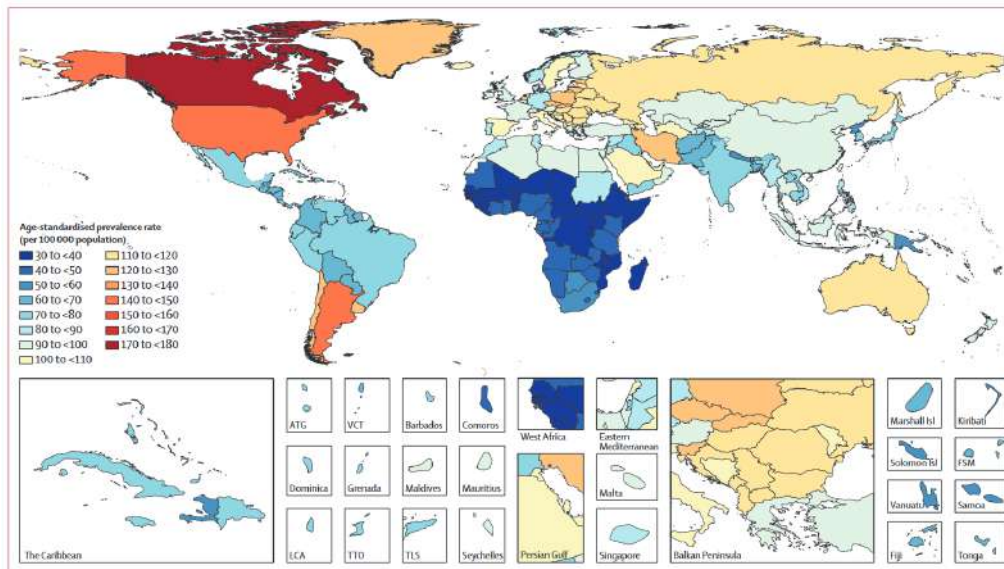


Figura 1.1: Prevalencia estandarizada por edad de la EP por cada 100.000 habitantes, por lugar de residencia en 2016. Los países que se muestran con colores cálidos tienen una mayor tasa de prevalencia estandarizada por edad que los países mostrados con colores fríos. Fuente: [1].

dopamina, que es fundamental en los circuitos de los ganglios basales que regularizan diversas funciones como la conducta motora, la emotividad y afectividad y la comunicación neuroendocrina [6].

Como corresponde a una enfermedad neurodegenerativa, la EP suele desarrollarse entre los 55 y 65 años, estando presente en el 1-2% de las personas mayores de 60 años, aumentando al 3,5% a la edad comprendida entre los 85-89 años [7, 8, 9]. Aunque existen casos juveniles, estos son casos singulares.

Las proyecciones dopaminérgicas desempeñan un papel fundamental en la modulación de los circuitos motores y cognitivos. El deterioro de los circuitos talamo-córtico-estriatales, afectan a los movimientos automáticos, y, en consecuencia, hay una mayor dependencia de los circuitos estriato frontales, que influyen en el control motor de los pacientes con EP. Además, la pérdida de dopamina en el circuito fronto-estriatal contribuye a las deficiencias cognitivas de esta enfermedad [10]. En la Figura 1.2 se puede observar un esquema de estos circuitos dopaminérgicos.

El avance de los síntomas motores y no motores asociados a la EP aumenta con la degeneración progresiva de las neuronas dopaminérgicas. Los síntomas más destacados de esta enfermedad fueron descritos por primera vez por James Parkinson en el siglo XIX, en la publicación científica titulada "An Essay on the Shaking Palsy" en español "Ensayo sobre la parálisis agitante" en 1817 [11]. Posteriormente, estos síntomas fueron

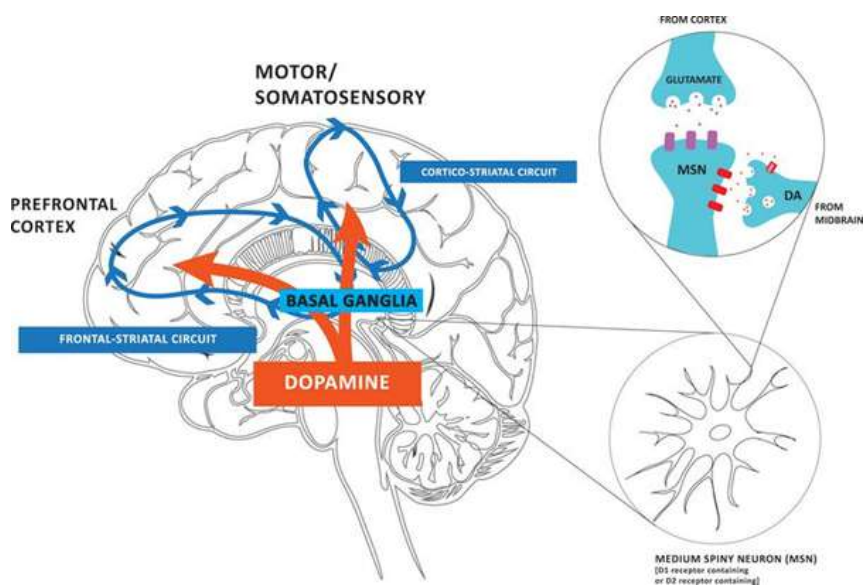


Figura 1.2: Proyecciones de dopamina de los circuitos córtico-estriales y frontal-estrial. Fuente: [10].

descritos por Jean-Martin Charcot [12], quien nombró esta enfermedad como enfermedad de Parkinson. Los síntomas por los que se conoce la enfermedad forman parte del cuadro denominado “parkinsonismo” [13], y se resumen a continuación:

- **Bradicinesia.** Es el término que se asigna a la lentitud de movimientos, y es el rasgo más característico de la EP. Puede manifestarse inicialmente como una reducción del parpadeo o una disminución de la expresión facial, percepción de torpeza o dificultad en el control de la motricidad fina, lentitud en la realización de las actividades de la vida diaria o reducción del balanceo de los brazos al caminar [14, 15].
- **Rigidez muscular.** Se caracteriza por un aumento continuo y uniforme de la resistencia al movimiento. Este signo puede manifestarse a menudo como dolor, normalmente en el hombro o en otras extremidades, alterar la postura y el equilibrio además de limitar el movimiento. El dolor asociado a la rigidez es un síntoma que puede preceder varios años a los signos clínicos más evidentes de la EP [16, 17].
- **Temblo en reposo.** El temblor característico de la EP es un temblor de reposo que suele estar presente en las manos, los labios, la barbilla, la mandíbula y las piernas, y se caracteriza por presentar una frecuencia de 4 a 6 Hz.

Los síntomas descritos son conocidos como síntomas motores de la EP. Sin embargo, estudios de autopsia y de neuroimagen indican que estos signos motores son una

manifestación tardía que se hace evidente cuando el grado de degeneración de las neuronas dopaminérgicas es del 50-70% [18, 19]. Según Braak, la neurodegeneración se inicia en fases muy tempranas y se caracteriza por el depósito de cuerpos de Lewy que progresan desde el tronco del encéfalo hacia el mesencéfalo y finalmente a la corteza [20], como puede verse en la Figura 1.3.

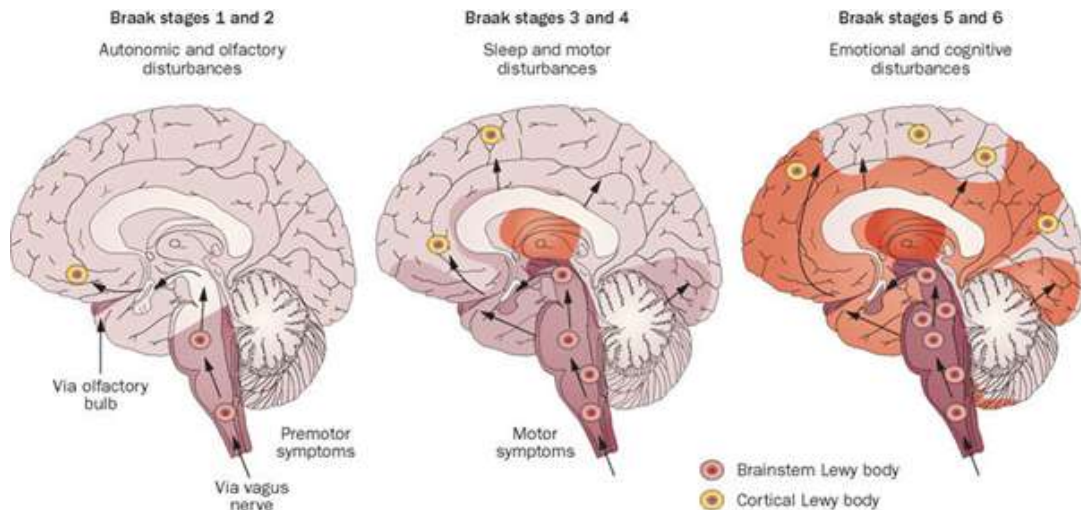


Figura 1.3: Se muestra el sistema de estadificación de Braak de la enfermedad de Parkinson, indicando los lugares de inicio en el bulbo olfatorio y la médula oblonga, hasta la posterior infiltración de la patología de Lewy en las regiones corticales. Fuente: [21].

Según el desarrollo de la evolución de la EP enunciado por Braak, se puede esperar que los síntomas que se presentan de una manera más temprana son habitualmente no motores, por esta razón también se les ha llamado "pre motores", y persisten durante toda la enfermedad. La fase pre-motora o prodrómica puede aparecer entre 12 y 14 años antes de iniciarse los síntomas motores parkinsonianos [22]. Entre los síntomas que caracterizan esta fase encontramos: disfunción olfativa, alteraciones cognitivas leves, síntomas psiquiátricos, trastornos del sueño, disfunción autonómica, dolor y fatiga [23].

Se debe notar, que la denervación se ha descrito como típicamente asimétrica en los estudios funcionales [24]. Consecuentemente, la EP presenta una asimetría de los síntomas en los primeros signos motores que después de los primeros años de enfermedad tiende a volverse bilateral. Esta lateralización, junto con la interacción de la dominancia hemisférica, aún no se han explicado completamente [25, 26].

### 1.1.3. Criterios diagnósticos

El diagnóstico de la EP es clínico y se basa en la identificación de cada uno de los síntomas que caracterizan la enfermedad por parte del médico especialista. Entre

las maniobras clínicas usadas por el neurólogo para evaluar los síntomas propios del parkinsonismo y servir de guía para el diagnóstico encontramos los siguientes:

- **Bradicinesia.** La evaluación en la exploración clínica se basa en que el paciente realice pruebas asociadas a movimientos repetitivos, tales como golpes con los dedos de las manos, con los dedos de los pies y movimientos abriendo y cerrando las manos. De modo que el especialista pueda evaluar la lentitud y la disminución de la amplitud de estos movimientos [27].
- **Rigidez muscular.** La exploración de este síntoma debe incluir una evaluación de la rigidez axial (cuello y tronco) y periférica (extremidades) mediante movimientos pasivos [27].
- **Temblo en reposo.** Su evaluación debe incluir la evaluación del temblor en reposo y el temblor cinético para identificar y distinguir otros tipos de temblores asociados a la EP o TE (temblor esencial) [28]. Por un lado, se realiza la observación del temblor en reposo situando al paciente con las extremidades en posición relajada y mientras camina, así como la evaluación del temblor postural con maniobras como hacer que el paciente extienda los brazos en posición horizontal delante del cuerpo y en posición de "batido de alas". Por otro lado, se evalúa el temblor cinético mediante maniobras como la prueba "de dedo a nariz" [27].

Actualmente, el diagnóstico se realiza clínicamente de acuerdo con los criterios diagnósticos de la Sociedad de Trastornos del Movimiento (MDS clinical criteria for Parkinson's disease) [29], que incluyen los criterios esenciales de bradicinesia con al menos una característica de temblor de reposo o rigidez, síntomas estrechamente relacionados con la neurotransmisión dopaminérgica. Por ello, entre los criterios de soporte al diagnóstico encontramos una respuesta clara y beneficiosa a la terapia dopaminérgica [29]. La Tabla A.1 recogida en el Apéndice A, presenta los criterios de diagnóstico clínico de la MDS [29].

Sin embargo, la precisión en el diagnóstico varía considerablemente en función de la duración de la enfermedad, ya que al ser una enfermedad neurodegenerativa, es más complicado identificarla en etapas tempranas que tras un seguimiento más prolongado de la sintomatología [30]. De modo que la edad del paciente, la experiencia del neurólogo y la evolución del conocimiento sobre la EP [29, 31], son factores clave para realizar el diagnóstico. Debemos resaltar que no existe una prueba específica y objetiva de diagnóstico de la EP. Aunque algunas técnicas como la gammagrafía cardíaca y DaTSCAN (descritas más adelante) pueden apoyar el diagnóstico clínico, que es el principal criterio de diagnóstico de la EP, este está sometido a la subjetividad del especialista y

su precisión se ve influenciada por la experiencia del neurólogo en el diagnóstico de la enfermedad [32]. Según puede verse en la publicación de Rizek et al 2016 [7], el diagnóstico de la enfermedad se ha formalizado mediante los criterios del Banco de Cerebros de la Sociedad de la Enfermedad de Parkinson del Reino Unido [33], con una precisión diagnóstica de hasta el 90 %. Por lo tanto, el diagnóstico clínico se mantiene como la metodología de referencia hasta que se disponga de marcadores diagnósticos validados [29].

#### 1.1.4. Evolución de la EP

La evolución de la enfermedad de Parkinson se caracteriza por un aumento progresivo de la degeneración de las neuronas dopaminérgicas que conlleva a un empeoramiento de los síntomas que padece el paciente [34]. Particularmente, en etapas tardías de la EP, las características motoras y no motoras prevalecen al tratamiento e incluyen síntomas motores axiales como inestabilidad postural, bloqueo de la marcha, caídas, disfagia y disfunción del habla. En estados más avanzados, tras unos 17 años de enfermedad, en torno al 80 % de los pacientes con EP tienen congelación de la marcha e inestabilidad postural, produciendo caídas y lesiones [34]. Continuando con la sintomatología según avanza la enfermedad, encontramos que la demencia es particularmente prevalente y se encuentra presente en el 83 % de los pacientes con EP con un promedio de 20 años de duración de la enfermedad [35].

El estadio y la gravedad de la EP es un factor importante a tener en cuenta para tomar decisiones terapéuticas eficaces [36]. Para conocer el grado de afectación de la enfermedad en el paciente y determinar su progresión [36, 37], el neurólogo realiza una exploración repetida del paciente mediante pruebas semejantes a las establecidas para el diagnóstico, con el fin de determinar su estado global basado en una puntuación o escala relacionada con la evolución de la EP. Las escalas que se usan con más frecuencia para determinar la progresión de la enfermedad son la UPDRS (por sus siglas en inglés “Unified Parkinson’s Disease Rating Scale”) y la escala de Hoehn-Yahr (HY) [33].

La UPDRS se desarrolló inicialmente para evaluar globalmente los principales síntomas de la EP (tanto motores como no motores), permitiendo realizar una evaluación de la discapacidad y el deterioro relacionado con esta enfermedad [38]. A pesar de ser utilizada frecuentemente tanto para la investigación como para la práctica clínica diaria, una revisión posterior de la UPDRS realizada por el Grupo de Trabajo sobre Escalas de Calificación de la Sociedad de Trastornos del Movimiento (MDS) identificó varias deficiencias de la escala original en cuanto a la evaluación de los síntomas no motores [39] además de no permitir la estadificación de la EP [36]. Por tanto, la UPDRS fue modificada y renombrada como MDS-UPDRS. Esta escala consta de 65 ítems divididos en 4 partes.



La parte I evalúa las experiencias no motoras de la vida diaria, y contiene 13 ítems, los 6 primeros son evaluados por un especialista, y los 7 siguientes son un cuestionario para el paciente. La parte II se refiere a las experiencias motoras de la vida diaria, y contiene 13 ítems, todos ellos basados en un cuestionario para el paciente. La parte III es el examen motor, que contiene 33 ítems y es evaluado completamente por un especialista. Finalmente, la parte IV evalúa las complicaciones motoras, contiene 6 ítems y se lleva a cabo por un especialista una vez que el paciente ha empezado a tomar la medicación para la EP. En particular, evalúa dos complicaciones motoras, las discinesias y las fluctuaciones motoras, que están muy relacionadas con la duración de la enfermedad, la dosis de levodopa y la duración del tratamiento con este fármaco [40, 36].

La escala más utilizada para realizar una estadificación de la gravedad de la EP es la escala HY, ya que proporciona una evaluación de la progresión de la enfermedad en varios grados, desde 0 (sin signos de enfermedad) a 5 (estado grave de EP) [41]. Se ha demostrado que las etapas de la escala HY se correlacionan con los estudios de neuroimagen de pérdida dopaminérgica y con los años de progresión de la enfermedad [42, 41]. Consecuentemente, los cambios en el estadio HY de un paciente tienen importancia pronóstica, influyen en la evaluación clínica del especialista y repercuten en el tratamiento. La escala HY modificada se muestra en la Tabla 1.1.

Estadio	Síntomas implicados en cada estado
0.0	No hay signos de enfermedad
1.0	Enfermedad exclusivamente unilateral
1.5	Afectación unilateral y axial
2.0	Afectación bilateral sin alteraciones del equilibrio
2.5	Afectación bilateral leve con recuperación en la prueba de retro pulsión (pull test)
3.0	Afectación bilateral leve a moderada; cierta inestabilidad postural, pero físicamente independiente
4.0	Incapacidad grave; aún capaz de caminar o de permanecer en pie sin ayuda
5.0	Permanece en una silla de ruedas o encamado si no tiene ayuda

Tabla 1.1: Descripción de los ítems que componen la escala HY en cada uno de sus estados. Fuente: [41].

Para establecer el marco clínico actual del paciente, se realizan pruebas en base a estas dos escalas, UPDRS y HY, para ofrecer una visión clínica acerca de varios aspectos de la progresión de la enfermedad. Ambas escalas dependen de la valoración subjetiva de los síntomas y signos del paciente que pueden estar influenciados por el tratamiento farmacológico, pudiendo afectar positiva o negativamente a los resultados de estas pruebas.

Entre las limitaciones de las escalas encontramos que los cambios en las puntuaciones de la MDS-UPDRS contienen una cantidad significativa de varianza de error, destacando la necesidad de desarrollar técnicas o instrumentos más fiables para avanzar en la comprensión de la heterogeneidad en la progresión de la EP [43].

Con respecto a la escala HY, encontramos que las etapas de esta escala mezclan el deterioro y discapacidad además de presentar una evolución no lineal. Como la escala HY se basa en la evolución de los síntomas motores, no ofrece información acerca de los problemas no motores de los pacientes. Además, no capta completamente las deficiencias o la discapacidad de otras características motoras de la EP debido a que esta escala se inclina en gran medida hacia la inestabilidad postural como índice principal de la gravedad de la enfermedad [41].

Por tanto, más allá de la exploración clínica utilizada para apoyar el diagnóstico de la enfermedad, cuyos hallazgos también son utilizados para evaluar la progresión, no encontramos un método de fácil acceso que permita de manera objetiva diagnosticar la enfermedad ni evaluar su evolución en el paciente con EP.

### **1.1.5. Tratamiento de la enfermedad de Parkinson**

La pérdida de las neuronas dopaminérgicas producida en la EP no puede restablecerse, por lo que el tratamiento va destinado a disminuir la sintomatología, es decir, se basa en reducir los síntomas que presenta el paciente para mejorar su calidad de vida. Se lleva a cabo principalmente por medio de fármacos que aumentan las concentraciones de dopamina o estimulan directamente los receptores dopaminérgicos. Por tanto, las deficiencias funcionales producidas por la enfermedad se restauran en gran medida con la administración de levodopa (precursor de la dopamina) y/o con agonistas de dopamina [44, 45]. Por ello es muy importante evaluar de forma objetiva la progresión de la enfermedad en el paciente, ya que la medicación se ajusta en función de los síntomas, que a su vez se correlacionan con los años de enfermedad.

Durante el diagnóstico de la EP el paciente puede iniciar un tratamiento con levodopa, por medio del cual el neurólogo evalúa la respuesta del paciente al medicamento sirviendo como confirmación y apoyo al diagnóstico clínico. La dopamina logra restablecer en gran medida las deficiencias funcionales [44, 46]. Sin embargo, a medida que avanza la EP la variación de respuesta a la medicación produce fluctuaciones motoras y no motoras entre el estado OFF (cuando no hay efectos de la medicación) y el estado ON (cuando los síntomas disminuyen por la acción de la medicación) [34, 47]. Además, hay que tener en cuenta que tras el tratamiento crónico con levodopa (entre 5-10 años) aparecen algunos efectos secundarios, como las discinesias (movimientos involuntarios y anormales) condicionadas por la evolución de la enfermedad [48]. Por ello, el trata-

miento en el transcurso de la EP se puede dividir principalmente en dos grupos. Por un lado, encontramos las terapias de primera línea que se basan en un procedimiento farmacológico, principalmente compuesto por la ingesta de levodopa y agonistas dopaminérgicos [45, 13]. Por otro lado, encontramos las terapias de segunda línea, donde se encuentran los métodos invasivos como perfusión subcutánea continua de apomorfina, perfusión intestinal continua/cabidopa y tratamientos quirúrgicos como la Estimulación Cerebral Profunda (ECP o DBS por sus siglas en inglés “Deep Brain Stimulation”) bilateral del núcleo subtalámico o del globo pálido interno, que son utilizadas en algunos casos cuando el tratamiento farmacológico falla o ha perdido su efectividad [49, 50, 51].

Aunque el tratamiento de la EP es sintomático, la existencia de una prueba de alta precisión para la detección de esta enfermedad en estados tempranos permitiría el tratamiento sintomático precoz y aumentar la calidad de vida de los pacientes.

#### 1.1.6. Pruebas Complementarias

Las características clínicas de la EP pueden compartirse con otros trastornos, creando serias dificultades para determinar un diagnóstico correcto. Entre las patologías que se confunden a menudo con la EP en sus primeras fases encontramos el temblor esencial, los parkinsonismos atípicos como la atrofia multisistémica (AMS) y la parálisis supranuclear progresiva (PSP), la degeneración ganglionar corticobasal, el parkinsonismo inducido por fármacos y la demencia con cuerpos de Lewy. Todas estas patologías comparten algunas características comunes con la EP y por tanto dificultan su diagnóstico, además de que algunos de estos trastornos producen inicialmente una respuesta de leve a moderada al tratamiento con dopamina aumentando la dificultad para realizar un diagnóstico correcto [52].

Para esclarecer el diagnóstico de la EP, existen algunas técnicas complementarias en el campo de la neurología que se utilizan individualmente o en combinación para apoyar el diagnóstico clínico permitiendo descartar otras patologías, siendo esta su principal función. Entre las técnicas más utilizadas se encuentran las pruebas nucleares de imagen como son el PET (por sus siglas en inglés “Positron Emission Tomography”), el SPECT (por sus siglas en inglés “Single-Photon Emission Computed Tomography”) o la gammagrafía cardíaca con MIBG (metayodobenzilguanidina), que muestran las consecuencias fisiológicas indirectas de la denervación dopaminérgica. Sin embargo, estas pruebas son costosas y no siempre son accesibles para los pacientes. A continuación, se especifican algunas de las características de estas pruebas:

- **Gammagrafía cardíaca con 123I-MIBG.** Es una prueba nuclear que refleja la densidad e integridad de las terminaciones nerviosas simpáticas postganglionares. Se

utiliza para discriminar la EP y la AMS [53, 54, 55], ya que en la EP se muestra la zona deteriorada mientras que en la AMS la captación se preserva. Esta prueba tiene una alta sensibilidad, superior al 89 % [56], y una especificidad superior al 77 % [57, 56]. Consecuentemente, es recomendada por la Federación Europea de Sociedades Neurológicas y la Sociedad de Trastornos del Movimiento [58] para el diagnóstico diferencial de los síndromes parkinsonianos.

- **PET.** Es una técnica de imagen nuclear utilizada tanto para evaluar los mecanismos subyacentes de la EP (aumentando la comprensión en torno al diagnóstico diferencial, la progresión de la enfermedad, las complicaciones derivadas de la medicación, y los síntomas no motores de la enfermedad), como para diferenciar la EP de otros trastornos [52]. La PET puede utilizarse como medida de la funcionalidad de la terminal dopaminérgica, cuya pérdida se correlaciona con un aumento de la discapacidad motora. Ha sido utilizada con  $^{18}\text{F}$ -DOPA para estudiar a pacientes con problemas de sueño [59], para monitorizar el resultado de los trasplantes de injertos estriatales en humanos [60], y con  $^{11}\text{C}$ -RTI 32 para evaluar la depresión presente en pacientes con EP [61]. Por tanto, las imágenes PET pueden ayudar a evaluar la enfermedad longitudinalmente [52]. Estas características junto con la alta sensibilidad, y gran resolución espacial y temporal en comparación con otras técnicas de imagen, hacen de esta prueba una herramienta con un gran valor clínico.
- **SPECT.** Desde que se introdujeron las técnicas de imagen molecular in vivo mediante SPECT, el diagnóstico de la EP se hizo más fiable al evaluar las formas degenerativas de parkinsonismo y estimar la pérdida de células dopaminérgicas [62, 63, 64]. Los ligandos SPECT específicos para el transportador de dopamina (tales como FP-CIT,  $\beta$ -CIT, IPT, TRODAT-1) proporcionan un marcador de la degeneración neuronal presináptica permitiendo una detección más temprana [65]. Específicamente, el SPECT con  $^{99m}\text{Tc}$ -TRODAT-1 (TRODAT) se utiliza con este propósito, ya que varios estudios han resaltado la relación entre el nivel de captación de TRODAT y la gravedad de la EP [66, 63, 67]. Esta prueba presenta una alta sensibilidad (del 100 %) y una especificidad en torno al 70 % [62].

Las pruebas de imagen utilizadas para respaldar el diagnóstico de la EP facilitan el diagnóstico de la enfermedad descartando otras patologías con las que comparte la sintomatología. Sin embargo, estas técnicas son complejas, costosas y presentan limitaciones que hacen que no siempre estén accesibles para los pacientes. Algunas de las limitaciones son:

- **Gammagrafía cardíaca con 123I-MIBG.** Es una herramienta muy específica y valiosa utilizada para discriminar la EP de la AMS y acceder a la alteración autonómica postganglionar (cardíaca) [55]. Aunque no siempre es posible distinguir ambas patologías, como se muestra en la publicación de Skowronek et al 2019 [68], hay estudios que asocian la relación 123I-MIBG corazón/mediastino reducida en la EP con una mayor duración de la enfermedad, la aparición de síntomas no motores y el deterioro autonómico [69, 70]. Otros estudios, como los publicados por Matsui et al 2006 y Umemura et al 2013 respectivamente [71, 72], no pudieron confirmar esta relación. Consecuentemente, la interpretación de los resultados de esta prueba debe realizarse por médicos muy experimentados y de especialistas en trastornos del movimiento [42]. Además, aunque la precisión diagnóstica de esta prueba en la EP es alta, la sensibilidad es insuficiente en cuadros de corta duración [73], limitando su uso en el diagnóstico en etapas tempranas de la enfermedad. A la controversia entre los resultados de los diferentes estudios, debemos añadir las restricciones y limitaciones que presenta esta técnica. Entre ellas encontramos los efectos de la medicación concomitante, la disminución progresiva de la captación de 123I-MIBG ( $\sim 3\%/año$  [69]) y el aumento de la degeneración presente en los síndromes de Parkinson, de modo que sigue siendo necesaria la verificación clínica de forma individual.
- **PET.** Las imágenes obtenidas por medio de esta prueba han favorecido el diagnóstico de la enfermedad, diferenciando la EP de otros trastornos, además de evaluar complicaciones derivadas de la medicación, eficacia de una terapia y síntomas no motores [52]. Sin embargo, la PET es una técnica cara y hasta la fecha no está ampliamente disponible para los pacientes. Aunque la PET con  $^{18}F$ -DOPA se ha utilizado en el diagnóstico de la EP, no está claro hasta qué punto es útil este método en el diagnóstico diferencial de otros trastornos del parkinsonismo idiopático [74]. Las futuras aplicaciones de PET dependen del desarrollo de nuevos radiotrazadores específicos.
- **SPECT.** A diferencia de la PET, esta prueba no necesita un ciclotrón ni instalaciones radioquímicas debido a una vida media más larga de los radiotrazadores. Los estudios de SPECT tienen la ventaja de utilizar una producción industrial de trazadores que reduce el coste, siendo accesible a un mayor número de pacientes [75]. El transportador de dopamina en SPECT es lo suficientemente sensible como para detectar la pérdida de neuronas nigroestriales in vivo incluso en las fases preclínicas de la EP [62]. Sin embargo, la precisión diagnóstica de esta prueba en los síndromes parkinsonianos sigue siendo controvertida. La especificidad varió

en función de la edad de los pacientes, pasando de 100% en pacientes jóvenes (edad <55 años), a un 68,5% en pacientes mayores (edad >55 años) [76]. En la publicación de Marshall et al 2009 [77], se realizó un estudio prospectivo y longitudinal con 123I-FP-CIT con 99 pacientes con temblor y/o parkinsonismo durante 3 años, y los resultados mostraron una sensibilidad media del 78% y una especificidad del 97%. En la publicación de Sawana et al 2009 [73], se realizó un estudio de seguimiento durante 2 años utilizando TRODAT en pacientes con síndromes parkinsonianos clínicamente poco claros, en el que se obtuvo una tasa de discordancia de la SPECT en los pacientes del 20%, con una sensibilidad del 100% y una especificidad del 70% en la prueba. En los estudios de Sasannezhad et al 2017, Nagasawa et al 1996 y De la Fuente et al 2003 respectivamente [75, 78, 79], no se observaron diferencias significativas entre los pacientes con EP de inicio tardío y de inicio temprano en cuanto al estadio de la enfermedad y los síntomas, y tampoco se observaron diferencias significativas entre ambos grupos en el análisis visual. Mientras que en las publicaciones de Chou et al 2004, Uhl 1992 y Kaufman et al 1991 [63, 80, 81], afirman que la imagen SPECT de TRODAT puede distinguir con precisión a los pacientes con EP temprana, proponiendo este método para el diagnóstico de pacientes con síntomas y signos tempranos de EP.

Aunque todavía no se recomienda específicamente ninguna técnica de neuroimagen para la práctica clínica de la EP, tanto el PET como el SPECT son los más utilizados debido a su capacidad para detectar cambios in vivo en la función dopaminérgica pre-sináptica dentro del cerebro de pacientes en comparación con los de controles sanos [82]. Aunque la SPECT es un procedimiento más rutinario que la PET, queda patente la controversia acerca de la precisión de esta prueba. Consecuentemente, esta técnica puede ser combinada con otras como la ecografía transcraneal y la prueba de olfato, que pueden ayudar a obtener una mayor precisión en el diagnóstico de la EP [62].

Aún contando con estas técnicas, el proceso de diagnóstico de la EP puede ser largo y estar influido por la subjetividad de las pruebas clínicas. Aunque se utilicen pruebas de neuroimagen, estas no siempre están disponibles para los pacientes, son costosas y algunas de ellas pueden proporcionar dosis de radiación al paciente. El futuro del desarrollo de estas técnicas se encuentra en proporcionar una alta precisión con una mejor resolución de imagen, junto con el desarrollo de nuevos radiotrazadores para ayudar a comprender el papel de los sistemas de neurotransmisores, ya que los agentes de imagen como el transportador de dopamina o los ligandos del receptor D2 evalúan sólo una parte de los aspectos de las neuronas dopaminérgicas [83]. Sin embargo, este nuevo enfoque puede encarecer el proceso para la realización de la prueba haciéndola aún menos accesible para los pacientes. En resumen, entre las limitaciones del diagnóstico

de la EP encontramos:

1. Los síntomas de los pacientes con EP presentan una gran heterogeneidad.
2. Existen síndromes parkinsonianos atípicos y otros trastornos que en ciertas etapas evolutivas comparten muchos rasgos con la EP y su diferenciación requiere una gran experiencia clínica.
3. Las pruebas clínicas que se usan para el diagnóstico son subjetivas.
4. Los síntomas de los pacientes pueden estar encubiertos por alteraciones producidas por diferentes medicamentos.
5. Los métodos complementarios comúnmente utilizados para apoyar el diagnóstico son caros y no siempre están disponibles para los pacientes.
6. La identificación de los signos necesarios para el diagnóstico en fases tempranas es más difícil por lo que se puede producir un retraso del diagnóstico.

Por tanto, se hace patente una vez más la necesidad de una prueba diagnóstica de menor importe económico, rápida y fiable, con una alta precisión que ayude al neurólogo a detectar la enfermedad y que permita el diagnóstico en etapas tempranas y potencialmente en fase prodrómica, ya que eventualmente podría permitir una ventana terapéutica crítica para tratamientos neuroprotectores que detengan o incluso reviertan el proceso neurodegenerativo de la EP [84].

## 1.2. Electroencefalografía

### 1.2.1. Orígenes del electroencefalograma

Según el diccionario Collins, la electrofisiología es la rama de la fisiología que se ocupa de los fenómenos eléctricos relacionados con el cuerpo y sus funciones. Los inicios de la fisiología experimental moderna se atribuyen a Albrecht von Haller [85], con su tratado “Elementa physiologiae corporis humani” [86].

Coincidiendo con el nacimiento de la fisiología, en la segunda mitad del siglo XVIII, comenzaron los primeros estudios relacionados con la electrofisiología, como los estudios de John Walsh de 1773 acerca de la naturaleza eléctrica de las descargas producidas por el pez torpedo [87]. Pero fue Luigi Galvani en 1780, quien descubrió que aplicando una pequeña corriente eléctrica a la médula espinal de una rana muerta, se producían grandes contracciones musculares [88]. Los resultados de sus estudios fisiológicos en las ranas, relacionados con los fenómenos eléctricos, quedaron registrados



en “De viribus electricitatis in motu musculari” [89], llegando a demostrar la existencia de corrientes eléctricas en los músculos de animales, [85].

Esta línea de pensamiento condujo a numerosos investigadores a indagar en el campo de la electrofisiología, favoreciendo el desarrollo de nuevos mecanismos y teorías que dieran explicación al término conocido como “electricidad animal”. Particularmente, cabe señalar la contribución de Julius Bernstein en 1870, quien sentó las bases de la teoría iónica del potencial celular, del que se deriva el potencial de acción, siendo esta una de las contribuciones más importantes de este campo [85].

El potencial de acción es una señal eléctrica que se propaga a lo largo de la superficie de la membrana de una neurona o fibra muscular [90]. La fibra muscular es una célula multinucleada, cilíndrica y con capacidad contráctil y de la cual está compuesto el músculo esquelético [91]. Las neuronas son células nerviosas que tienen excitabilidad eléctrica, esto es la capacidad para responder a un estímulo y convertirlo en un potencial de acción o impulso eléctrico, por medio de un cambio rápido en el potencial de membrana que comprende una despolarización seguida de una repolarización. Un esquema de este fenómeno se muestra en la Figura 1.4. Una vez producido el impulso, este se desplaza debido al intercambio de iones entre el líquido intersticial y el interior de la neurona a través de canales iónicos específicos en su membrana plasmática [92, 90], facilitando la propagación de los impulsos eléctricos a lo largo del axón.

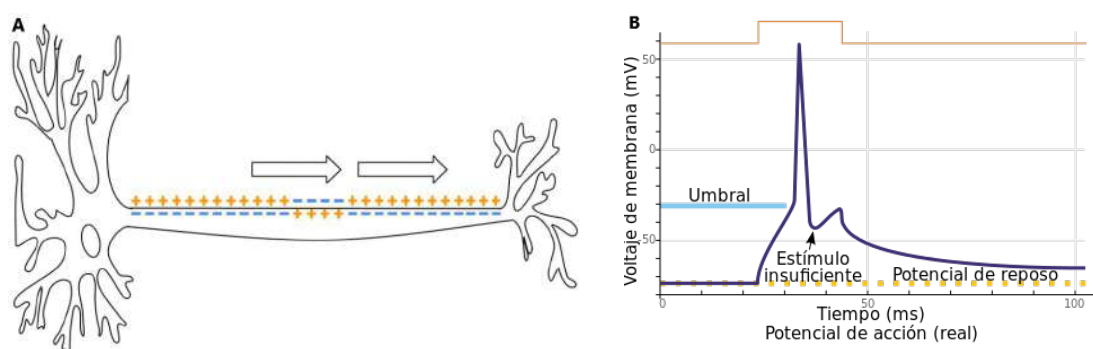


Figura 1.4: A. Muestra un esquema del cambio de polaridad en la membrana del axón de una neurona dando lugar a la propagación del potencial de acción. Fuente: Laurentaylorj, CC BY-SA 3.0, via Wikimedia Commons. B. Registro real de un potencial de acción, que muestra los procesos de polarización, despolarización y valor umbral a partir del cual se propaga el impulso. Fuente: Mememen, CC BY-SA 3.0, via Wikimedia Commons.

Con el inicio de los estudios electrofisiológicos, se desarrollaron nuevos dispositivos capaces de medir estas corrientes eléctricas. En 1872, Gabriel Lippmann diseñó un instrumento capaz de medir las corrientes bioeléctricas, y en 1876, Etienne-Jules Marey desarrolló un sistema capaz de fotografiar las variaciones de corriente registradas con



un electrómetro capilar, obteniendo de este modo, un dispositivo capaz de proporcionar registros gráficos de los fenómenos bioeléctricos [85].

El interés por captar los procesos electrofisiológicos producidos por el cerebro, comenzó en un estudio conjunto por G. Fritsch y Julius Eduard Hitzig en 1870, quienes descubrieron la capacidad de la corteza cerebral humana para ser estimulada eléctricamente. En 1875 Richard Caton, basó sus estudios en los fenómenos eléctricos de los hemisferios cerebrales de conejos y monos. Las investigaciones realizadas por Caton pueden ser consideradas como el inicio del campo de la electroencefalografía [88]. Posteriormente, en 1890, Adolf Beck publicó una investigación sobre la actividad eléctrica espontánea del cerebro de conejos y perros, colocando electrodos directamente en la superficie del cerebro para probar la estimulación sensorial, que le condujo al descubrimiento de las ondas cerebrales [93].

Pero no fue hasta 1912, cuando Vladimir Vladimirovich Pravdich-Neminsky publicó el primer electroencefalograma animal (de un perro), junto con la respuesta a un estímulo externo [94]. Otras investigaciones de este campo emergieron, como la de Napoleón Cybulski y Jelenska-Macieszyna en 1914, que realizaron los registros electroencefalográficos de convulsiones inducidas experimentalmente. En 1924, ampliando el trabajo realizado previamente en animales por Richard Caton, el fisiólogo y psiquiatra alemán Hans Berger registró el primer electroencefalograma humano mostrando ondas que cambiaban cuando el paciente realizaba diferentes tareas cognitivas o experimentaba cambios en las emociones o en la estimulación sensorial [95]. La Figura 1.5 muestra uno de los primeros registros que realizó Berger. También desarrolló el dispositivo conocido como electroencefalograma, descrito como “uno de los desarrollos más sorprendentes, notables y trascendentales de la historia de la neurología clínica” [96], dando origen a la técnica que conocemos actualmente. Berger fue el primero en describir las diferentes ondas o ritmos que estaban presentes en el cerebro del paciente, como el ritmo de ondas alpha y su supresión cuando el sujeto abre los ojos, además de estudiar la naturaleza de las alteraciones del EEG en enfermedades cerebrales como la epilepsia.

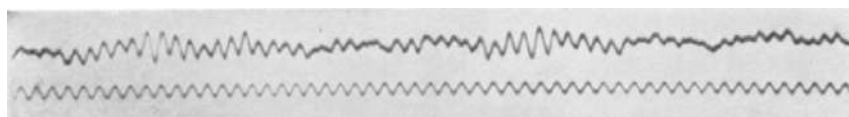


Figura 1.5: Registro de uno de los primeros electroencefalogramas grabados por Hans Berger. Fuente: [97].

Dentro de la electroencefalografía encontramos dos vertientes, una basada en una metodología invasiva, por medio de electrodos intracraneales, y otra no invasiva que registra la actividad por medio de electrodos de cuero cabelludo situados en diferentes

posiciones de la superficie de la cabeza del paciente. En este trabajo nos centraremos en la técnica no invasiva. Cada uno de los electrodos registra la actividad de las neuronas situadas en la corteza cerebral (neuronas piramidales) generando una señal por cada electrodo, asociada a la zona en la que se encuentra situado, con las características medidas. El conjunto de estas señales compone el electroencefalograma (EEG), y representa un mapa de la actividad eléctrica del cerebro del paciente. Un ejemplo de este dispositivo se muestra en la Figura 1.6.

### SCHEMATIC DIAGRAM OF AN EEG MACHINE

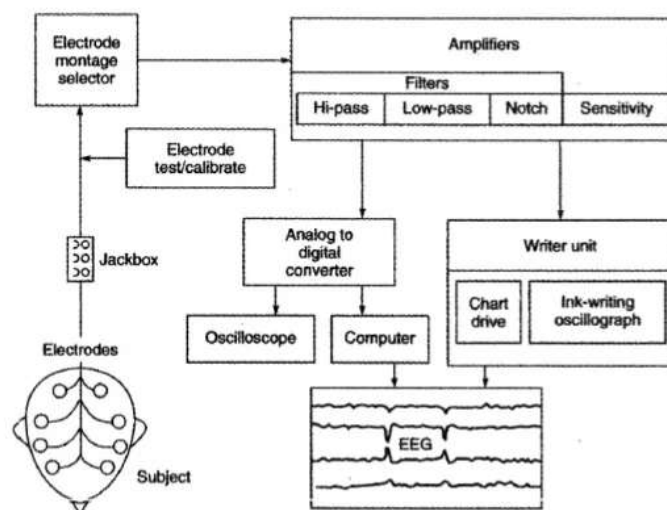


Figura 1.6: Esquema de un dispositivo de electroencefalografía junto con un ejemplo de electroencefalograma registrado. Fuente: [98].

#### 1.2.2. Análisis e interpretación convencional del EEG

Tras el nacimiento de la electroencefalografía, sus múltiples y crecientes aplicaciones promovieron el uso clínico de esta prueba, favoreciendo el desarrollo de dispositivos cada vez más sensibles a las señales electrofisiológicas del cerebro. Su aplicación en pacientes con enfermedades neurológicas comenzó en 1935 con los estudios de Gibbs, Davis y Lennox, en pacientes con epilepsia [99], pero pronto se extendió a otros trastornos, entre ellos encontramos la descripción del sueño REM (por sus siglas en inglés “Rapid Eye movement”) en 1953 por Aserinsky y Kleitman y, en la década de 1980, también se utilizó en el campo de la psiquiatría.

El análisis y la extracción de información a partir de las señales de los EEGs se realiza

de forma visual, buscando en las señales posibles alteraciones que se relacionen con alguna anomalía. Por ello, su principal aplicación se centró en la epilepsia [100], ya que las lecturas del registro electroencefalográfico durante el periodo del ataque muestran grandes alteraciones fácilmente identificables [101]. Un ejemplo del EEG de un paciente durante un ataque epiléptico se muestra en la Figura 1.7. Además de utilizar el EEG para el diagnóstico de la epilepsia, la electroencefalografía también era utilizada para el diagnóstico de tumores, derrames cerebrales y otros trastornos cerebrales focales [102]. Con la llegada de las técnicas de imagen anatómica, se limitó el uso de esta prueba, aunque sigue utilizándose en investigación y como prueba clínica de algunas afectaciones.

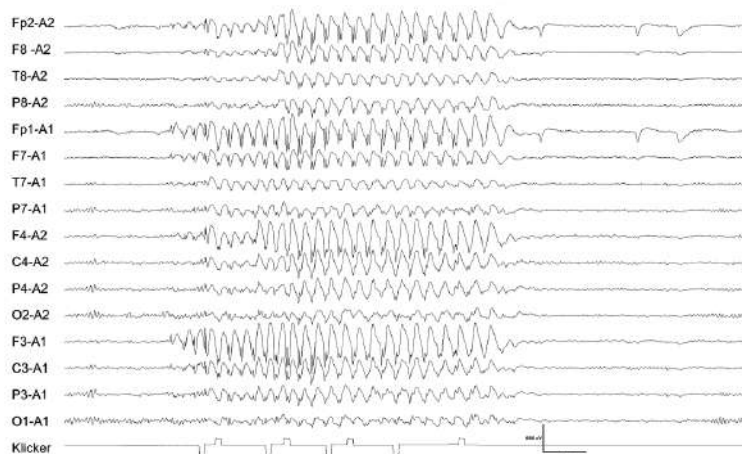


Figura 1.7: Ejemplo de EEG de un paciente de 19 años con epilepsia. Fuente: [100].

Aunque el EEG dispone de una baja resolución espacial en comparación con las pruebas de imagen mencionadas, es una de las pocas técnicas disponibles que ofrecen una gran resolución temporal (de milisegundos), además de permitir realizar los registros de la actividad cerebral en tiempo real. Consecuentemente, hay veces que es necesario realizar la grabación de esta actividad durante varios días para determinar alguna anomalía [103].

El análisis visual del registro electroencefalográfico es un proceso largo y meticuloso, especialmente en EEGs de larga duración, como es el caso de los EEGs registrados para el análisis del sueño [104]. Este tipo de análisis limita las aplicaciones de esta prueba al añadir subjetividad al diagnóstico, ya que las alteraciones visuales son interpretadas por el especialista en función de sus conocimientos y experiencia.

Para minimizar la subjetividad en la interpretación del registro electroencefalográfico, se han ideado métodos para detectar alteraciones en las señales. Entre estas metodologías encontramos:

- **Potenciales evocados.** Consiste en registrar el patrón específico que presenta el sujeto tras producirse un estímulo somatosensorial, auditivo o visual [105].
- **Potenciales relacionados con eventos.** Investigan las fluctuaciones de potencial vinculadas a un evento sensorial, cognitivo o motor específico [106]
- **Contenido espectral.** Analiza las ondas cerebrales que se obtienen al pasar las señales del EEG al dominio de la frecuencia. Estas oscilaciones y su sincronización han sido relacionadas con muchas funciones cognitivas, como la transferencia de información, la percepción, el control motor y la memoria [107, 108, 109, 110].

### 1.2.3. Análisis cuantitativo de EEG

Marc Nuwer, en su estudio de 1997 [111], introdujo el concepto de EEG digital, adquisición electroencefalográfica registrada por medio de ordenadores que almacena la información de las señales en formato digital. Esto permite utilizar técnicas de análisis de señales, como la aplicación de filtros y cambios de escala que facilitan la extracción de la información. Además, el almacenamiento digital permite el acceso a distancia, favorece la conservación de la prueba y permite tratar el conjunto de las señales del EEG de forma conjunta [111].

Marc Nuwer también introdujo el concepto de EEG cuantitativo o QEEG (por sus siglas en inglés “Quantitative EEG”), siendo este un tipo de análisis de las señales electroencefalográficas que consiste en realizar un registro digital para después procesar, transformar y analizar las señales del EEG mediante algoritmos matemáticos [112]. Esta nueva metodología proporciona nuevas técnicas de extracción de características, tanto en los dominios temporal como espectral, como el análisis de bandas específicas de frecuencia, el estudio de la complejidad de la señal, el análisis de la conectividad de los canales y el análisis de patrones [112].

El análisis realizado mediante QEEG elimina la subjetividad presente en esta prueba al descartar el análisis visual convencional realizado por el facultativo. Esta técnica permite la aplicación de un amplio abanico de métodos para extraer las características del EEG que van más allá de la inspección visual. Particularmente, en 1929 se sugirió que los cambios en los ritmos eléctricos del cerebro tendrían valor diagnóstico [113], facilitando la interpretación clínica de esta prueba. Esto favoreció la aplicación del EEG, no sólo como prueba diagnóstica, sino como una prueba complementaria en otros trastornos, aportando información adicional con el fin de obtener un diagnóstico preciso, una evaluación correcta de la gravedad de la enfermedad y una evaluación específica de la respuesta al tratamiento [112], aunque su uso en este sentido no se extiende actualmente más allá de la epilepsia y las alteraciones de consciencia.

Hay varios tipos de análisis que se pueden llevar a cabo por medio del QEEG. Con el EEG digital, el registro queda almacenado como una serie temporal, esto es una serie de números que representa el voltaje registrado en función del tiempo en el que se toma el registro. Este formato facilita la detección automática de eventos, análisis de fuentes (comparando los valores registrados del EEG con modelos predeterminados permitiendo la localización y orientación de posibles fuentes del pico analizado), análisis de la frecuencia de las señales (transformando los datos registrados del EEG a una representación de su espectro de frecuencias), mapas cerebrales (representando visualmente el valor de algún parámetro característico en una imagen de la cabeza del paciente) o análisis por medio de parámetros estadísticos [111]. Las diferentes técnicas que pueden emplearse para llevar a cabo este análisis pueden dividirse en dos grupos, aquellas que analizan las señales en su dominio temporal y las que lo hacen en su dominio espectral.

- **Análisis en el dominio temporal.** Relaciona los valores del potencial registrado en función del tiempo proporcionando la dinámica de la serie. Entre estas técnicas encontramos la identificación de eventos, detección de alteraciones en las señales, búsqueda de patrones repetitivos, estudio de la simetría entre señales y grado de aleatoriedad, aplicación de técnicas de series temporales e implementación de medidas de conectividad. Particularmente, los estudios de conectividad tienen una gran importancia dentro del campo de la neurociencia, ya que permiten evaluar la organización tanto anatómica como funcional del cerebro para detectar anomalías [114].
- **Análisis en el dominio espectral.** Consiste en estudiar las señales del EEG por medio de su espectro de frecuencias a través de la Transformada de Fourier. Este enfoque facilitó en gran medida la interpretación del EEG, permitiendo descubrir otros ritmos de las bandas espectrales, como el caso de la banda theta identificada en 1943 [113].

El análisis de frecuencias se realiza sobre diferentes intervalos o bandas, que se relacionan con diversas funciones cerebrales para estudiar características específicas de cada una de ellas. Las bandas más representativas son:

- **Delta (0,5 - 4) Hz.** Este ritmo se caracterizan por tener una gran amplitud. Por medio de la actividad delta se puede localizar lesiones focales cerebrales, ya que surge en regiones alteradas por lesiones, y predomina en las etapas profundas del sueño [113].
- **Theta (4 - 8) Hz.** Estas ondas se relacionan con varios aspectos de la formación

de la memoria [115], y con el control cognitivo, al estar presentes en múltiples trastornos psiquiátricos [116].

- **Alpha (8 - 12) Hz.** Las oscilaciones alpha fueron los primeros ritmos cerebrales registrados en humanos por Berger [117]. Están presentes cuando el paciente se encuentra en estado de relajación con ojos cerrados o de poca actividad mental y tienen un predominio occipital [97].
- **Beta (12 - 30) Hz.** El ritmo beta está relacionado con el estado de vigilia. Fue descubierto por Berger al registrar el EEG de un paciente cuando este abría los ojos en estado de reposo [117]. Las ondas beta también se asocian con las contracciones musculares y se suprimen antes y durante los cambios de movimiento [118].
- **Gamma (>30) Hz.** Las ondas gamma se relacionan con fenómenos cognitivos como la memoria y la atención. Las alteraciones en este ritmo se han asociado con trastornos cognitivos, como la enfermedad de Alzheimer [119], la epilepsia [120] y la esquizofrenia [121].

La digitalización del EEG y, consecuentemente el QEEG, han permitido que el análisis electroencefalográfico adquiriera una mayor objetividad frente a la valoración visual. Además, esta técnica permite extraer información y características adicionales del EEG sin que sea necesario una representación gráfica clara, extendiendo la aplicación de esta prueba a otros trastornos neurológicos a parte de la epilepsia, tanto en el proceso de diagnóstico como en el control de su evolución.

#### 1.2.4. Aplicaciones del EEG cuantitativo en trastornos neurológicos

La aplicación clínica del QEEG, aunque no tiene el mismo uso generalizado que el EEG, se ha extendido a otros trastornos como ictus, demencia, lesiones cerebrales traumáticas y trastornos psiquiátricos [112].

En la epilepsia, su aplicación ha ayudado a proporcionar un diagnóstico rápido y diferencial entre los distintos subtipos, además de evaluar la respuesta al tratamiento [112]. Varios estudios han detectado alteraciones en las características de las ondas cerebrales producidas por fármacos, de modo que el análisis espectral del EEG se convierte en una herramienta esencial para evaluar sus efectos sobre el paciente y la enfermedad [122]. Por ejemplo, en la publicación de Tedrus et al 2019 [123], se encontró un aumento de la potencia absoluta en todas las bandas de frecuencia en los pacientes epilépticos, y que la coherencia intra e interhemisférica en la banda theta era mayor en los pacientes con epilepsia que en los sanos, permitiendo su clasificación.

Los pacientes con ictus, debido a lesiones focales cerebrales, suelen presentar anomalías en los ritmos cerebrales [112]. La aplicación del análisis del QEEG para el diagnóstico o monitorización de las anomalías producidas por el ictus se utilizó por primera vez en 1984, obteniendo como resultado un aumento de la relación theta/beta en el hemisferio dañado [124]. En el estudio aportado por Kanna et al 2009 [125], encontraron que la potencia relativa alpha de los pacientes con ictus se redujo en ambos hemisferios. Dado que la actividad alpha frontal se asocia con la progresión del deterioro cognitivo [112], la recuperación tras el ataque puede evaluarse analizando esta medida.

En el caso de la enfermedad de Alzheimer, las anomalías del QEEG se suelen identificar en los estadios moderados y avanzados de la enfermedad [112]. Las alteraciones más comunes incluyen las anomalías en las bandas delta, theta y alpha [126]. Particularmente, una reducción de la banda alpha en pacientes con enfermedad de Alzheimer leve se ha propuesto como marcador de diagnóstico de alteraciones cognitivas [127].

El análisis por medio del QEEG también ha sido utilizado para evaluar las conexiones funcionales en los pacientes con depresión [112], al utilizar la asimetría alpha frontal como marcador de la respuesta emocional y de los trastornos emocionales [128]. El análisis convencional del EEG revela anomalías presentes en los pacientes con depresión en el 20 %–40 % de los casos [129], mientras que por medio de la implementación de las técnicas de análisis del QEEG, la precisión en el diagnóstico aumenta, obteniendo valores de sensibilidad comprendidos entre el 72 %–93 % y una especificidad entre el 75 %–88 %, según la Asociación Americana de Neuropsiquiatría [130]. Aunque también se ha utilizado para evaluar el efecto de la medicación, el QEEG no muestra cambios significativos en respuesta al tratamiento antidepresivo, y por tanto, aún no se recomienda para el seguimiento del tratamiento psiquiátrico [131].

Otros estudios han utilizado el análisis de QEEG como herramienta de diagnóstico en los trastornos del aprendizaje [129, 132], mediante la potencia espectral y la coherencia, con una precisión del 46 %–98 % [130], ya que la potencia del EEG puede reflejar la capacidad de procesamiento de la información cortical [129]. También encontramos casos de estudio de trastorno por déficit de atención e hiperactividad, en los que se puede observar un aumento de la potencia en las bandas theta y delta para adultos y niños, y una reducción de la potencia en la banda beta en adolescentes [133].

Como podemos observar, el análisis del QEEG ha permitido ampliar la aplicación de la electroencefalografía a varios trastornos neurológicos. De hecho, en la publicación de Livint et al 2020 [112] se realizó una revisión del QEEG en trastornos neuropsiquiátricos, quedando recogidas otras aplicaciones de esta técnica, como en la lesión cerebral traumática, encefalopatías, ansiedad y la enfermedad de Parkinson como se menciona más adelante. El QEEG representa una herramienta con un gran potencial para mejorar el



diagnóstico clínico y la evaluación de la respuesta al tratamiento. Sin embargo, aunque existe una amplia literatura publicada sobre el análisis del QEEG, todavía hay muchos debates científicos y polémicos sobre su contribución a la práctica clínica, ya que el EEG está influenciado por una serie de factores biológicos (como la edad, el grosor de los tejidos), técnicos (como los equipos electroencefalográficos, electrodos) y alteraciones en las señales por fuentes externas o artefactos [129]. Esto conlleva a la necesidad un mayor estudio y protocolización del análisis del QEEG para su posterior aplicación a la práctica clínica [129]. En conclusión, la electroencefalografía basada en el análisis cuantitativo no se centra en el diagnóstico como prueba individual, sino como prueba complementaria a otras investigaciones para proporcionar información objetiva acerca de un diagnóstico preciso, una evaluación correcta de la gravedad de la enfermedad y una evaluación específica de la respuesta al tratamiento [112].

### 1.2.5. Electroencefalografía y Parkinson

En el caso de la EP, la sintomatología de los pacientes no se ve reflejada en las señales del EEG de manera evidente a simple vista y, por tanto, el análisis visual no ofrece información relevante acerca de la enfermedad o el estado del paciente. Consecuentemente, el análisis convencional ofrece una gran tasa de error y la electroencefalografía no parece la ser la técnica más adecuada para apoyar el diagnóstico de esta enfermedad.

Sin embargo, existen cambios en la conectividad cerebral entre controles y pacientes con EP, que hacen posible mediante el análisis cuantitativo del EEG extraer resultados e información relevante para el neurólogo, que le ayuden tanto en el diagnóstico como para evaluar el estado del paciente. Concretamente, los mayores hallazgos se han hecho mediante el análisis espectral de las señales del EEG. El estudio de estas señales en su espacio de frecuencias puede mostrar anomalías que no son apreciables en el dominio del tiempo mediante una inspección visual, permitiendo la distinción de algunas características diferenciales de los pacientes con EP [134, 135].

Particularmente, en la publicación de Yuvaraj et al 2014 [136], se observa una reducción de la potencia relativa en las bandas delta, theta, alpha, beta y de la potencia absoluta theta, alpha y beta en las regiones anteriores y la asimetría interhemisférica en las bandas theta, alpha, beta en el procesamiento de la felicidad, la tristeza, el miedo, la ira y la sorpresa. En el estudio de Soikkeli et al 1991 [137], se observó que los pacientes con EP sin demencia tienen más actividad theta y frecuencias bajas en comparación con un paciente sano. En la publicación aportada por Han et al 2013 [138], observaron que la potencia relativa de la banda delta se incrementa mientras que en las bandas alpha y beta, la potencia relativa disminuye.



Las diferencias encontradas en los resultados de los artículos mencionados, son una muestra de la dificultad existente en el proceso de identificación de los pacientes con EP por medio de esta técnica. Aunque la heterogeneidad de la sintomatología de los pacientes podría influenciar los resultados de las diferentes medidas, la dificultad para extraer la información a partir de estas medidas reside en las propias características de las señales de EEG: alto contenido en ruido y naturaleza estocástica de las señales.

Las señales electroencefalográficas sin procesar pueden estar contaminadas por señales no relacionadas con la actividad cerebral, denominadas artefactos o ruidos que ocultan las características registradas por las neuronas piramidales [139, 140]. Estos artefactos pueden ser causados por fuentes biológicas y/o fuentes externas. La aparición de artefactos en las señales del EEG produce una atenuación de la señal de origen cerebral lo que podría llevar a una falsa interpretación de la actividad producida por las neuronas y a un diagnóstico erróneo del trastorno cerebral [141].

En el caso de la EP, la aparición de artefactos cobra mayor importancia ya que no hay cambios visibles en las señales de EEG entre pacientes y controles sanos y, por tanto, la alteración de las señales puede producir resultados erróneos y dificultar la interpretación de los resultados.

### 1.2.6. Ventajas y desventajas de la electroencefalografía

El análisis proporcionado mediante el EEG cuantitativo impulsa el uso de la electroencefalografía favoreciendo su aplicación en la práctica clínica diaria más allá de la epilepsia. Particularmente en el caso de la EP, esta técnica podría proporcionar a los neurólogos una prueba alternativa a las pruebas de imagen, favoreciendo el diagnóstico de la EP además de ofrecer información acerca del tratamiento y seguimiento del paciente.

La electroencefalografía es una prueba con grandes beneficios tanto para los pacientes como para los sistemas sanitarios e investigadores en neurociencia. Entre los beneficios que ofrece esta prueba encontramos:

- **No invasiva.** Los electrodos que contiene el casco de EEGs se colocan sobre el cuero cabelludo, por lo que esta prueba no requiere de cirugía ni otras intervenciones invasivas, aumentando el rango de aplicabilidad sobre los pacientes.
- **Alta resolución temporal.** El registro de valores por cada segundo es alto y depende de las características técnicas de cada dispositivo. Las grabaciones suelen comprender entre 256 a 1000 registros por segundo.
- **Bajo coste.** Es una técnica de bajo coste en comparación con las técnicas de

imagen. Además, tanto el mantenimiento como el coste de cada registro también es menor, lo que promueve la utilización de esta prueba.

- **Fácil accesibilidad.** Al ser una prueba de bajo coste, está disponible en la mayoría de centros sanitarios.
- **Rápida adquisición.** El registro se realiza en tiempo real, por lo que la duración de la prueba coincide con la duración del registro del EEG.
- **Alto test-retest.** La electroencefalografía es una prueba con una alta reproducibilidad, esto es que los resultados pueden ser replicados por otros especialistas en otras condiciones externas.
- **Movilidad.** Es una prueba que permite su desplazamiento favoreciendo su acceso a pacientes que no tengan posibilidad de desplazarse hasta el centro sanitario, además de permitir su traslado a países en vías de desarrollo, donde la aparatología de neuroimagen es costosa de implementar.

Estos beneficios hacen que consideremos el EEG como una prueba versátil con capacidad de proporcionar información del estado del cerebro del paciente y evaluar los daños producidos por la enfermedad. Es especialmente relevante su buena resolución temporal, que proporcionan información dinámica de la actividad eléctrica cerebral y de la conectividad. Por esta razón, el EEG se ha registrado en diferentes condiciones fisiológicas, esperando identificar diferentes patrones que ofrezcan información acerca del estado del paciente. En el caso de la EP, en el que la sintomatología del paciente no se refleja mediante cambios visuales de las señales del EEG, es necesario emplear las técnicas de análisis cuantitativo sobre estas señales. Sin embargo, estas técnicas quedan limitadas por las características intrínsecas de estas señales, estas son:

- **Baja resolución espacial.** Los campos eléctricos generados por la actividad de las neuronas, son amortiguados por los tejidos cerebrales situados entre la fuente (las neuronas) y los receptores (los electrodos). Este efecto produce una correlación entre las señales de electrodos cercanos además de una atenuación de las mismas. La actividad registrada por los electrodos recoge las conexiones de las zonas externas del cerebro, de modo que los procesos más profundos quedan atenuados.
- **Baja relación señal-ruido.** La actividad cerebral (en escala de  $\mu V$ ) registrada por los electrodos, contiene múltiples fuentes de ruidos de diversas procedencias, como ruidos ambientales, ruidos fisiológicos y otros ruidos de amplitud similar, que ocultan las características registradas por las neuronas piramidales.

- **Naturaleza estocástica.** Según el diccionario Collins, la estocasticidad se define como “un proceso que implica una secuencia aleatoria de observaciones, cada una de las cuales se considera como una muestra de un elemento de una distribución de probabilidad”. Dicho de otro modo, un conjunto de variables aleatorias ( $X_t$ ) cuya distribución varía de acuerdo al tiempo  $t$ . Además, son de carácter no estacionario y no determinista, esto es que los parámetros estadísticos que describen su dinámica varían con el tiempo [142]. Por lo tanto, para extraer las características esenciales de las señales del EEG, se deben utilizar técnicas avanzadas para el estudio de la dinámica no lineal, que requieren métodos computacionales más costosos [139, 143].

Estas desventajas ofrecen unas fuertes limitaciones que repercuten en la aplicación clínica de esta prueba para detectar enfermedades neurológicas. Sin embargo, existen métodos capaces de hacer frente a estas desventajas.

Para minimizar el impacto del ruido y artefactos de las señales del EEG, existen una serie de procedimientos semi-automáticos que permiten identificar artefactos en las señales para después eliminarlos, realizar filtrados, detectar canales anómalos y minimizar el ruido. Cada uno de estos pasos deben ser evaluados por un especialista, ya que estos procesos de limpieza pueden conllevar la eliminación de componentes relevantes relacionadas con la actividad cerebral de las señales grabadas, conduciendo a un diagnóstico erróneo. Sin embargo, el principal problema para filtrar las señales del EEG es que existen varios protocolos validados sin un estándar definido para realizar la limpieza.

Con respecto a la baja resolución espacial del EEG, cabe señalar que un casco con un mayor número de electrodos ofrecerá una mejor resolución espacial que un dispositivo que contenga menos. De esta forma, aunque haya correlación entre las señales de electrodos cercanos, el aumento en el número de electrodos puede salvar esta deficiencia permitiendo realizar una reconstrucción de la actividad cerebral del paciente.

En cuanto a la naturaleza estocástica de las señales de EEG, para extraer las características esenciales relacionadas con la enfermedad, debemos utilizar técnicas avanzadas de análisis de dinámica no lineal y métodos estadísticos complejos, que requieren la implementación de métodos computacionales más costosos [139, 143]. Entre estas técnicas, podemos destacar el análisis de series temporales. Sin embargo, presentan una alta dependencia con respecto a los procedimientos de limpieza realizados.

Por tanto, para poder extraer las características esenciales relacionadas con la enfermedad a partir de las señales de los EEGs, es necesario utilizar una metodología capaz de salvar estas dificultades. Las técnicas de Inteligencia Artificial permiten aprender relaciones de los datos y llevar a cabo decisiones basadas en dichas relaciones sin conocerlas a priori [144]. De hecho, son especialmente adecuadas para el manejo de

grandes cantidades de datos de manera eficiente y el análisis de información compleja y multimodal, permitiendo detectar y reconocer patrones ocultos en ella e identificar irregulares, siendo también apropiadas para el manejo de datos que presenten alteraciones. Por ello, proponemos estas técnicas para analizar y obtener información de las señales de EEG.

## 1.3. Inteligencia Artificial

### 1.3.1. Definición de Inteligencia Artificial

El concepto de inteligencia ha sido objeto de debate a lo largo de la historia debido a la dificultad de su definición. De hecho, el diccionario de la Real Academia Española (RAE) presenta distintas definiciones [145], como “*capacidad de entender o comprender*”, “*capacidad de resolver problemas*”, “*conocimiento, comprensión, acto de entender*”, o “*habilidad, destreza y experiencia*”. Todas ellas heredan algunos de los postulados de la teoría de las inteligencias múltiples de Howard Gardner [146], quien propuso en 1983 que la inteligencia no se puede identificar unívocamente, sino que existen diferentes competencias intelectuales o inteligencias humanas relativamente autónomas que se pueden combinar de distintos modos (inteligencia lingüístico-verbal, lógico-matemática, visual-espacial, emocional, etc.).

La definición de la rama de las Ciencias de la Computación conocida como *Inteligencia Artificial* (IA) hereda las dificultades de definir Inteligencia. El término como tal fue acuñado en 1956 en la serie de conferencias *The Dartmouth Summer Research Project of Artificial Intelligence* convocadas por John McCarthy, junto con Marvin L. Minsky, Nathaniel Rochester y Claude Shannon [147]. En la Figura 1.8 se muestra la placa conmemorativa situada en el Dartmouth College con motivo de la conmemoración de los 50 años de este acto. En ella se recoge el propósito fundacional de esta disciplina: “*avanzar sobre la base de la conjetura de que cada aspecto del aprendizaje o cualquier otra característica de la inteligencia puede ser en principio descrito de manera tan precisa que se puede crear una máquina que lo simule*”.

Dejando de lado la gran variedad de definiciones acerca del concepto de *Inteligencia* e *Inteligencia Artificial*, aún queda por resolver la cuestión acerca de cuándo una máquina podría considerarse inteligente. En 1950 Alan Turing, considerado el padre de la IA, propuso en su artículo *Computing Machinery and Intelligence* [148] el juego de la imitación, una prueba concreta para determinar si una máquina era inteligente o no, también conocida como Test de Turing. Esta prueba se basa no en una definición de inteligencia o en una métrica de la misma, sino en la comparación, mediante un juego, de las habili-

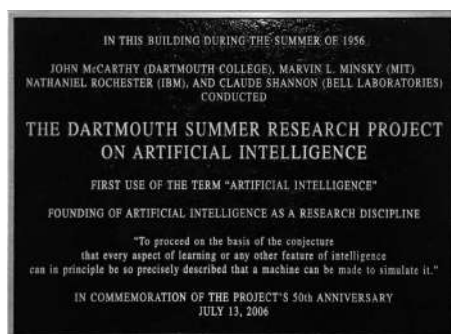


Figura 1.8: Placa conmemorativa de la celebración de las conferencias *The Dartmouth Summer Research Project of Artificial Intelligence* en 1956. Fuente: [147].

dades de una (supuesta) máquina inteligente y un humano. Si ambos muestran el mismo nivel de competencia en el juego, Turing deduce que habrá que conceder el adjetivo de “inteligente” a la máquina. Esta afirmación cuenta con seguidores y detractores, combinando conceptos filosóficos, antropológicos y científicos, como es el caso de Searle con su célebre experimento mental de la habitación china [149], que a su vez generó un intenso debate aún de plena actualidad, materializado en la distinción entre IA fuerte (sistema que puede realizar con éxito cualquier tarea intelectual que pueda realizar un humano) e IA débil (sistema diseñado específicamente para enfocarse en una tarea específica).

Esta incapacidad de medir la inteligencia (en un humano o en una máquina) junto con la dualidad mente-cerebro, puesta de manifiesto por las aproximaciones psicológica y neurocientífica al proceso de la cognición, ha generado las dos corrientes de la Inteligencia Artificial que, turnándose en la historia, han dado lugar a los avances alcanzados hoy en día en este campo:

- Por un lado, encontramos el enfoque racional, basado en el procesamiento por parte de nuestra mente de símbolos mediante funciones lógico-matemáticas, conocido como procesamiento simbólico o IA simbólica.
- Por otro lado, encontramos un enfoque antropológico, centrado en el procesamiento que el cerebro hace de las señales electroquímicas que constituyen el impulso nervioso dando lugar al procesamiento subsimbólico o IA subsimbólica.

La corriente subsimbólica dentro de la IA busca recrear mediante un ordenador la forma en que trabaja el cerebro humano, concretamente las neuronas. El resultado de estos trabajos es la definición y desarrollo de lo que se conoce como Redes Neuronales Artificiales (ANN por sus siglas en inglés “Artificial Neural Networks”), explicadas más en detalle en apartados posteriores. Dentro de la IA subsimbólica se incluyen también otras

aproximaciones de inspiración biológica, como son la computación evolutiva, basada en la replicación y mutación de las cepas víricas [150] o la inteligencia de enjambre, basados en modelos de auto-organización de insectos sociales. Todas estas técnicas de IA subsimbólica que superan los enfoques deterministas tradicionales a expensas de perder precisión y generalización son conocidas como “soft computing” [151].

En general podemos hablar de una jerarquía dentro de la Inteligencia Artificial que queda reflejada en la Fig. 1.9 Dicha figura sirve a su vez para ubicar las técnicas que se han usado en esta tesis para el análisis de los EEGs de enfermos de Parkinson.

Así, el Machine Learning o Aprendizaje Automático es un campo de la Inteligencia Artificial que nace en 1959 y que estudia la habilidad de los sistemas computacionales de aprender sin ser explícitamente programados para ello. De entre las diversas técnicas de Machine Learning las más relevantes y que ahora mismo suponen una revolución en prácticamente todos los aspectos científicos y sociales son las redes neuronales artificiales (ANN), sistemas compuestos por un conjunto de procesadores elementales (las neuronas) conectadas entre sí y cuyo comportamiento está determinado por la topología y pesos de sus conexiones. Sus fundamentos se enraízan en la neurobiología y su origen data de 1943.

Por último, el Deep learning es un subconjunto de arquitecturas neuronales capaces de llevar a cabo abstracciones de alto nivel a partir de datos complejos usando arquitecturas jerárquicas llamadas Redes Neuronales Profundas, considerándose su nacimiento oficial en 2006.

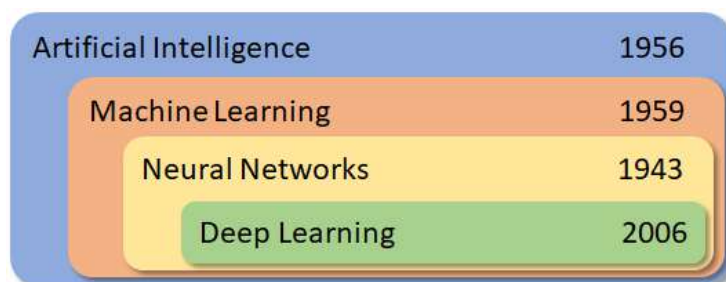


Figura 1.9: Esquema de la estructura jerárquica de los subcampos de la IA y su fecha fundacional. Fuente: elaboración propia.

### 1.3.2. Machine Learning

El aprendizaje automático o Machine Learning (ML por sus siglas en inglés), es una disciplina dentro del campo de la IA encargada de desarrollar algoritmos con la capacidad de generalizar comportamientos y reconocer patrones ocultos en una gran cantidad de datos. Fue definida por Arthur Samuel en 1959 como “campo de estudio que otor-

ga a las máquinas la capacidad de aprender sin ser programados explícitamente para ello” [152]. Por tanto, las técnicas de ML permiten al algoritmo modificar su comportamiento basándose en datos, o en aciertos o errores cometidos previamente, es decir, en la experiencia, para que tenga lugar el proceso de aprendizaje. Un algoritmo de ML no es programado para resolver un problema, sino para aprender, de forma autónoma, a resolverlo. Como subcampo de la IA, el ML hereda las dos corrientes de procesamiento:

- **Procesamiento simbólico.** Engloba los modelos de ML que utilizan procesamiento simbólico como árboles de decisión (DT por sus siglas en inglés “Decision Trees”) o bosques aleatorios (RF por sus siglas en inglés “Random Forest”).
- **Procesamiento subsimbólico.** Comprende aquellos modelos de ML diseñados para estimar las relaciones funcionales entre los datos. Entre ellos están los modelos de clustering, métodos de regresión, redes Bayesianas, sistemas evolutivos, SVM (por sus siglas en inglés “Support Vector Machine”), y las ANN, siendo estas últimas los modelos más utilizados actualmente.

Si bien el aprendizaje es una característica intrínseca a todos los modelos de ML, este se puede llevar a cabo desde diferentes perspectivas. Podemos clasificar los modelos de acuerdo con el tipo de aprendizaje, o con la regla de aprendizaje que implementan, esto es, la función matemática que lleva a cabo la auto-organización de los parámetros del modelo. En lo que respecta a los tipos de aprendizaje, podemos distinguir [153, 154]:

- **Aprendizaje supervisado.** El modelo recibe pares de ejemplos etiquetados (pares entrada-salida) con los que auto-organiza sus parámetros internos de manera que encuentra la forma de, dadas las variables de entrada, asignar la correspondiente etiqueta de salida. La salida puede ser categórica (cualitativa), si el problema es de clasificación, o numérica (cuantitativa), si el problema es de regresión. Los algoritmos más usados dentro de este tipo de aprendizaje son: DT, clasificador de Naïve Bayes, Regresión por mínimos cuadrados, regresión logística, SM (por sus siglas en inglés “Statistical Models”) y métodos “ensemble”. Este tipo de entrenamiento también se puede aplicar a datos no estructurados.
- **Aprendizaje no supervisado.** El algoritmo detecta patrones ocultos en los datos de entrada sin etiquetar (no hay retroalimentación explícita). Dentro de este tipo de aprendizaje encontramos los modelos de clustering, cuyo objetivo es encontrar y diferenciar grupos con características similares o los de reducción de dimensionalidad, que utilizan técnicas matemáticas y estadísticas para convertir el dataset original en uno nuevo con menos dimensiones a cambio de perder información.

Ejemplos de algoritmos de reducción de dimensionalidad son Análisis de Componentes Principales (PCA por sus siglas en inglés “Principal Component Analysis”), Descomposición en Valores Singulares (SVD por sus siglas en inglés “Singular Value Decomposition”) o Análisis de Componentes Independientes (ICA por sus siglas en inglés “Independent Component Analysis”). Además, en este grupo encontramos los modelos de mapas auto-organizativos o mapas de Kohonen [155].

- **Aprendizaje semi-supervisado.** El modelo aprende a partir de un conjunto de datos de los que sólo una parte están etiquetados, pudiendo seguir para ello algún tipo de criterio de supervisión. Juega un papel importante en casos en los que se necesita procesar una gran colección de ejemplos carentes de etiqueta y el etiquetado es costoso o difícil de realizar.
- **Aprendizaje reforzado.** El algoritmo aprende a tomar decisiones mediante un proceso de retroalimentación de prueba y error, con el objetivo de maximizar el beneficio por medio de una serie de refuerzos (recompensas) y minimizar los fallos mediante penalizaciones (castigos). Este tipo de aprendizaje tiene especial importancia en situaciones en las que se desconoce el error preciso cometido, pero se dispone de información global acerca del desempeño del modelo.

En la Figura 1.10 se muestra un esquema con algunos de los modelos más utilizados dentro del campo del ML, etiquetados en función del tipo de procesamiento y tipo de aprendizaje anterior.

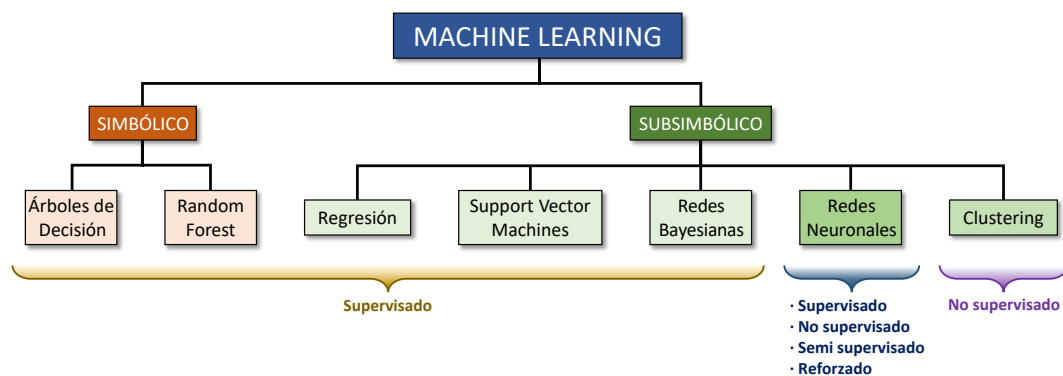


Figura 1.10: Esquema con algunos modelos de ML etiquetados por el tipo de procesamiento y tipo de aprendizaje. Fuente: elaboración propia.

La Fig. 1.10 muestra cómo las redes neuronales artificiales son los modelos que proporcionan una mayor versatilidad en cuanto al tipo de aprendizaje en función de los datos de entrada a la red. Por ello, centraremos el estudio en estas redes además de ser las que utilizamos en el transcurso de esta tesis.



### 1.3.3. Redes Neuronales Artificiales

Una Red Neuronal Artificial se define como un sistema compuesto por un conjunto de elementos de procesamiento simples (neuronas), conectados entre sí (red), cuyo comportamiento está determinado por la topología y los pesos de sus conexiones.

En 1943 Warren McCulloch y Walter Pitts presentaron un modelo teórico basado en neuronas artificiales [156], el cual se considera el primer trabajo dentro del campo de la IA y, en particular, artículo fundacional de la IA subsimbólica. Este modelo se basó en la fisiología y el funcionamiento de las neuronas y sus conexiones en el cerebro humano para recrear estas redes de forma artificial. En él las neuronas presentaban dos estados, “*all*” o “*none*”, esto es, activado y desactivado, en el que cada estado de activación dependía del estado de las neuronas *vecinas*. Este tipo de redes mostraron ser capaces de recrear el comportamiento de los operadores lógicos, permitiendo recrear cualquier función de cómputo aumentando la complejidad de la red [157].

En 1951 Marvin Minsky y Dean Edmonds construyeron SNARC, el primer computador basado en una red neuronal. Posteriormente, en 1958, Frank Rosenblatt desarrolló y simuló la primera red neuronal artificial, evolución del modelo de Warren McCulloch y Walter Pitts, a la que llamó “Perceptron” [158]. Este es considerado como el primer computador capaz de aprender mediante prueba y error.

Sin embargo, en 1969 Marvin Minsky y Seymour Papert demostraron en su libro *Perceptrons: an introduction to computational geometry* [159], que el perceptrón propuesto por Rosenblatt (de una capa), no podía aprender si los patrones de entrenamiento no eran linealmente separables (*Teorema de convergencia del perceptrón*). Además, señalaban que no había razones para suponer que las virtudes del perceptrón de una capa se pudiesen trasladar a la versión multicapa. A esto se sumaba que el algoritmo utilizado para entrenar al perceptrón de una capa no era extensible de manera eficiente a otros tipos de redes más complejas, esto es, con un aumento en el número de capas ocultas. Estos hechos generaron una gran controversia acerca de las posibilidades de las ANN, produciendo, a finales de los años 60, un estancamiento en el desarrollo de estas técnicas [160].

Estas limitaciones fueron solventadas en 1986 por David E. Rumelhart, Geoffrey E. Hinton y Ronald J. Williams en su libro seminal *Learning representations by back-propagating errors* [161], donde proponen una arquitectura más compleja, el Perceptrón Multicapa (MLP por sus siglas en inglés “MultiLayer Perceptron”) y describen formalmente el algoritmo de Backpropagation, que se verá con detalle en siguientes apartados, y que junto con el algoritmo del descenso del gradiente, daban solución a este problema.

El funcionamiento de cada neurona es un proceso de 2 pasos representado en la

Figura 1.11. En primer lugar, cada neurona  $j$  recibe los valores de salida  $x_1, \dots, x_n$  de otras  $n$  neuronas (o los datos de entrada), que son ponderados por los pesos  $w_{1j}, \dots, w_{nj}$ , para realizar la suma ponderada. En este paso se puede introducir un término independiente o bias  $b_j$ . Este conjunto de valores  $\{w_{ij}, b_j\}$  representa un hiperplano mediante el cual se pueden separar dos conjuntos de datos. Los términos  $w_{ij}$  proporcionan la intensidad con la que la neurona  $i$  se ve afectada por la neurona  $j$ , siendo estos los parámetros del modelo que se modifican durante el entrenamiento de la red. En segundo lugar, se obtiene la salida de la neurona,  $y_j$ , aplicando una función de activación al resultado de la suma ponderada llevada a cabo en el paso anterior, representada como  $f$  en la Fig. 1.11. Esta función puede ser de distintos tipos, dando lugar a los distintos tipos de neurona, cada una de ellas con unas propiedades características.

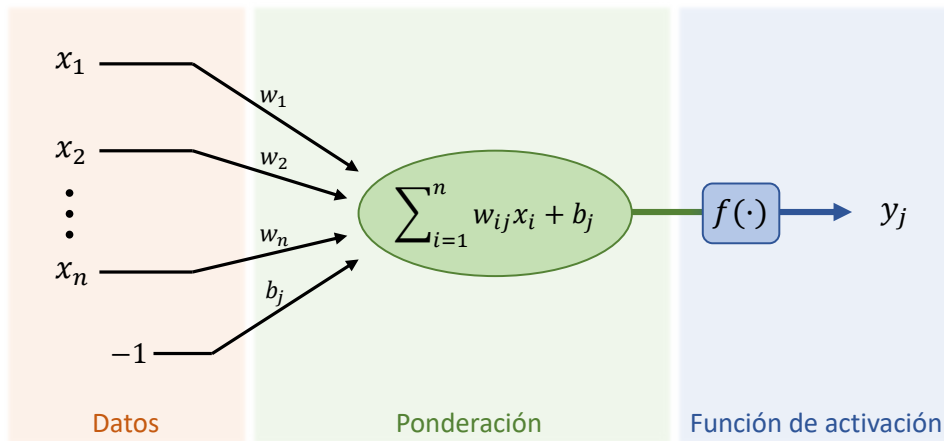


Figura 1.11: Esquema del funcionamiento de una neurona, indicando los pasos de suma ponderada y función de activación. Fuente: elaboración propia.

Las funciones de activación más comunes pueden dividirse en dos grupos, dependiendo de su continuidad. Entre las discontinuas destacan la función escalón y la función signo, mostradas en la Figura 1.12.

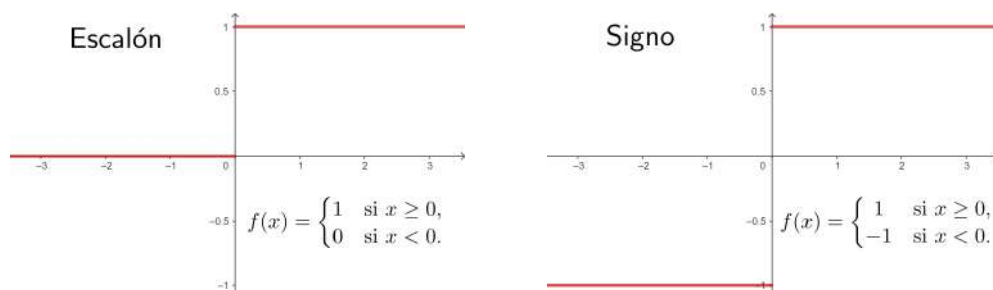


Figura 1.12: Ejemplo de funciones de activación discontinua. Se muestra la función escalón y la función signo. Fuente: elaboración propia.

Entre las funciones de activación continuas más comunes encontramos lineales, sigmoide, ReLU (por sus siglas en inglés “Rectified Linear Unit”) y tangente hiperbólica, que actúan a nivel de componente, y Softmax, que tiene en cuenta la información global del tensor que contiene los datos. Un tensor no es más que una matriz multidimensional, definido por Goodfellow et al 2016 [162] como “una colección de valores ordenados en una retícula regular con una cantidad variable de ejes”.

La función sigmoide, cuya expresión viene dada por

$$f(x) = \frac{1}{1 + e^{-x}}, \quad x \in \mathbb{R}, \quad (1.1)$$

hace que los valores grandes se saturan en 1 y los pequeños en 0, por lo que es muy útil para representar probabilidades independientes, esto es que cada valor se representa por medio de una probabilidad global. La función tangente hiperbólica, por su parte, toma valores entre  $-1$  a  $1$ , y la función ReLU se comporta como una función lineal cuando es positiva y constante a cero cuando el valor de entrada es negativo. Estas funciones se muestran en la Figura 1.13.

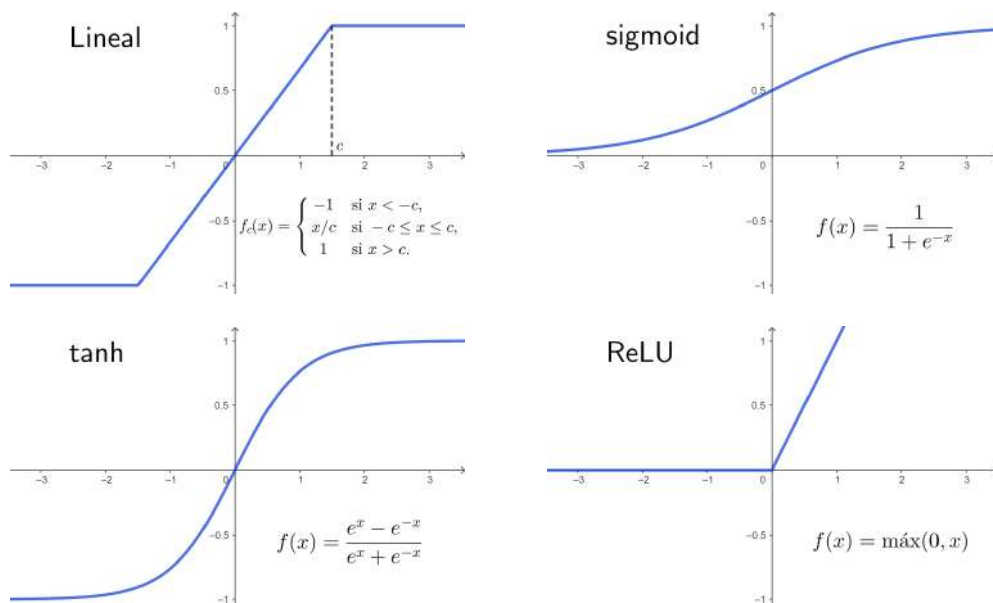


Figura 1.13: Ejemplo de funciones de activación continua. Se muestran las funciones lineal, sigmoide, tangente hiperbólica y ReLU. Fuente: elaboración propia.

La función Softmax, o función exponencial normalizada, permite transformar tensores de números reales en tensores de probabilidades en los que las componentes toman valores entre 0 y 1 y la suma de todos los elementos del tensor es 1, siguiendo por tanto

una distribución de probabilidad. Su expresión viene dada por

$$f_j(x_1, \dots, x_n) = \frac{e^{x_j}}{\sum_{k=1}^n e^{x_k}}, \quad (x_1, \dots, x_n) \in \mathbb{R}^n, \quad j = 1, \dots, n. \quad (1.2)$$

Continuando con la estructura interna de estos modelos, podemos organizar las neuronas en capas, de forma que las neuronas en una misma capa reciben la misma información de entrada, que puede ser la salida de las neuronas de una capa anterior o los datos de entrada a la red, y transfieren la información calculada a las neuronas de la capa siguiente por medio de la misma función de activación. Todas las neuronas de una misma capa son del mismo tipo, es decir, tienen la misma función de activación. Por otro lado, cuando colocamos las neuronas de forma secuencial, esto es, en diferentes capas, una de ellas recibe la información procesada por la neurona anterior, consiguiendo un conocimiento jerarquizado. De ese modo, la red aprende conceptos básicos en las primeras capas, y el conocimiento obtenido en cada una de ellas es procesado nuevamente por la siguiente capa elaborando cada vez conceptos más complejos, abstractos e interesantes. Cuantas más capas se añaden a la red más complejo puede ser el conocimiento elaborado. La profundidad en la cantidad de capas se conoce como aprendizaje profundo o Deep Learning (DL).

Por tanto, la estructura de una red neuronal tiene tres partes diferenciadas, la primera capa denominada capa de entrada, donde se encuentran las variables de entrada, la última capa denominada capa de salida y proporciona la salida de la red, y las capas intermedias denominadas capas ocultas. La Figura 1.14 muestra un ejemplo de esta arquitectura.

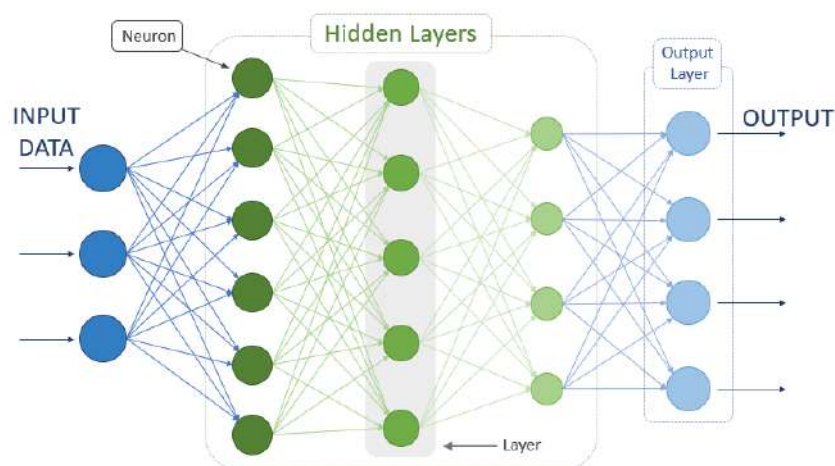


Figura 1.14: Ejemplo de arquitectura de una red neuronal, indicando la capa de entrada a la red, la capa de salida y las capas ocultas. Fuente: elaboración propia.

Para llevar a cabo el proceso de entrenamiento de estos modelos, como hemos visto en el apartado anterior, podemos hacer uso de una gran variedad de métodos y reglas de aprendizaje. En el caso de los modelos de aprendizaje supervisado, que son los que utilizaremos en este estudio, el aprendizaje se basa en la minimización del error obtenido por la red cuando, durante el entrenamiento se compara la salida producida por la red con la salida esperada (etiqueta asociada a los datos de entrada). El entrenamiento consiste, por lo tanto, en utilizar la superficie de error asociada a la red para buscar el estado estable de mínimo error descendiendo por el gradiente de esa superficie, evaluada por medio de una *función de coste* (“loss function” en inglés) que puede presentar distintas expresiones que dependen de las particularidades de cada problema. Entre las funciones más comunes destacan el *error cuadrático medio*, calculado mediante la fórmula

$$ECM = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i) \quad (1.3)$$

donde  $\hat{Y}$  son los valores predichos por el modelo,  $Y$  los valores reales y  $n$  el número de valores, la *binary cross-entropy*, dada por

$$BCE = -\frac{1}{n} \sum_{i=1}^n (Y_i \log(\hat{Y}_i) + (1 - Y_i) \log(1 - \hat{Y}_i)) \quad (1.4)$$

y la *categorical cross-entropy*, dada por

$$CCE = -\frac{1}{n} \sum_{i=1}^n Y_i \log(\hat{Y}_i). \quad (1.5)$$

Para minimizar el error bastaría con calcular los mínimos locales de la función de coste,  $g(\theta)$ , mediante la derivada con respecto a los parámetros o pesos del modelo,  $\theta = (\theta_1, \dots, \theta_m)$ . Sin embargo, debido a la gran cantidad de parámetros involucrados, obtener esta información puede no ser trivial ni eficiente. Por ello, se recurre a un proceso iterativo en el que, en cada paso, se calculan las derivadas parciales de la función de coste,  $g(\theta)$ , con el fin de obtener el vector gradiente negativo  $-\nabla g(\theta)$ , que proporciona la dirección de máximo decrecimiento.

Para aproximarse al valor mínimo se introduce un parámetro  $\alpha$ , denominado tasa de aprendizaje (“learning rate” en inglés), que indica el desplazamiento efectuado en la dirección opuesta al gradiente, y que, por tanto, indica cuánto afecta el gradiente a la actualización de los pesos. Consecuentemente, en cada iteración, los nuevos parámetros,  $\theta' = (\theta'_1, \dots, \theta'_m)$  vienen dados por la expresión:

$$\theta' = \theta - \alpha \nabla g(\theta). \quad (1.6)$$

Los primeros modelos de MLP llevaban a cabo este proceso mediante la Regla Delta Generalizada o Backpropagation, definida como el método de obtención del vector gradiente en Redes Neuronales de múltiples capas y funciones de transferencia no lineales y diferenciables para poder aplicar a cada capa el descenso de gradiente. De forma general, a los algoritmos que llevan a cabo esta función se los conoce como *optimizadores*.

El objetivo del método Backpropagation es el de estimar el gradiente con respecto a los parámetros del modelo, por medio de la regla de la cadena para plantear un esquema iterativo que se orienta desde la salida hacia la capa de entrada. Los errores de una capa, es decir, las derivadas parciales de la función de loss con respecto a la salida de las funciones de activación de las neuronas de dicha capa, se calculan a partir de los errores de la capa siguiente. Esta técnica permite a la red auto-ajustar sus pesos de forma eficiente, incluso en arquitecturas complejas, haciendo posible el crecimiento de los modelos para adentrarnos en el aprendizaje profundo.

Este método se conoce como el descenso del gradiente y es la base de estos algoritmos de ML. Debemos notar que un valor alto de  $\alpha$  puede hacer que no converja en el mínimo y, consecuentemente, el algoritmo no se ajuste adecuadamente al problema planteado. Por el contrario, un valor pequeño de  $\alpha$  puede requerir de muchas iteraciones, haciendo el algoritmo ineficiente.

En lo que respecta a los optimizadores, encargados del cálculo de la función de coste, su desarrollo ha avanzado en los últimos años para dar lugar a algoritmos cada vez más eficientes, como es el caso de los optimizadores adaptativos. Entre ellos encontramos RMSProp (por sus siglas en inglés “Root Mean Square Propagation”), que usa el signo del gradiente para ver la dirección de máximo decrecimiento y adapta el tamaño del paso  $\alpha$  para cada peso, y el optimizador Adam (por sus siglas en inglés “Adaptive Moment Estimation”) [163], que es una modificación del RMSProp que adapta la tasa de aprendizaje en función de la media y varianza de la distribución de los parámetros.

#### 1.3.4. Deep Learning

Los modelos de aprendizaje profundo o Deep Learning (DL por sus siglas en inglés) son modelos de ML basados en arquitecturas jerárquicas de varias capas que pueden aprender representaciones de datos con múltiples niveles de abstracción [164]. La Figura 1.9 muestra cómo los modelos DL son un subconjunto de redes neuronales que cuentan con un alto número de capas ocultas, según la estructura mostrada en la Figura 1.14. Las técnicas de DL han tenido un creciente desarrollo en los últimos años en diferentes disciplinas con resultados muy prometedores. Sin embargo, las arquitecturas más complejas necesitan una mayor cantidad de datos de entrada y un cuidadoso proceso de entrenamiento, que requiere de más capacidad de cómputo para llevarse a cabo. Por lo

tanto, la selección adecuada del conjunto de datos es primordial, así como la transformación de los datos de entrada mediante operaciones matemáticas. Consecuentemente, el uso de modelos más complejos permite extraer información cada vez más abstracta e interesante.

Las redes existentes dentro del campo del DL tienen arquitecturas complejas y muy diversas que permiten adaptarse a los diferentes tipos de datos de entrada. Entre estos modelos encontramos las redes neuronales convolucionales (CNN por sus siglas en inglés “Convolutional Neural Networks”) para el análisis de imagen, las redes neuronales recurrentes (RNN por sus siglas en inglés “Recurrent Neural Networks”) para el análisis de datos secuenciales, modelos GAN (por sus siglas en inglés “Generative Adversarial Networks”) para generación de imagen, modelos *autoencoders* que se utilizan con diferentes finalidades (como la eliminación de ruido o identificación de anomalías), o modelos Transformers dentro del campo del Procesamiento del Lenguaje Natural (PLN, o en inglés “Natural Language Processing”) para el análisis de texto. Aunque estos modelos tienen una arquitectura adaptada para un tipo concreto de dato de entrada, pueden ser modificados y combinados entre ellos, generando modelos híbridos, para abarcar otro tipo de problemas más complejos.

A continuación, vamos a presentar con más detalle los conceptos básicos de los modelos que utilizaremos para desarrollar el marcador de la enfermedad de Parkinson. Por un lado, para realizar estudios de conectividad funcional por medio de grafos, utilizaremos las GCNN (por sus siglas en inglés “Graph Convolutional Neural Networks”), que son una generalización de las CNN, por lo que introduciremos los conceptos principales de las redes convolucionales para facilitar la comprensión de las GCNN. Por otro lado, aplicaremos las técnicas de PLN para analizar las señales de electroencefalografía realizando una equivalencia entre texto y los canales del EEG. Este análisis se realizará por medio de un modelo tipo BERT, basado en la arquitectura Transformer, que es el modelo más representativo dentro del campo del PLN y que introduciremos más adelante.

#### 1.3.4.1. Redes Convolucionales

Las CNN son un tipo de red neuronal que se aplica principalmente para el análisis de imagen [165]. Estas redes se inspiran en el proceso de visión humano, que sigue un procesamiento en cascada donde en primer lugar se identifican los elementos básicos, y a partir de estos se generan patrones cada vez más complejos. Yann LeCun, en 1989, introdujo el primer diseño de una red convolucional para la detección de números escritos en cheques bancarios [166, 167].

Las CNN permiten obtener información acerca del entorno de cada píxel, teniendo en cuenta su información local [164]. Esto es muy importante en el tratamiento de imágenes

ya que los píxeles forman patrones independientemente de su posición dentro de la imagen. Las neuronas de estas arquitecturas implementan la operación convolución. Esta operación se aplica a dos funciones y se define matemáticamente como

$$(f * g)(m, n) := \int_{-\infty}^{+\infty} f(\tau, \lambda)g(m - \tau, n - \lambda)d\tau d\lambda. \quad (1.7)$$

y que en su versión discreta (necesaria para trabajar con tensores, no con funciones) es

$$(f * g)(m, n) := \sum_{\tau, \lambda} f(\tau, \lambda)g(m - \tau, n - \lambda), \quad (1.8)$$

donde  $f$  es la función que define la imagen original y  $g$  está asociada al filtro implementado.

La operación convolución da como resultado una tercera función que expresa el modo en que  $f$  es modificada por  $g$ , esto es, cómo la imagen es alterada por el filtro. La representación tanto de las imágenes (original y filtrada) como del filtro, es tensorial. En el caso de una representación bidimensional, cada píxel de la imagen es un elemento de su matriz asociada, y cada elemento del filtro es un parámetro de su correspondiente matriz. Esta operación representa el desplazamiento del filtro sobre los coeficientes de la matriz de la imagen, realizando, para cada píxel, una suma ponderada de los valores de los píxeles circundantes usando los pesos del filtro. En la Figura 1.15 se muestra un esquema de este proceso.

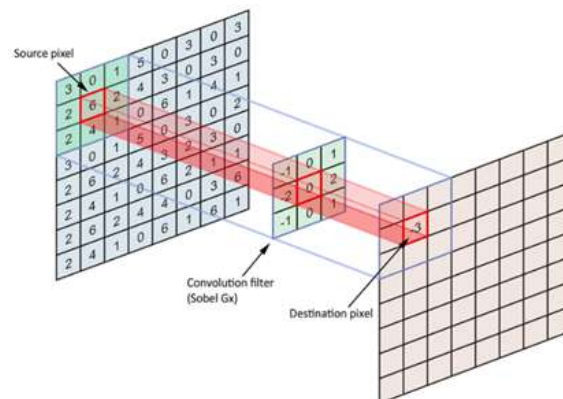


Figura 1.15: Ejecución de un filtro sobre una imagen representada por una matriz. Fuente: An intuitive guide to Convolutional Neural Networks <https://freecodecamp.org>

Los diferentes resultados dependerán, por tanto, de los parámetros que definan el filtro, permitiendo detectar bordes, texturas, contrastes, o promediar la intensidad de los píxeles para desenfocar la imagen, entre otras muchas aplicaciones. Los parámetros que conforman la matriz del filtro son los parámetros que la red debe modificar mediante



el proceso de aprendizaje para poder detectar patrones en la imagen. Cada una de las imágenes generadas se conoce como mapa de características, ya que indica las características asociadas al filtro aplicado.

Sin embargo, las características que podemos encontrar tras la aplicación del filtro son básicas o elementales. Para identificar patrones más complejos se hace necesario realizar esta operación de forma secuencial, apilando varias capas convolucionales, dispuestas de modo que la salida de cada una de ellas se convierta en la entrada de la siguiente. Esta sucesión de operaciones aumentará la cantidad de mapas de características, permitiendo identificar patrones cada vez más complejos y abstractos al componer patrones simples previos. Tras la aplicación de la convolución, se busca sintetizar la información de píxeles próximos, reduciendo, en el proceso, el tamaño de la imagen. Esto se consigue, bien identificando el máximo (MaxPooling), o el valor promedio (Average-Pooling), de entornos de píxeles. De este modo, la arquitectura típica de una CNN, que podemos ver en la Figura 1.16, se representa como un embudo, donde la imagen inicial se va comprimiendo, es decir, se reduce su tamaño, al mismo tiempo que el número de mapas de características va aumentando.

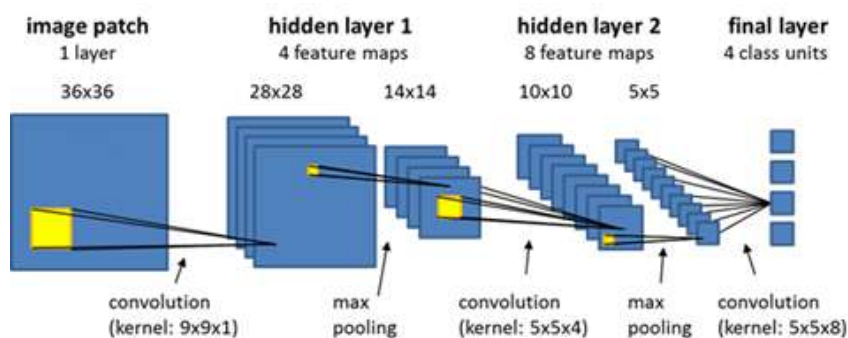


Figura 1.16: Ejemplo de una arquitectura CNN. Fuente: Convolutional Neural Networks Algorithms <https://docs.ecognition.com>

Por tanto, las CNN permiten el análisis de imagen por medio de la extracción de patrones que van adquiriendo cada vez mayor complejidad. Esta misma metodología se puede aplicar sobre estructuras menos regulares, tales como los grafos.

Un grafo  $G$  se define en [168] como un par  $G = (V, E)$ , donde  $V$  es el conjunto de vértices o nodos y  $E$  el conjunto de aristas. Un ejemplo visual de un grafo puede verse en la Fig. 1.17. En el caso de los grafos *ponderados* (*weighted graphs*), cada arista tiene asociado un peso, que representa una medida de la relación entre los dos nodos conectados por ella.

El análisis de este tipo de estructuras más irregulares y heterogéneas se basa en la búsqueda de patrones o clasificación de grafos. Para realizar este tipo de tareas, se

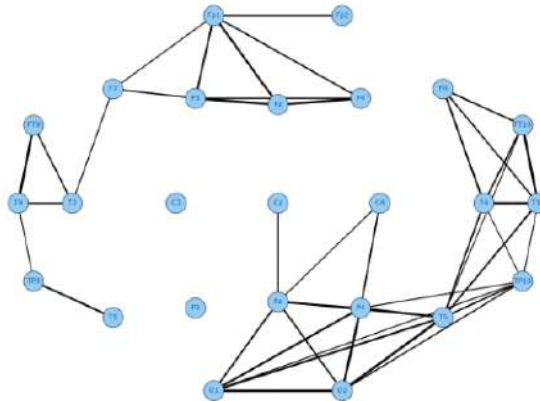


Figura 1.17: Grafo obtenido por medio de la librería EEGraph de un paciente con epilepsia para la medida de la correlación de Pearson con el umbral especificado por defecto. Fuente: elaboración propia.

han desarrollado las GCNN, una generalización de las CNN que agregan en cada nodo del grafo, expresado mediante su matriz de adyacencia (esto es su expresión matricial), la información de sus conexiones por medio de la operación convolución. Sin embargo, hay numerosos tipos de redes neuronales de grafos basadas en CNN que permiten la adaptación de la red al problema o estudio que se desee abordar [169], lo que convierte a las GCNN en uno de los modelos de aprendizaje profundo de grafos más destacados.

#### 1.3.4.2. Modelos de procesamiento del lenguaje natural

Los modelos dentro del campo del PLN buscan dar solución al problema de cómo una máquina puede entender y comprender el lenguaje humano. Estos modelos abarcan tareas como la traducción, la generación de texto o la interacción con humanos por medio de una conversación. Este campo de la IA se ha ido desarrollando paralelamente al de herramientas estadísticas en primer lugar y neuronales posteriormente. Entre las primeras aproximaciones para abordar este problema encontramos los modelos probabilísticos, pero el avance en el campo de las ANN hizo que a partir de 2013 [170, 171] se empezasen a desarrollar modelos de traducción automática neuronal (NMT, por sus siglas en inglés “Neural Machine Translation”) que pronto se extendieron a otros ámbitos del PLN. Se trata de un enfoque que utiliza una red neuronal artificial para predecir la probabilidad de una secuencia de palabras, generalmente modelando oraciones completas en un único modelo integrado.

Los primeros modelos de ANN utilizados fueron las RNN por medio de las capas LSTM (por sus siglas en inglés “Long Short-Term Memory”) [172] aunque los problemas de la falta de memoria limitaron su aplicación y desarrollo. Estas limitaciones fueron

superadas al añadir los módulos de atención a la arquitectura anterior, ya que estos manejaban la información del texto de manera global, mientras que las capas LSTM lo hacían de forma secuencial. En el artículo “Attention is all you need” de Google (2017), se presentan los modelos Transformers [173], cuya arquitectura estaba basada en una estructura encoder-decoder (codificador-decodificador) que se compone únicamente de módulos de atención. Estos modelos además de solventar las limitaciones anteriores, reducen los tiempos de entrenamiento y mejoran los resultados de modelos previos. La arquitectura de este modelo se muestra en la Figura 1.18.

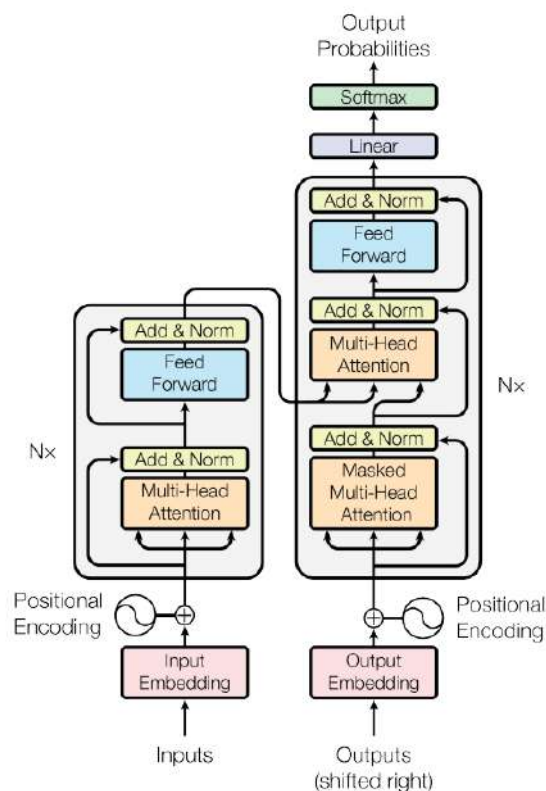


Figura 1.18: Arquitectura de un modelo Transformer. Fuente: [173].

Para poder trabajar con texto, primero debemos vectorizar cada una de las palabras para que el modelo sea capaz de analizarlas, esto es, dar a cada palabra una representación numérica que, de alguna manera, recoja el significado de la misma en relación con otras palabras. Uno de los enfoques fue utilizar la codificación conocida como one-hot-encoding, en la que la representación se realiza por medio de un vector de dimensión igual al número total de palabras diferentes del texto (vocabulario). Consecuentemente, la posición del vector que coincida con la de la palabra tendrá un valor de 1, mientras que los restantes valores del vector tomarán un valor de 0, es decir, para la primera palabra

$w_1$  su vector sería  $[1, 0, 0, \dots, 0]$ , para la segunda palabra  $w_2$  su representación vendría dada por  $[0, 1, 0, \dots, 0]$  y para la última palabra  $w_n$  por  $[0, 0, 0, \dots, 1]$ . Esta representación equivale a generar un espacio  $n$ -dimensional con  $n$  vectores independientes, de modo que cada palabra es equidistante al resto. Sin embargo, las palabras pueden agruparse mediante familias o grupos semánticos, de manera que podemos decir que dos palabras como “perro” y “gato” están más relacionadas que “perro” y “coche”. Por tanto, el siguiente paso fue buscar una codificación capaz de relacionar palabras semánticamente parecidas. Esta representación se realiza a través de los embeddings, es decir, por medio de un vector con dimension menor que el tamaño del vocabulario y cuyos elementos son números que asigna la red en función de la relación semántica con otras palabras. De este modo, en su representación en un espacio  $m$ -dimensional (con  $m < n$ ) tendremos clusters de palabras relacionadas semánticamente, mientras que palabras alejadas serán palabras sin este tipo de relación.

A partir de estos embeddings, el mecanismo de atención proporciona información acerca de la relación entre las palabras de una frase o un texto, es decir, realiza la tarea de comprender el texto. Para ello, el modelo genera durante el entrenamiento tres matrices de pesos: *query*  $W_Q$ , *key*  $W_K$  y *value*  $W_V$ , que multiplicadas por el embedding  $x_i$  (asociado a la palabra  $i$ ) dan lugar a los vectores *query*  $q_i = x_i W_Q$ , *key*  $k_i = x_i W_K$  y *value*  $v_i = x_i W_V$ . Con estos vectores, podemos estudiar la relación entre la palabra  $i$  y la palabra  $j$  por medio del producto escalar  $a_{ij} = q_i \cdot k_j$ , ya que cuanto mayor sea el valor obtenido por medio de esta operación mayor será la coincidencia de las direcciones de estos dos vectores y, por tanto, las palabras tendrán una mayor relación semántica. Los valores  $a_{ij}$  reciben el nombre de pesos de atención. Si calculamos la relación de la palabra  $i$  con las demás, conformamos el vector de pesos de atención  $a_i = (a_{i1}, a_{i2}, \dots, a_{in})$  que permite evaluar la importancia que nuestro modelo asigna al resto de palabras para dotar de contexto a la frase.

Realizando esta operación para todas las palabras, obtenemos la matriz de pesos de atención  $A$ , que contiene la importancia que el modelo asigna a cada palabra en función de una dada. Así,  $A = QK^T$ , con los vectores  $q_i$  y  $k_i$  constituyendo las filas de  $Q$  y  $K$ , respectivamente. Tras este paso, se normalizan mediante una capa softmax encargada de generar una distribución de probabilidad con los valores obtenidos. Para evitar que los gradientes tras la aplicación de la función *softmax* se desvanezcan durante el entrenamiento, los pesos de atención  $a_{ij}$  se dividen por la raíz cuadrada de la dimensión de los vectores *key*,  $\sqrt{d_k}$ .

Finalmente, la matriz  $V$ , constituida por los vectores  $v_i$  dispuestos en filas, contiene la información del texto dado como input, con tantas filas como palabras conforme la frase, de modo que la salida de la unidad de atención es la matriz  $V$  ponderada por los pesos

de atención  $A$  transformados según hemos descrito. La expresión formal se recoge en la ecuación 1.9. De este modo, el output contiene la información del contexto de la frase, dando mayor importancia a aquellas palabras a las que le haya prestado atención.

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V. \quad (1.9)$$

El alcance de esta arquitectura sigue explorándose hoy en día y ha evolucionado con modelos más grandes y complejos como BERT (Bidirectional Encoder Representations from Transformers) [174], y GPT (Generative Pre-trained Transformer) [175], que se entrenaron con grandes conjuntos de datos lingüísticos, como el corpus de Wikipedia, dotándolos de grandes capacidades para realizar tareas específicas con resultados sorprendentes.

### 1.3.5. Aplicaciones de Inteligencia Artificial en el campo de la medicina

El incremento de la cantidad de datos generados hoy en día permite la aplicación de los modelos de IA a nuevos retos y áreas. Debido al gran potencial que presentan estos modelos, están captando la atención de diversas disciplinas, entre ellas la medicina [176]. En este campo, el principal objetivo de estos artículos se centra en el diagnóstico [177], seguido de la evaluación del tratamiento de los pacientes y otras aplicaciones como la clasificación de una enfermedad, la cirugía robótica, o la predicción del pronóstico. Sin embargo, estas técnicas quedan restringidas al apoyo en el diagnóstico clínico debido a la controversia que radica en dejar que los sistemas inteligentes tomen decisiones autónomas acerca de la salud de los pacientes.

Las áreas terapéuticas en las que se han publicado más trabajos de IA, de acuerdo con [177], son la oncología seguida de la cardiología y de la neuropsiquiatría, si bien también se ha utilizado en estudios de oftalmología, neumología o dermatología. Dentro de la oncología, el cribado y apoyo al diagnóstico del cáncer de mama, junto con el estudio de la EAC (Enfermedad de las Arterias Coronarias) en el cáncer de pulmón han supuesto dos de los principales focos de las técnicas de Deep Learning. En cardiología, el principal objetivo es el apoyo al diagnóstico mediante imágenes de ecografía, resonancia magnética y perfusión arterial miocárdica o cerebral.

Esto hace que, en medicina, los modelos de IA se usen especialmente para el análisis de imágenes médicas, la detección de objetos o la segmentación por medio de las CNN [177, 178] aplicadas a diferentes trastornos [47]. El segundo tipo de dato sobre el que más se aplican estas técnicas son las series temporales [177], entre las que se encuentran las señales electrofisiológicas y, por tanto, los EEGs. Existen numerosos artículos dentro del campo de la neurociencia que aplican técnicas de ML para el es-

tudio del EEG en diferentes trastornos [179], como Alzheimer [35], esquizofrenia [180], alteraciones del sueño [181] y detección de ataques epilépticos [182]. Particularmente, aquellas relacionadas con el DL están ofreciendo resultados cada vez más interesantes, aunque su aplicación se ve limitada en función de la cantidad de datos disponibles, ya que en el campo de la medicina su adquisición es costosa y ligada a la disponibilidad de los pacientes.

El análisis del EEG específicamente por medio de técnicas de DL, ya se ha utilizado en enfermedades neurológicas como la epilepsia, alteraciones del sueño, y monitorización cognitiva entre otros [34, 182]. Destaca la capacidad de estos modelos para aprender representaciones de características a partir de datos en bruto, sin necesidad de un preprocesado previo de los EEGs, lo que permite extraer información de importancia pronóstica relacionada con la enfermedad. El creciente uso de estas técnicas queda reflejado en un aumento en el número de publicaciones en los últimos años, siendo el estudio de enfermedades neurológicas mediante IA un tema de investigación actualmente en desarrollo [47, 179]. Los avances en este campo favorecen el diagnóstico precoz de enfermedades, la evaluación farmacológica, además de tener fines terapéuticos como la rehabilitación de accidentes cerebrovasculares [183].

## 1.4. Inteligencia Artificial en el diagnóstico de la EP

El uso del DL en el análisis del EEG de pacientes con EP no ha sido totalmente validado en el entorno clínico. Sin embargo, los beneficios aportados por esta técnica han promovido su uso dentro de la investigación en neurociencia en los últimos años [184]. Es de esperar que el uso intensivo de la IA en las técnicas de diagnóstico de la EP pueda ayudar a detectar esta enfermedad en sus fases iniciales, es decir, en la etapa pre motora, favoreciendo el desarrollo de terapias preventivas que frenen el grado de avance del deterioro motor y cognitivo en la EP [185].

Para apoyar el diagnóstico de la enfermedad de Parkinson es deseable encontrar un marcador que sea objetivo, de bajo coste, de rápida adquisición, preciso y no invasivo para el paciente. La electroencefalografía, capaz de registrar el estado del cerebro del paciente en tiempo real, es un buen candidato ya que presenta la mayoría de estas características. Sin embargo, las señales del registro electroencefalográfico en la EP no muestran alteraciones que permitan su estudio de forma visual, por lo que se requiere de técnicas de análisis cuantitativo para poder obtener la información referente al estado de la enfermedad. Entre estas técnicas encontramos, como ya se ha comentado, el análisis de series temporales. Podemos adelantar que las características intrínsecas a las señales del EEG dificultan la obtención de resultados concluyentes por medio de

estos métodos. Además, su uso está condicionado por los procedimientos de limpieza empleados para minimizar el ruido de las señales y eliminar artefactos, ya que estos modifican las señales del EEG. Para realmente obtener un biomarcador objetivo, deberemos buscar otras técnicas que sean capaces de analizar los EEGs en bruto, esto es, sin realizar la limpieza de artefactos. Es aquí donde la potencia de los modelos neuronales y especialmente de las arquitecturas Deep Learning pueden constituir un elemento diferenciador respecto a otras aproximaciones, por lo que en esta tesis nos proponemos llevar a cabo un estudio cuantitativo de los EEGs de pacientes de Parkinson con estas técnicas de IA.

Una investigación exhaustiva del estado del arte del análisis de EEGs de enfermos de Parkinson con técnicas de IA confirma que se trata de un campo poco explorado, pero con un gran potencial, lo que justifica proponer los siguientes trabajos desarrollados en esta tesis:

- Por un lado, un estudio basado en la conectividad cerebral para evaluar las diferencias entre las conexiones de los pacientes con EP y controles. Un estudio similar no ha sido abordado con esta perspectiva anteriormente. Para analizar la conectividad modelaremos los EEGs como grafos para después utilizar GCNN para desarrollar el marcador de la enfermedad.
- Por otro lado, y como resultado de la experiencia investigadora acumulada en el grupo de investigación en el seno del cual se han desarrollado los trabajos de esta tesis, proponemos también un planteamiento innovador: el uso de modelos del campo del PLN (basado en BERT) para ese análisis cuantitativo. Ello implica realizar una equivalencia entre texto y EEGs de modo que podamos utilizar este modelo para analizar los canales del EEG y tener así otra vía que conduzca a la obtención del marcador de la enfermedad.





## Capítulo 2

# Objetivos e hipótesis

### 2.1. Objetivo general

*El objetivo de esta tesis es desarrollar un biomarcador basado en el registro electroencefalográfico para detectar la enfermedad de Parkinson usando técnicas de Deep Learning.*

### 2.2. Hipótesis del estudio

Las hipótesis que se han considerado para llevar a cabo el objetivo de esta tesis se enuncian a continuación:

- HIPÓTESIS 1: Las técnicas de análisis de series temporales son capaces de extraer información de la EP a partir de las señales del EEG.
- HIPÓTESIS 2: Las técnicas de DL tienen la potencia computacional necesaria para analizar señales del EEG y extraer características intrínsecas a la EP permitiendo la aplicación a un único individuo.
- HIPÓTESIS 3: Los estudios de conectividad funcional mediante el modelado de EEGs como grafos permiten identificar la presencia de la EP en un paciente.
- HIPÓTESIS 4: Las técnicas de PLN permiten relacionar segmentos del EEG detectando las características de la EP para identificar la enfermedad en sujetos individuales.

## 2.3. Objetivos específicos

### **OBJETIVO 1: Desarrollar un biomarcador para la EP basado en la aplicación de técnicas estadísticas a los EEGs**

Para verificar la *hipótesis 1*, se analizarán las señales del EEG mediante técnicas estadísticas y de análisis de series temporales. Esto nos permitirá profundizar en la comprensión de las señales de los EEGs y de la información que nos ofrecen. Entre estos métodos se encontrarán las medidas de conectividad, el análisis de resultados por medio de pruebas estadísticas para evaluar la significancia de los resultados, y otras medidas relacionadas con el análisis de señales, como la entropía de Shannon, utilizada en teoría de la información.

Las limitaciones inherentes a este estudio conducirán a la introducción de nuevas técnicas de análisis que favorecerán el desarrollo de un marcador objetivo de la enfermedad. Los resultados obtenidos nos indicarán la eficacia de las medidas de conectividad para detectar alteraciones producidas por la EP.

### **OBJETIVO 2: Evaluar de la capacidad de las técnicas de DL para extraer características del EEG que permitan identificar la EP**

Se comprobará el cumplimiento de la *hipótesis 2* mediante una revisión sistemática de la literatura que nos proporcionará un punto de partida acerca del desarrollo de modelos de DL empleados en este campo. Buscaremos establecer una metodología de trabajo que nos permita hacer una evaluación de los datos, la arquitectura de los modelos empleados junto con las técnicas de validación. El análisis de la literatura marcará el punto de partida de la investigación posterior que abordará aquellos campos inexplorados en el análisis de EEG de pacientes con EP mediante técnicas de DL. Estos son, estudios acerca de la conectividad funcional por medio de medidas de conectividad y teoría de grafos, y estudios que analizan las señales de EEG por medio de técnicas de PLN.

### **OBJETIVO 3: Desarrollar un biomarcador basado en la conectividad funcional obtenida a partir del modelado de EEGs como grafos y su análisis con técnicas de Deep Learning**

Se probará la *hipótesis 3*, realizando un estudio de conectividad modelando los EEGs de pacientes con EP como grafos. Dada la ausencia de herramientas que permiten realizar esta tarea se desarrollará una herramienta propia. Esta deberá, además, implementar diferentes medidas de conectividad, que permitirán definir las conexiones del grafo desde diferentes puntos de vista.

Una vez desarrollada la herramienta, se modelizarán los EEGs como grafos para evaluar las diferencias en la conectividad cerebral. Se usarán las GCNN con el fin de desarrollar un clasificador que identifique las conexiones asociadas a la EP y permita diagnosticar la enfermedad. También se realizará una evaluación acerca del efecto de la medicación en los pacientes con EP.

#### **OBJETIVO 4: Desarrollar un biomarcador basado en el análisis de EEGs de pacientes con EP mediante técnicas del campo del PLN**

Probaremos un modelo dentro del campo del PLN para validar la *hipótesis 4*. Para implementar los modelos del campo del PLN al análisis de las señales de EEG en pacientes con EP, se realizará una equivalencia entre texto y EEGs basada en que ambos tipos de datos son series temporales. Con este enfoque se buscará relacionar fragmentos de las señales de EEG de igual modo que estos modelos relacionan palabras en una frase. El modelo que se utilizará es de tipo BERT, adaptado al problema de clasificación binaria o diagnóstico que permita distinguir pacientes individuales con EP de controles.



## Capítulo 3

# Diseño experimental, Pacientes, Materiales y Métodos

### 3.1. Diseño experimental y cálculo de la muestra

#### 3.1.1. Diseño experimental

Se realiza un estudio de cohortes retrospectivo, encargado de comparar dos grupos de individuos, afectado y libre de una determinada condición (enfermedad), de acuerdo con la exposición a un factor [186]. El estudio es transversal, ya que se analiza dicho factor en un momento determinado que no tiene continuidad en el tiempo [187].

Las cohortes de pacientes incluidos en este estudio se obtienen de bases de datos con registros basales de pacientes con Enfermedad de Parkinson que se reclutan para otros estudios del grupo. Los criterios de inclusión de todos los sujetos son los mismos y el registro de datos demográficos y clínicos es uniforme. Cada estudio cuenta con el consentimiento informado de los pacientes para el uso de los datos en investigaciones adicionales derivadas de los mismos.

#### 3.1.2. Cálculo de la muestra

Dado que la aplicación de los resultados es de potencial relevancia en el campo clínico y aún a pesar de que el número de muestras requeridas por las técnicas de ML viene determinado por la complejidad del modelo utilizado, se hace un cálculo preliminar de la muestra basándonos en datos clínicos para garantizar que se incluyen como mínimo el número de sujetos calculados, por lo cual el cálculo muestral se usa como una referencia. Para determinar la muestra final de cada estudio, se ha tenido en cuenta el número de participantes involucrados en estudios con objetivos y arquitecturas similares.

En el caso de no tener referencias bibliográficas, el aprendizaje adquirido acerca de los modelos empleados y la evaluación de los procesos de entrenamiento y validación, nos permitirán identificar alteraciones en los resultados producidos por la escasez de datos.

Dado que los datos se obtienen de una base previamente registrada en el contexto de otros estudios de nuestro grupo de investigación, no se ha contravenido normas éticas realizando estudios a más pacientes de los necesarios.

Para determinar el tamaño de la muestra para los estudios que incluyen registros electroencefalográficos de pacientes Pre-Post LD, se realiza un análisis de potencia a priori con el software G\*Power teniendo en cuenta la diferencia de valores en la escala UPDRS-III con la ingesta de levodopa [188]. En este tipo de análisis se obtiene el tamaño de la muestra en función de la potencia ( $1 - \beta$ ) (donde  $\beta$  es el error de Tipo II o probabilidad de aceptar la hipótesis nula cuando no es cierta), significancia  $\alpha$  (o error de Tipo I que indica la probabilidad de rechazar la hipótesis nula cuando es cierta), y el índice del tamaño del efecto asociado a la población.

Considerando la hipótesis nula  $H_0$  de que las medias de los valores UPDRS-III Pre-LD y Post-LD son idénticas, realizamos un t-test de una cola tomando los valores  $\beta = 0,2$  y  $\alpha = 0,05$  considerados en [188]. El índice del tamaño del efecto queda determinado a partir de la mínima mejoría de UPDRS-III con importancia clínica, establecida en 5 puntos en dicha escala [188], y de la desviación estándar de la población. Al no disponer de este último dato, realizamos el análisis considerando un valor de 7,7 de std (por sus siglas en inglés "standard deviation"), determinada a partir de los datos de 42 pacientes con PD en [189], tal y como se indica en el material suplementario de [188]. Con estos valores, el análisis de potencia proporciona un tamaño de muestra de 17 pacientes tal y como se muestra en la Figura 3.1.

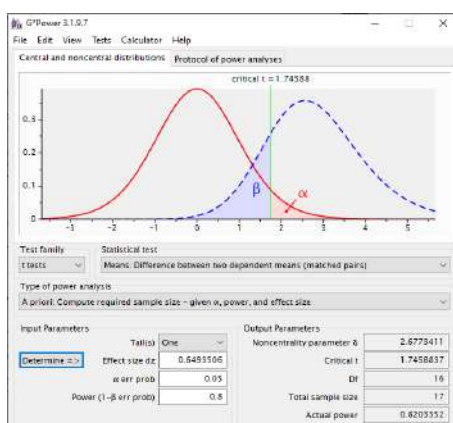


Figura 3.1: Ejecución del software G\*Power para el cálculo del tamaño de la muestra. Fuente: elaboración propia

Dado que el número de muestras requeridas por las técnicas de ML viene determinado por la complejidad del modelo utilizado, tomaremos el valor obtenido mediante el cálculo muestral como referencia para los estudios posteriores.

## **3.2. Pacientes**

### **3.2.1. Reclutamiento**

Los pacientes incluidos en este estudio formaron parte de otros dos proyectos del mismo grupo de investigación, cuyos protocolos encontramos en [190, 191]. Por tanto, los criterios de inclusión y exclusión a los que se sometieron los pacientes quedan definidos por las necesidades específicas de dichos proyectos.

Los participantes fueron reclutados en la consulta externa de EP del Hospital Beata María Ana de Madrid, España, y remitidos desde otros centros y hospitales colaboradores o autorreferidos que conocían los proyectos de investigación debido a la difusión a través de las redes sociales. Los sujetos incluidos fueron valorados por un neurólogo experto en trastornos del movimiento. Todos los sujetos fueron pacientes con EP idiopática diagnosticados según los criterios del Banco de Cerebros de la Sociedad de la Enfermedad de Parkinson del Reino Unido y tenían una prueba de screening cognitivo que descartaba deterioro cognitivo.

Durante el transcurso de los protocolos mencionados, se prohibieron los cambios en la medicación dopaminérgica y, en caso de ser necesario, conllevó a la exclusión del paciente. La dosis diaria total de fármacos dopaminérgicos se registró y controló en el análisis final de los resultados.

Los datos sensibles de todos los participantes se registraron por separado, se anonimizaron y se custodiaron de acuerdo con la legislación europea vigente en materia de protección de datos. Todos los datos registrados se verificaron dos veces en una base de datos diseñada específicamente para los estudios mencionados. Ambos proyectos cuentan con la aprobación de un comité de bioética sanitario y requieren la firma de los consentimientos informados por todos los participantes.

El reclutamiento de los pacientes sin EP que se tomaron como controles, siguieron las mismas pautas que los participantes con EP.

### **3.2.2. Criterios de inclusión**

Los criterios de inclusión comunes a los protocolos son:

- Mayor de 18 años.

- Estudio de neuroimagen sin patologías previas.
- Puntuación  $> 5$  en el ítem de transferencias (de la cama a la silla y viceversa) del Índice de Barthel.
- Puntuación  $\geq 24$  en el Mini-Mental State Examination.
- Enfermedad de Parkinson idiopática (diagnosticada según los criterios del Banco de Cerebros de la Sociedad de la Enfermedad de Parkinson del Reino Unido).
- Estadio  $< III$  de Hoehn-Yahr sin fluctuaciones motoras evidentes.
- Capacidad para dar el consentimiento informado para participar en el estudio.

Los criterios de inclusión para los controles fueron:

- Mayor de 18 años.
- Ausencia de enfermedades neurológicas.
- Puntuación  $\geq 24$  en el Mini-Mental State Examination.
- Capacidad para dar el consentimiento informado para participar en el estudio.

### 3.2.3. Criterios de exclusión

Los criterios de exclusión comunes a los estudios son:

- Antecedentes de enfermedades neurológicas diferentes a la EP.
- Presencia de implantes o piezas metálicas en la cabeza.
- Presencia de marcapasos cardíaco, estimuladores vagales, cerebrales o transcutáneos, bombas de medicación, derivaciones ventriculoperitoneales o clips de aneurisma.
- Dificultades significativas en el lenguaje.
- Antecedentes de abuso de alcohol o drogas.
- Problemas médicos no controlados.
- Embarazo.
- Epilepsia.



- Depresión activa moderada o grave (BDI > 14).
- Cambios de medicación dopaminérgica en los últimos 30 días.
- Datos atípicos para la EP idiopática.

Los criterios de exclusión aplicados a los controles fueron los mismos.

#### **3.2.4. Aspectos éticos**

Todos los participantes incluidos en el estudio dieron su consentimiento informado por escrito después de una explicación completa del procedimiento. Los estudios se llevaron a cabo de acuerdo con los principios de la Declaración de Helsinki, revisada en 2013, y fueron aprobados por el comité de normas éticas en experimentación humana del Hospital de Fuenlabrada Código 16-37 y el Comité de Ética Institucional del Hospital Universitario 12 de Octubre con el código 20-514.

### **3.3. Materiales**

#### **3.3.1. Adquisición de electroencefalogramas**

Para llevar a cabo el registro de los EEGs se utilizó un amplificador actiCHamp (Brain Vision LLC, Morrisville, NC, EE.UU.) para amplificar y digitalizar los datos del EEG a una frecuencia de muestreo de 512 Hz [192]. La actividad cerebral se registró desde 64 posiciones con electrodos de cuero cabelludo Ag/AgCl activos (electrodos actiCAP, Brain Vision LLC, NC, EE.UU.) y los datos del EEG se almacenaron en un PC con Windows 7 (Microsoft Corporation, Washington, DC, EE.UU.). Los electrodos de tierra y de referencia se colocaron en AFz y FCz, respectivamente.

La adquisición del EEG se llevó a cabo mediante el software NeuroRT Studio (Mensia Technologies SA, París, Francia). El procesamiento de la señal de EEG se realizó utilizando funciones de MATLAB (MathWorks Inc., Natick, MA, EE.UU.), específicamente la extensión EEGLab [193].

Los electrodos se encuentran colocados de acuerdo con el sistema internacional 10-10 de alta resolución. La nomenclatura del sistema hace referencia a la distancia entre electrodos adyacentes, siendo esta de un 10% de la distancia total [194]. Un esquema de la distribución de los electrodos del casco que hemos utilizado se muestra en la Figura 3.2.

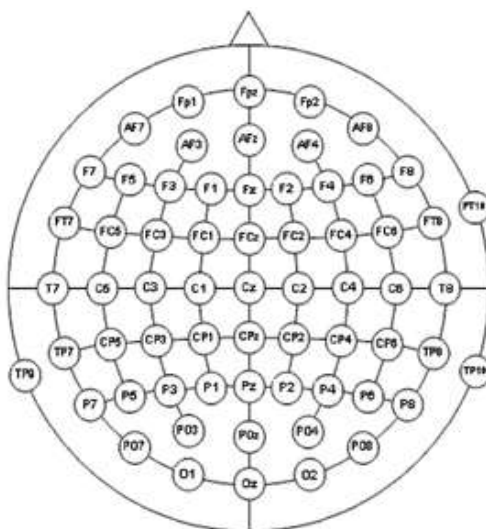


Figura 3.2: Representación esquemática de la distribución de los electrodos de un casco de electroencefalografía de 64 canales. Fuente: <https://www.brainlatam.com/manufacturers/eeg-cap/extended-1020-system-with-30-channels-191>

### 3.3.2. Procesamiento de señales biomédicas

#### 3.3.2.1. Dataset

Cada uno de los registros electroencefalográficos realizados por el software NeuroRT Studio dio lugar a un fichero en formato “.GDF” (General Data Format) que contenía el valor numérico de los registros. Además, se dispone de un fichero en formato “.set” que contiene la posición de los electrodos junto con la etiqueta que asocia el nombre con el número del canal.

Aquellos EEGs que se sometieron al proceso de limpieza fueron exportados al formato “.dat” por medio del software utilizado para realizar la limpieza de artefactos.

En la Tabla 3.1 se muestra un resumen de las características del dataset completo, aunque en cada proyecto llevado a cabo se especificarán los participantes incluidos dentro de cada estudio debido a que el registro de los EEGs se llevó a cabo durante todo el transcurso de esta tesis.

Item	Pacientes con EP	Controles
Participantes	80	24
Frecuencia de muestreo	512 Hz	512 Hz
Número de canales	64	64

Item	Pacientes con EP	Controles
Número total de pruebas de EEG consideradas	3	3
Número total de EEGs	240	72
Tamaño de cada EEG de Resting	15,5 MB	15,5 MB
Número de puntos de cada EEG de Resting	32.000	32.000
Tamaño de cada EEG de FT	8 MB	8 MB
Número de puntos de cada EEG de FT	16.000	16.000

Tabla 3.1: Resumen de las características del dataset especificadas para los EEGs de pacientes con EP y los EEGs de controles. Fuente: elaboración propia.

### 3.3.2.2. Protocolo de limpieza de artefactos

Para cada uno de los EEGs registrados, se llevó a cabo un control de calidad para validar los datos antes de incluirlos en cada uno de los estudios propuestos. Este control se realizó mediante el software Matlab en su versión 2015b, junto con el paquete EEGLab en su versión 15\_0b.

Aquellos EEGs que se sometieron al proceso de eliminación de artefactos, detección y reconstrucción de canales anómalos y técnicas de filtrado, se procesaron mediante el mismo software Matlab, con versión 2015b, y EEGLab en su versión 15\_0b.

### 3.3.2.3. Medidas de conectividad

Las medidas de conectividad implementadas para el análisis de las señales del EEG mediante los métodos de análisis de series temporales fueron programadas en Matlab en su versión 2018a.

Para unificar los formatos y facilitar el desarrollo de estudios basados en medidas de conectividad, posteriormente se implementaron todas las medidas en el entorno de Python, junto con varias librerías relacionadas con el análisis, manipulación y visualización de datos. Entre las librerías importadas encontramos: *Numpy*, un paquete especializado en el cálculo numérico y el análisis de datos [195]; *Pandas*, una librería especializada en el manejo y análisis de estructuras de datos [196]; *matplotlib*, una librería diseñada para la visualización de datos [197]; *Mne*, es un paquete diseñado para la exploración, visualización y análisis de datos neurofisiológicos [198]; *NetworkX*, un paquete utilizado para la creación, manipulación y estudio de la estructura, dinámica y funciones de redes complejas [199]; *Plotly*<sup>1</sup>, una librería diseñada para la visualización de gráficos; *Scipy*,

<sup>1</sup><https://plotly.com/python/>

un paquete que proporciona herramientas matemáticas y algoritmos [200]; *SCoT*, un paquete de Python para el análisis de señales de EEG/MEG y estimación de fuentes de conectividad [201]; *AntroPy*<sup>2</sup>, un paquete de Python para implementar diferentes tipos de entropías y métricas complejas de series temporales.

### 3.3.3. Revisión sistemática de la literatura

Se utilizó el software Zotero en su versión 6.0.12, para almacenar los resultados de las búsquedas bibliográficas y como herramienta para manejar los documentos descargados.

### 3.3.4. Desarrollo de técnicas de Deep Learning

Para el desarrollo de los modelos de DL, se utilizó el software Python en diferentes versiones, junto con varias librerías cuyas versiones fueron actualizándose para incluir las mejoras implementadas. Además de las librerías anteriormente mencionadas, se utilizaron: *keras* [202]; *tensorflow* [203, 204]; *scikitlearn* [205, 206]; *PyMeta* [207] basado en el paquete PythonMeta en Python para el metanálisis.

En cuanto a hardware empleado, los modelos de DL se diseñaron en un ordenador con procesador intel-core i5-8500, memoria RAM de 32 GB, con ubuntu en su versión 20.04 y un disco SSD de 250 GB y otro HDD de 1 TB. Posteriormente fueron entrenados en un servidor con procesador AMD Ryzen Threadripper 2950X 16-Core, de 128 GB de memoria RAM, con versión de CUDA 11.6 y ubuntu 20.04, un disco SSD de 250 GB y 4 GPUs Nvidia GeForce RTX 2080 de 12 GB cada una.

## 3.4. Métodos

### 3.4.1. Adquisición de electroencefalogramas

El protocolo de registro del EEG empieza informando al paciente acerca de las pruebas que se van a realizar para obtener su consentimiento por escrito tal y como se especifica en los protocolos publicados por Gonzalez-Zamorano et al 2021 y Arroyo-Ferrer et al 2021 [190, 191].

Las pruebas electroencefalográficas comienzan colocando el gorro con los electrodos sobre la cabeza del paciente y ajustando el dispositivo de modo que todos los electrodos hagan contacto con el cuero cabelludo. Tras encender el dispositivo, se visualizan las impedancias de cada electrodo para regularlas mediante Neurogel. Dicho gel es introdu-

---

<sup>2</sup><https://raphaelvallat.com/antropy/build/html/index.html>

cido por medio de una jeringa de irrigación, a través de una abertura presente en cada uno de los electrodos. Tras la verificación de las impedancias, contenidas en el rango  $< 10 \text{ k}\Omega$ , se procedió al registro de la señal.

Para cada paciente, se realizaron tres registros electroencefalográficos asociados a tres pruebas distintas, diseñadas para obtener el registro de la actividad cerebral en su estado de reposo y durante la activación motora de ambos hemisferios. La tarea motora escogida es una de las más informativas que se incluyen en las evaluaciones clínicas, como la UPDRS, además de tener una representación somatotópica predominante en los ganglios basales [208]. La prueba de reposo se ha utilizado ampliamente en imágenes de resonancia magnética para estudiar la conectividad funcional entre regiones cerebrales específicas organizadas en redes [209]. El estado de reposo se ha utilizado ampliamente para estudiar los microestados de EEG [210] que se alteran en la EP en función de la administración de dopamina, [211]. El procedimiento llevado a cabo por medio de cada una de estas pruebas se describe a continuación.

- **Prueba de reposo o “Resting”**. El registro comienza con esta prueba, en la cual el paciente permanece sentado mirando una pared blanca, en estado relajado, sin hablar ni realizar movimientos, con los ojos abiertos. Esta prueba tiene una duración de 1 min.
- **Prueba de activación motora del hemisferio izquierdo o “FTD”**. Seguido de la prueba de Resting, se realiza la prueba motora Finger Tapping Derecho (FTD), que tiene como objetivo la activación motora del hemisferio izquierdo. En el transcurso de esta prueba, el paciente continúa sentado, eleva la mano derecha hasta la altura de su cabeza y realiza un golpeteo repetitivo con los dedos índice y pulgar de la mano elevada. El golpeteo debe ser tan rápido como pueda el paciente y con la mayor amplitud posible. Esta prueba tiene una duración de 30 s para evitar artefactos producidos por la fatiga.
- **Prueba de activación motora del hemisferio derecho o “FTI”**. El tercer registro se corresponde con la prueba Finger Tapping Izquierdo (FTI), que tiene como objetivo la activación motora del hemisferio derecho. Esta prueba, es equivalente a la anterior pero el paciente usa los dedos índice y pulgar de la mano derecha.

El orden de estas pruebas se mantuvo en todas las sesiones para todos los pacientes. El protocolo de registro fue diseñado para obtener la máxima información posible acerca del estado del paciente, teniendo en cuenta que la EP tiene una afectación motora bilateral. Todas las pruebas fueron registradas a una frecuencia de muestreo de 512 Hz, esto es, para cada segundo se produce un registro de 512 valores de potencial.

Un subconjunto de los participantes con EP realizaron estas pruebas antes y después de la ingesta de la medicación farmacológica. En estos casos, se pidió a los pacientes que acudieran al hospital a primera hora de la mañana sin su correspondiente ingesta diaria de levodopa o agonista (al menos 12 h después de la última dosis de medicación dopaminérgica). La actividad electroencefalográfica previa a la levodopa “Pre-LD”, se registró siguiendo el protocolo de registro anterior. A continuación, se retiró el casco con los electrodos del EEG y los participantes tomaron su dosis matinal habitual de levodopa o agonista con un vaso de agua, 30 minutos antes de tomar un desayuno ligero y disponer de tiempo libre. Una vez que los pacientes se encontraban en su estado “ON” habitual, cuando la levodopa se encuentra activa (esto es una hora después de la toma de levodopa “Post-LD”) se realizó el registro siguiendo el protocolo especificado.

### 3.4.2. Procesamiento de señales biomédicas

Uno de los principales problemas para analizar las señales de EEG mediante técnicas de análisis de series temporales, es el ruido. Para minimizar su influencia se realiza un pre-procesamiento en las señales denominado limpieza de EEGs. Sin embargo, los modelos dentro del campo del DL han demostrado ser robustos al ruido y artefactos de estas señales, por lo que el procesamiento realizado para estos proyectos no contienen el proceso de limpieza.

Debemos señalar que todos los EEGs incluidos en el estudio, fueron validados mediante una inspección visual para descartar aquellos que tuvieron fallos en el registro electroencefalográfico. Los protocolos específicos se describen a continuación.

#### 3.4.2.1. Dataset

El conjunto de datos o dataset se compone de los EEGs de los pacientes con EP y controles, registrados con un casco de 64 canales a una frecuencia de muestreo de 512 Hz.

Todos los EEGs recolectados fueron validados mediante un análisis visual. En este proceso, los EEGs que mostraron alteraciones significativas en las señales fueron descartados del estudio. Los criterios de exclusión fueron:

- **Un 20% o más de electrodos anómalos.** En este conjunto incluimos electrodos planos, es decir, electrodos que no han registrado señal, electrodos mal calibrados, estos son, canales cuya impedancia no ha sido bien calibrada, y electrodos cuyo registro contiene artefactos significativos.
- **Un conjunto de electrodos anómalos consecutivos menor que el anterior.** En

este caso se procedió a descartar el EEG, ya que la información asociada a esa región del cerebro no fue registrada y por tanto no puede ser reconstruida con la información de los electrodos adyacentes.

- **Las características no se mantienen durante el registro del EEG.** Durante la grabación del EEG, los electrodos pueden perder el contacto con el cuero cabelludo mientras se realiza el registro y variar su impedancia durante las pruebas. Si estas alteraciones son significativas, esto es, una alta variación entre los tramos del EEG, se procederá a la exclusión del EEG.
- **Registro de la señal inadecuado.** Si el EEG no se ha calibrado bien, la escala de registro de las señales variará con respecto a los otros registros. El rango que hemos considerado es de  $(250-750) \mu V$  pudiendo ser ampliado en algunos casos a  $(100-1000) \mu V$ .
- **Artefactos significativos.** Si el conjunto de las señales del EEG están alteradas de forma significativa por artefactos externos.

El protocolo diseñado para realizar la validación de los EEGs se incluye en el Apéndice B. Todos los EEGs validados quedaron registrados en un fichero con formato “.xlsx” donde se indicaron los canales anómalos, escala de registro y las características particulares de cada EEG.

#### 3.4.2.2. Protocolo de limpieza de artefactos

Para analizar las señales de EEG mediante las técnicas de análisis de series temporales, es necesario realizar un procesamiento del EEG para minimizar el ruido y los artefactos que contienen las señales. Se estableció un protocolo basado en referencias bibliográficas de limpieza de EEGs, este protocolo se especifica en el Apéndice C. Los pasos realizados fueron:

- **Localización de canales.** Si bien EEGLab reconoce automáticamente el número de canales, es necesario introducir un fichero con las posiciones y etiquetas de los nombres asociados a las posiciones de los electrodos del casco utilizado.
- **Filtro FIR (por sus siglas en inglés “Finite Impulse Response”).** Se realizó un filtro en el intervalo (1-100) Hz para acotar el intervalo de la señal y así disminuir el tiempo de computo.
- **Reconstrucción con ASR (por sus siglas en inglés “Artifact Subspace Reconstruction”).** En el caso de datos muy ruidosos, ASR puede corregir porciones de datos en bruto en lugar de rechazarlos.

- **Filtro FIR.** Se realizó un filtrado en el intervalo (1-49) Hz. Se realizó a 49 Hz para eliminar el pico de corriente a 50 Hz.
- **Análisis de componentes independientes ICA (por sus siglas en inglés “Independent component analysis”).** Es uno de los métodos basados en componentes que se utilizan ampliamente en el análisis de patrones y de bioseñales. ICA descompone las señales de EEG en componentes independientes en función de la independencia estadística de las señales [141].
- **Eliminación de artefactos mediante MARA (por sus siglas en inglés “Multiple Artifact Rejection Algorithm”).** Tras realizar la identificación de artefactos procedemos a su eliminación [212].
- **Identificación de canales anómalos.** Para seleccionar los canales anómalos se utiliza la función “Automatic Channel Rejection” seleccionando la Medida de Probabilidad de la caja de herramientas EEGLab [193], con un coeficiente por defecto de 5. En este paso se indican los canales que son seleccionados como anómalos en función de las especificaciones indicadas.
- **Eliminación de canales anómalos.** Se seleccionan los canales anómalos anteriores para reconstruirlos por medio de la herramienta “Interpolate Electrodes” utilizando el método esférico [213].
- **Exportación del archivo.** El archivo final es exportado para no sobrescribir el documento original.

Debemos notar que las señales no se han referenciado con respecto a la media, esto es, restar la señal promedio de todos los canales en cada instante de tiempo. De este modo se preserva la integridad de la señal, ya que algunas pequeñas características de la señal podrían quedar ocultas o perderse, e incluso difundir artefactos aislados de la señal dando lugar a estimaciones no realistas de los valores de las medidas [214].

Debido a la falta de un protocolo estandarizado por la comunidad de usuarios que utilizan los EEGs, cada grupo de investigación utiliza diferentes procedimientos de limpieza dificultando la comparativa entre resultados. Sin embargo, esta falta de estandarización permite a los investigadores limitar el tratamiento de las señales en caso de querer trabajar con señales menos procesadas. Debemos notar que los resultados estadísticos de las medidas dependen del proceso de limpieza realizado, ya que al eliminar el ruido de las señales también se pueden eliminar componentes relevantes asociadas a la enfermedad, dando lugar a un diagnóstico erróneo. Consecuentemente, el protocolo que hemos descrito elimina las componentes mínimas.



Las técnicas de ML, debido a su capacidad de abstracción y su gran complejidad, permiten analizar los EEGs sin realizar este proceso de limpieza de artefactos, esto es, analizando las señales de EEG en bruto. Para evitar el sesgo producido por la adquisición de la señal, estas se introdujeron tras ser validadas por el protocolo mostrado en el Apéndice B.

### 3.4.2.3. Medidas de conectividad

Los estudios de conectividad son de gran importancia dentro del campo de la neurociencia ya que permiten evaluar la organización tanto anatómica como funcional del cerebro. Estos estudios se realizan identificando y clasificando los patrones producidos por las conexiones entre las diferentes regiones cerebrales [215]. Sin embargo, los cambios de conectividad no pueden estudiarse mediante rasgos diagnósticos del EEG identificables visualmente y se han utilizado varios métodos de análisis de señales para resolver esta cuestión. Para obtener y representar información precisa acerca de la conectividad cerebral, se consideran una serie de medidas neurobiológicamente significativas sensibles a los cambios de conectividad. Algunas de las medidas más utilizadas aparecen en [215], donde podemos encontrar la coherencia, que se ha utilizado ampliamente para analizar la conectividad del EEG [216], proporcionando una medida de asociación lineal con una interpretación relativamente sencilla.

El estudio de las conexiones (tanto anatómicas como funcionales) de las diferentes regiones cerebrales permite identificar alteraciones en los patrones de conectividad y relacionar estas alteraciones con anomalías o trastornos neuropsiquiátricos [217, 218, 219, 220]. Se ha demostrado que los sujetos con EP tienen un aumento de la conectividad funcional cortico-cortical en estado de reposo en varias bandas de frecuencia [221, 222]. Este aumento de la conectividad también se ha correlacionado con manifestaciones clínicas como la gravedad del temblor y la demencia [223].

Dada la irregularidad de las señales de electroencefalográficas, los análisis que proporcionan una medida que indique los cambios de frecuencia en un rango son muy útiles para identificar las alteraciones en los patrones de conectividad. Teniendo en cuenta que la mayoría de las señales fisiológicas son no lineales, la entropía parece ser una medida ideal para estudiar las señales neuronales [224]. Además, la entropía ya ha sido utilizada previamente para analizar las señales del EEG, especialmente en la epilepsia [225].

Se seleccionaron las 12 medidas de conectividad más utilizadas en la bibliografía [226, 227, 228, 229, 230], para abordar los estudios relacionados con la conectividad funcional del cerebro. Las medidas implementadas buscan analizar las señales de EEG tanto en su dominio temporal, con medidas relacionadas con la conectividad entre electrodos, simetría entre las señales y grado de aleatoriedad de la señal, como en su do-

minio espectral, enfocadas a trabajar con las bandas de frecuencias, no con todo el espectro del EEG.

Para cada una de las medidas seleccionadas se buscó la función implementada por defecto en el entorno de software escogido (en este caso Python), y se comparó la función que la ejecutaba con la fórmula matemática que la describía. En el caso de no encontrar una medida implementada por defecto, esta fue desarrollada a partir de su definición matemática. Las medidas de conectividad consideradas son:

- **Dominio temporal:** correlación de Pearson, correlación cruzada, correlación cruzada corregida, PLI (por sus siglas en inglés “Phase Lag Index”) y entropía de Shannon.
- **Dominio espectral:** coherencia, coherencia imaginaria, PLV (por sus siglas en inglés “Phase Locking Value”), PLI, WPLI (por sus siglas en inglés “Weighted Phase Lag Index”), DTF (por sus siglas en inglés “Directed Transfer Function”), potencia espectral y entropía espectral.

Todas estas medidas de conectividad fueron implementadas en el entorno de Python, de las que proporcionamos su definición matemática, junto con una breve interpretación de su aplicación a las series temporales. También especificaremos el intervalo en el que se encuentra acotada la medida junto con un breve significado de su interpretación. Además, también se indica el umbral por defecto asociado a cada medida (basado en valores bibliográficos) que se utilizará más adelante. En lo sucesivo,  $x$  e  $y$  harán referencia a dos series temporales.

## Medidas definidas en el dominio temporal

### Correlación de Pearson

El coeficiente de correlación de Pearson ( $C$ ) mide la correlación lineal entre dos variables, dos conjuntos de datos o dos series temporales. Se define como:

$$C_{x,y} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}, \quad (3.1)$$

donde  $\sigma_{xy}$  representa la covarianza,  $\sigma_x$  es la desviación estándar de  $x$ , y  $\sigma_y$  es la desviación estándar de  $y$ . La ecuación (3.1) proporciona un escalar en el rango  $[-1, 1]$ , donde 1 significa una correlación perfecta,  $-1$  representa una correlación inversa perfecta, mientras que 0 significa que no hay correlación entre  $x$  e  $y$ . Se ha utilizado la función *pearsonr* dentro del módulo *scipy.stats* para el cálculo de este valor. El umbral por defecto para esta medida es de 0,7.

### **Correlación cruzada**

La correlación cruzada ( $CC$ ) mide la similitud de dos series temporales en función del desplazamiento relativo o desfase (lag) de una de ellas con respecto a la otra. Su valor se obtiene mediante la siguiente expresión:

$$CC_{xy}(m) = \frac{R_{xy}(m)}{\sqrt{R_{xx}(0)R_{yy}(0)}}, \quad (3.2)$$

con

$$R_{xy}(m) := \sum_{j=0}^{N-1} (x_j - \mu_x)(y_{j-m} - \mu_y)^*,$$

donde \* indica el conjugado,  $\mu_x$  y  $\mu_y$  son las medias respectivas de  $x$  e  $y$ ,  $N$  denota la longitud de  $x$  e  $y$ , que coinciden entre sí ya que estas series temporales pertenecen al mismo EEG, y  $y_{j-m} := 0$  si  $j - m < 0$  o  $j - m \geq N$ . Aquí  $m \in \mathbb{Z}$  y  $R_{xy}(m) = 0$  si  $m \leq -N$  o  $m \geq N$ . El cálculo del valor de  $R_{xx}$ ,  $R_{yy}$  y  $R_{xy}$  se ha realizado mediante la función *correlate* dentro del módulo *scipy.signal*. Hay que tener en cuenta que la definición de la correlación cruzada introducida en la Ecu. (3.2) está normalizada de forma que el valor de cero-lag se corresponda con el coeficiente de correlación de Pearson. De este modo, para pequeños lags su valor se encuentra en una vecindad del intervalo  $[-1, 1]$ . Para obtener un valor más representativo asociado al  $CC$ , se calcula la media del 10% de los valores positivos del lag para cada par de series temporales, y se aplica un umbral de 0,5 a estas medidas.

### **Correlación cruzada corregida**

La correlación cruzada corregida ( $CCC$ ) mide la simetría de la correlación cruzada de  $x$  e  $y$  con respecto al lag (esto es, el desplazamiento de  $y$  con respecto a  $x$ ). Su expresión está dada por:

$$CCC_{xy}(m) = CC_{xy}(m) - CC_{xy}(-m), \quad (3.3)$$

donde  $m$  representa el tiempo de desfase entre  $x$  e  $y$ , y  $CC_{xy}$  es su correlación cruzada como se ha establecido. Para la computación de la Ecu. 3.3, se usa la función *correlate* dentro del módulo *scipy.signal*. Al igual que para  $CC$ , para obtener un valor más representativo asociado al  $CCC$ , se calcula la media del 10% del lag para cada par de series temporales. Se aplica un umbral de 0,1 para esta medida.

### **Phase Lag Index**

El Phase Lag Index ( $PLI$ ) mide los cambios de signo de las diferencias de fase de

dos series temporales, y por tanto, los cambios relativos de la fase de una serie con respecto a la otra. Esta medida es menos sensible a la influencia de ciertas fuentes que el *PLV*, y su valor viene dado por la expresión:

$$PLI = |E[\text{sign}(\Delta\phi(t))]| \quad (3.4)$$

donde  $\Delta\phi(t)$  es la diferencia de fase de  $x$  e  $y$  en el tiempo  $t$ , desplazada para estar en el intervalo  $[-\pi, \pi)$ . El significado de  $E[\cdot]$ ,  $|\cdot|$  y las funciones utilizadas para realizar el cálculo de la Ecu. (3.4) se detallan en la medida de *PLV*. El valor del *PLI* obtenido en la Ecu. (3.4) pertenece al intervalo  $[0, 1]$ , donde 1 representa una posición relativa uniforme de la fase de  $x$  con respecto a la de  $y$ , mientras que 0 revela un número igual de cambios en la posición relativa de dichas fases. El umbral por defecto para esta medida es 0, 1, y puede ser llamado para toda la serie temporal con la opción *pli*, o para una o múltiples bandas (que deben ser especificadas) con la opción *pli\_bands*.

### **Entropía de Shannon**

La Entropía ( $H$ ) o entropía de Shannon mide la cantidad media de información contenida en una serie temporal  $x$ . Su expresión viene dada por:

$$H(x) = - \sum_i p(x_i) \log(p(x_i)) \quad (3.5)$$

donde  $\log$  denota el logaritmo natural, y  $p(x_i)$  es la probabilidad no normalizada del evento  $x_i$  en  $x$ . Para los cálculos en (3.5), se utiliza la función *entropy* dentro del módulo *scipy.stats*. Hay que tener en cuenta que el valor obtenido en la Ecu. (3.5) no está normalizado, por lo que depende del número total de ocurrencias,  $N$ . Si  $H$  se normaliza por el factor  $\log(N)$ , entonces el valor resultante pertenece al intervalo  $[0, 1]$ . El estado de máxima entropía, es decir, entropía normalizada igual a 1, se obtiene en sucesos equiprobables y significa información mínima, mientras que la entropía mínima, es decir,  $H(x) \simeq 0$  representa la máxima cantidad de información. De forma similar a la medida *PS*, el 25 % de los canales con los valores más altos son seleccionados y conectados en el grafo final. De este modo, sus aristas no se ven influidas por la normalización.

## **Medidas definidas en el dominio espectral**

### **Coherencia cuadrática**

La coherencia cuadrática (*SC*) mide la relación entre el espectro de dos series temporales, definiéndose como:

$$SC_{xy}(f) = \frac{|G_{xy}(f)|^2}{G_{xx}(f)G_{yy}(f)} \quad (3.6)$$

donde  $G_{xy}(f)$  es la densidad espectral cruzada entre  $x$  e  $y$ , y  $G_{xx}(f)$  y  $G_{yy}(f)$  representan las densidades espectrales de potencia de  $x$  e  $y$  respectivamente. La ecuación (3.6) proporciona una serie de escalares en el rango  $[0, 1]$ , donde 1 representa una correlación perfecta en el espectro de frecuencias, mientras que 0 significa que no hay correlación entre el espectro de ambas señales,  $x$  e  $y$ . Para el cálculo de este valor se utiliza la función *coherence* dentro del módulo *scipy.signal*. Una vez obtenido, se calcula su media en cada banda de frecuencia (delta, theta, alfa, beta y gamma). El umbral por defecto para esta medida es de 0,65.

### **Coherencia imaginaria**

Al igual que la coherencia cuadrática, la coherencia imaginaria (*CI*) mide la relación entre el espectro de dos series temporales. Sin embargo, es menos sensible a ciertos efectos externos en comparación con la parte real de la coherencia y, por tanto, en comparación con la *CI*. Su expresión viene dada por:

$$IC_{xy}(f) = \frac{\text{Imag}(G_{xy}(f))}{\sqrt{G_{xx}(f)G_{yy}(f)}} \quad (3.7)$$

donde  $G_{xy}(f)$  es la densidad espectral cruzada entre  $x$  e  $y$ , y  $G_{xx}(f)$  y  $G_{yy}(f)$  representan las densidades espectrales de potencia de  $x$  e  $y$  respectivamente. Cabe señalar que la Ecu. (3.7) proporciona una serie de valores en el rango  $[-1, 1]$ , cuya media se calcula entonces en cada banda de frecuencia (delta, theta, alfa, beta y gamma). Para los cálculos de la Ecu. (3.7), se utilizan las funciones *welch* y *csd* dentro del módulo *scipy.signal*, y la función *imag* dentro del módulo *numpy*. El umbral por defecto para esta medida es de 0,4.

### **Phase Locking Value**

El Phase Locking Value (PLV) mide las fluctuaciones de la diferencia de fase de dos series temporales. Su valor se obtiene mediante:

$$PLV = |E[\exp(i \Delta\phi_{rel}(t))]| \quad (3.8)$$

donde  $\Delta\phi_{rel}(t)$  es la diferencia de fase de  $x$  e  $y$  en el tiempo  $t$ , desplazada para estar en el intervalo  $[0, 2\pi)$ ,  $E[\cdot]$  representa una media en tiempos, y  $|\cdot|$  significa módulo, de modo que para cada par de canales se devuelve un único valor real. Según la definición proporcionada por la Ecu. (3.8), el PLV pertenece al intervalo  $[0, 1]$ , donde 1 representa

la sincronía entre las dos series temporales (pequeñas fluctuaciones en la diferencia de fase), y 0 la ausencia de sincronía (grandes fluctuaciones en la diferencia de fase). Como la fase de las señales se obtiene a través de la transformada de Hilbert, se utiliza la función *hilbert* del módulo *scipy.signal*, y la función *angle* dentro del módulo *numpy*. El umbral por defecto para esta medida es de 0,8. Hay que tener en cuenta que la implementación del PLV en este trabajo proporciona el valor asociado a una o varias bandas que es necesario especificar.

### **Weighted Phase Lag Index**

El Weighted Phase Lag Index (*WPLI*) tiene un objetivo similar al del *PLI*, aunque es menos sensible a ciertos efectos externos en comparación con éste. Se calcula mediante la expresión:

$$\begin{aligned} WPLI &= \frac{|E[Imag(G_{xy}(f))]|}{E[|Imag(G_{xy}(f))|]} \\ &= \frac{|E[|Imag(G_{xy}(f))| * sign(Imag(G_{xy}(f)))]|}{E[|Imag(G_{xy}(f))|]} \end{aligned} \quad (3.9)$$

donde  $G_{xy}(f)$  es la densidad espectral cruzada entre  $x$  e  $y$ ,  $|\cdot|$  es el módulo, y  $E[\cdot]$  representa un promedio en frecuencias, de modo que para cada par de canales devuelve un único valor real. El valor proporcionado por la Ecu. (3.9) se encuentra en el intervalo  $[0, 1]$ . El umbral por defecto para esta medida es de 0,45, y, al igual que el *PLV*, la implementación del *WPLI* en este trabajo proporciona el valor asociado a una o varias bandas que deben ser especificadas.

### **Directed Transfer Function**

La directed transfer function (*DTF*) describe la influencia causal de un canal  $j$  sobre el canal  $i$  para cada frecuencia a través de la expresión:

$$DTF_{j \rightarrow i}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_{m=1}^k |H_{im}(f)|^2} \quad (3.10)$$

donde  $H_{ij}(f)$  es un elemento de una matriz de transferencia de un modelo autorregresivo multivariante MVAR (por sus siglas en inglés “multivariate autoregressive model”). En particular, la Ecu. (3.10) constituye una versión normalizada de la *DTF*, que toma valores de 0 a 1, indicando una relación entre la afluencia del canal  $j$  al canal  $i$  con respecto a todas las afluencias al canal  $i$ . En este trabajo, el cálculo del *DTF* se realiza a través del paquete de python *SCoT*, y la implementación de dicha medida está asociada a una

o múltiples bandas que deben ser especificadas. Para cada banda se considera el valor medio en el rango de frecuencias asociado a la misma. El umbral por defecto para esta medida es de 0,3.

### **Potencia espectral**

El espectro de potencia (PS por sus siglas en inglés “Power Spectrum”) de una serie temporal,  $x$ , describe su distribución en el dominio de la frecuencia. Se calcula mediante la expresión:

$$PS(f) = |X(f)|^2 \quad (3.11)$$

donde  $X(f)$  es la transformada discreta de Fourier unidimensional de  $x$ , que se calcula utilizando la función `rfft` del módulo `numpy.fft`. La implementación de la Ecu. (3.11) permite al usuario determinar la  $PS$  para una o múltiples bandas (que deben ser especificadas), y devuelve el valor medio de la  $PS$  en el rango de frecuencias asociado a cada banda. Para establecer una relación entre pares de electrodos, se selecciona el 25 % de los canales con los valores más altos y se conectan en el grafo final.

### **Entropía espectral**

La Entropía espectral ( $SE$  por sus siglas en inglés “Spectral Entropy”) es la entropía de Shannon en el dominio de las frecuencias de una serie temporal  $x$ . Por lo tanto, la  $SE$  mide la cantidad de información contenida en su espectro. Su expresión viene dada por:

$$SE(x) = \frac{-\sum_{f=0}^{f_s/2} p(f) \log_2(p(f))}{\log_2(\text{psd\_size})} \quad (3.12)$$

donde  $\log_2$  denota el logaritmo en base 2,  $f_s$  es la frecuencia de muestreo,  $p(f)$  representa la densidad espectral de potencia normalizada de  $x$  estimada a través del método Welch, y `psd_size` es el tamaño de la densidad espectral de potencia de  $x$ . A diferencia de la medida de entropía de Shannon, el valor obtenido a través de la Ecu. (3.12) está normalizado para pertenecer al intervalo  $[0, 1]$ . Los cálculos completos se realizan utilizando la función `spectral_entropy` del paquete de python `AntroPy`. Una vez que se obtiene un valor único para el  $SE$  para cada canal, se selecciona el 25 % de los canales con los valores más altos y se conectan en el grafo final.

Las medidas descritas se incluyen en la librería desarrollada que después se utilizará para el estudio de conectividad funcional para evaluar las diferencias de conectividad por medio de la teoría de grafos y técnicas de DL en pacientes con EP.

### 3.4.3. Revisión sistemática de la literatura

Se realizó una revisión sistemática de las técnicas de ML utilizadas para tratar los EEGs de pacientes con EP y así marcar un punto de partida en la investigación. Debido a la temática multidisciplinar de esta tesis, la revisión de la literatura se realizó desde dos puntos de vista. Desde el punto de vista clínico se acotó el problema a aquellas pruebas electroencefalográficas que eran equivalentes a las pruebas que se registraron en este estudio. Desde el punto de vista de las ciencias computacionales, predominó el estudio de los modelos utilizados, por lo que no se incluyó esta limitación.

El proceso llevado a cabo para realizar la revisión sistemática siguió las pautas PRISMA, (por sus siglas en inglés "*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*"), definidas en [231] como "un conjunto mínimo de ítems basados en la evidencia destinados a ayudar a los autores a informar de una amplia gama de revisiones sistemáticas y metanálisis que evalúan los beneficios y los daños de una intervención sanitaria". Por tanto, esta metodología proporciona estándares de calidad a través de ítems relacionados con el título, el resumen, los métodos, los resultados y la discusión, a través de los cuales los autores pueden realizar evaluaciones críticas y objetivas de diversos estudios. Se puede encontrar más información sobre los protocolos de revisión en los artículos "Statement" [232] y "Explanation and Elaboration" [233].

A continuación, se detallan los pasos seguidos para la obtención de los artículos seleccionados.

- **Identificación.** Paso en el cual se realiza la búsqueda bibliográfica en las bases de datos escogidas.
- **Selección.** Eliminación de elementos duplicados y resultados que no consistieron en artículos científicos, como libros, resúmenes o pósteres.
- **Elegibilidad.** Se aplican los criterios de inclusión y exclusión propuestos para cada revisión. Se revisó el título, resumen y el artículo completo, en caso de ser necesario para determinar el cumplimiento de los criterios establecidos.
- **Inclusión.** Descartamos aquellos artículos que no han sido encontrados o han sido retractados para obtener el resultado final del proceso.

Para realizar las búsquedas en las bases de datos se determinaron las palabras clave a utilizar de manera que se ajustaran al tema propuesto de la revisión tanto desde el punto de vista médico como computacional. Desde el punto de vista médico, los temas de búsqueda fueron la enfermedad de Parkinson y la electroencefalografía. Por tanto, los términos específicos elegidos fueron: 1. Parkinson's, 2. EEG, 3. electroencephalogram.



En cuanto a los términos computacionales, las publicaciones analizadas deben utilizar herramientas y técnicas del campo de la Inteligencia Artificial, pero la terminología de ML no siempre aparece en los artículos de la misma manera, por lo que consideramos algunos de los términos genéricos más utilizados para cubrir la mayor cantidad de trabajos posible. Estos fueron: 4. machine learning, 5. deep learning y 6. neural networks. La combinación de estos términos dio lugar a la expresión de búsqueda que se introdujo en las bases de datos aportadas para cada una de las revisiones, siendo esta: “Parkinson’s” AND (“machine learning” OR “deep learning” OR “neural networks”) AND (“EEG” OR “electroencephalogram”). La expresión simplificada bajo la notación numérica especificada es: 1 AND (4 OR 5 OR 6) AND (2 OR 3).

### 3.4.3.1. Revisión sistemática desde un punto de vista clínico

La búsqueda se realizó el 19 de mayo de 2020, sin límite de tiempo, en las siguientes 6 bases de datos: Web of Science, PUBMED, Scopus, MEDLINE, CINAHL y Science Direct.

Tras la eliminación de duplicados, con el objetivo de eliminar aquellos estudios que no habían sido revisados por pares, sólo se consideraron aquellas publicaciones catalogadas como artículos de investigación, aunque no estuvieran indexadas en JCR (por sus siglas en inglés “Journal Citation Reports”) de Clarivate Analytics. Consecuentemente, se excluyeron actas, artículos de congresos, capítulos de libros, pósteres y editoriales.

Los criterios de inclusión y exclusión se fijaron de acuerdo al objetivo de esta revisión y se recogen en la Tabla 3.2.

Criterios de exclusión	Criterios de inclusión
<ul style="list-style-type: none"> <li>· Estudios que no consideran EEGs.</li> <li>· Estudios que no utilizan técnicas de ML para el análisis de EEGs.</li> <li>· Estudios que centran su análisis en otras enfermedades neurológicas.</li> <li>· Estudios en animales.</li> <li>· Estudios farmacológicos.</li> <li>· Artículos que estudian cambios evocados en EEG debido a estímulos exógenos.</li> <li>· Registro de EEG invasivos y del sueño</li> </ul>	<ul style="list-style-type: none"> <li>· Estudios prospectivos o retrospectivos usando EEGs para evaluar la progresión o el diagnóstico de la EP usando diferentes arquitecturas de ML en registros electroencefalográficos de superficie en estado despierto</li> </ul>

Tabla 3.2: Criterios de exclusión e inclusión definidos en la revisión literaria desde el punto de vista de la medicina. Fuente: [234].

Finalmente, se prescindió de aquellos estudios que no incluyeron información sobre la metodología empleada. En consecuencia, los artículos seleccionados que resultaron de este proceso consistían en estudios de la EP que buscaban diagnosticar o determinar la evolución de esta enfermedad mediante el uso de técnicas de ML en EEG de reposo o EEG de activación motora.

Se examinó el texto completo de los artículos seleccionados por dos revisores de forma independiente para obtener una puntuación en la lista de verificación propuesta por “Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research” [235]. Esta lista consta de 12 puntos que informan acerca de los elementos que deben incluirse en un artículo de investigación para garantizar que el estudio sea de buena calidad. Las categorías evaluadas incluyen requisitos dentro del campo de la biomedicina, el campo de la informática y requisitos sobre cómo estos dos campos se superponen entre sí. La descripción de estos puntos se muestra en la publicación [235]. Para facilitar su comprensión, hemos resumido estos ítems en 7 puntos que han marcado las guías para la evaluación de los artículos seleccionados.

- **Punto 1.** Contiene los ítems 1 y 2. Se evalúa la estructura y el contenido del título y el resumen del artículo.
- **Punto 2.** Contiene los ítems 3 y 4. Identifica el objetivo clínico, se revisa el estado del arte de los modelos existentes y se justifica el estudio.
- **Punto 3.** Contiene los ítems 5 y 6. Se describe el conjunto de datos aportando una valoración de su calidad y justificando el modelo elegido.
- **Punto 4.** Contiene el ítem 7. Se describe el pre-procesamiento de datos y las métricas de validación empleadas.
- **Punto 5.** Contiene el ítem 8. Se describe el modelo, proporcionando información suficiente sobre los parámetros que definen su arquitectura para su reproducibilidad. Se evalúa si los datos disponibles son suficientes para un buen ajuste del modelo.
- **Punto 6.** Contiene el ítem 9. El rendimiento predictivo del modelo se proporciona en términos de las métricas de validación.
- **Punto 7.** Contiene los ítems 10, 11 y 12: Se proporcionan las implicaciones clínicas de los resultados obtenidos, se discuten las limitaciones del estudio y se informa de los resultados inesperados.

La evaluación de cada artículo a través de la lista de verificación de [235], fue realizada por 2 evaluadores independientes, minimizando el sesgo producido por un único revisor. Para medir el consenso entre ambas evaluaciones, teniendo en cuenta la opción de acuerdo por casualidad, se calculó el valor kappa ( $\kappa$ ) entre ambas evaluaciones. Un valor de  $\kappa > 0,7$  significa alto nivel de acuerdo entre evaluadores, un valor de  $\kappa$  acotado en el intervalo (0,5-0,7) se asocia a un nivel de acuerdo moderado y un valor  $\kappa < 0,5$  se relaciona con un bajo nivel de acuerdo. Este procedimiento dio lugar a una evaluación objetiva del contenido de cada artículo para proceder a la comparativa de la información incluida en cada uno de ellos.

Para facilitar la evaluación del contenido de cada artículo incluido en la revisión, de cada uno se extrajo la información asociada a los siguientes temas:

- **La calidad del dataset.** A través de parámetros clínicos y técnicos como el número de pacientes incluidos en el estudio, el estadio de la enfermedad, el tipo de prueba electroencefalográfica realizada y los parámetros asociados con el registro del EEG.
- **Pre-procesamiento de datos.** Por medio del proceso de limpieza del EEG y los métodos de extracción de características.
- **Técnicas de ML utilizadas.** A través de criterios de validación, calidad del proceso de entrenamiento/validación, métricas utilizadas y resultados de cada modelo.

Estos campos fueron elegidos con el fin de sintetizar la información más relevante dentro de cada uno de los artículos de acuerdo con los ítems de la lista de verificación que aparece en [235]. Además, facilita el estudio comparativo de los pares modelo-parámetros para los cuales, dependiendo del objetivo del estudio, se han obtenido mejores resultados. Las conclusiones se obtuvieron, por un lado, comparando, para cada uno de estos puntos, la información recogida en los diferentes artículos, y, por otro lado, valorando los resultados obtenidos por cada artículo en relación a los parámetros utilizados.

#### **3.4.3.2. Revisión sistemática desde el punto de vista de las ciencias computacionales**

La búsqueda de la revisión desde el punto de vista de las ciencias computacionales se realizó el 14 de febrero de 2022 sin límite de tiempo, en las siguientes bases de datos: Web of Science, PUBMED, Scopus, CINAHL y Science Direct. Nótese que, con respecto a la revisión anterior, hemos considerado una base de datos menos (MEDLINE) ya que esta ha sido absorbida por PUBMED.

Tras la eliminación de los elementos duplicados, se descartaron contenidos basados en libros académicos, capítulos de libros, resúmenes y pósteres por considerarlos fuera del alcance de esta revisión. No se excluyeron otras publicaciones para favorecer la evaluación de las técnicas utilizadas en el campo del ML.

Los criterios de exclusión aplicados se muestran en la Tabla 3.3, ordenados de acuerdo con los objetivos propuestos para esta revisión. Las razones por las cuales se eligieron estos criterios también han sido incluidas en la tabla.

<b>Criterios de exclusión</b>	<b>Razón de la exclusión</b>
Estudios centrados en enfermedades neurológicas diferentes a la EP	Fuera del objetivo de esta revisión.
Estudios que no utilizan técnicas de ML	Fuera del objetivo de esta revisión.
Estudios que no consideran EEGs	Fuera del objetivo de esta revisión.
Estudios con EEGs invasivos	Las señales de EEG invasivo no son comparables con las de EEG no invasivo. Además, se eligió el EEG no invasivo por ser de bajo coste, ampliamente disponible y fácil de adquirir, propiedades que no comparte el EEG invasivo.
Estudios en animales	Los resultados obtenidos en estudios con animales no siempre pueden extrapolarse a los seres humanos. Además, los electroencefalogramas de animales y humanos no son comparables.
Estudios farmacológicos	Los estudios centrados en el desarrollo y análisis de los componentes de los medicamentos quedan fuera de los objetivos de esta revisión.
Artículos de revisión	Fuera de los objetivos de esta revisión.

Tabla 3.3: Criterios de exclusión junto con la razón de la exclusión, definidos en la revisión literaria desde el punto de vista de las ciencias computacionales. Fuente: [236].

Aunque todos los estudios contenían los términos de búsqueda especificados, algunos de ellos sólo los utilizaron como referencia a otros trabajos, o los mencionaron sin desarrollarlos. Estos casos se clasificaron como estudios que no utilizaron técnicas de ML, estudios que no consideraron el EEG o estudios que no se centraron en la EP.

Tras esta fase, se descartaron los estudios que no fueron encontrados o fueron re-tractados y, por lo tanto, no pudieron pasar por la fase de elegibilidad. El resultado de este último paso proporcionó los artículos que constituyeron esta revisión y que fueron considerados para su posterior análisis. Para cada uno de los artículos seleccionados, se extrajo la información descrita en la Tabla 3.4:

Ítem	Descripción
1. Calidad del dataset	A través de parámetros clínicos y técnicos como el número de pacientes incluidos en el estudio, el estadio de la enfermedad, la administración de la medicación y el tipo de pruebas de EEG realizadas.
2. Pre-procesamiento de datos	A través del protocolo de limpieza del EEG y los métodos de extracción de características.
3. Análisis de las técnicas de ML	A través de los tipos de modelos, la arquitectura del modelo, la evaluación de la calidad del proceso de entrenamiento/validación, las métricas utilizadas y los resultados de cada modelo.

Tabla 3.4: Descripción de los ítems considerados para el análisis de los artículos. Fuente: [236].

Estos aspectos fueron elegidos para sintetizar la información más relevante dentro de cada uno de los artículos incluidos en esta revisión, con el fin de analizar cada ítem y brindar un punto de partida para futuros estudios.

Las conclusiones derivadas de esta revisión se obtuvieron, por un lado, comparando, para cada uno de estos puntos, la información recogida en los diferentes artículos, y, por otro lado, valorando los resultados obtenidos por cada artículo en relación con los parámetros utilizados.

#### 3.4.4. Modelado de EEGs como grafos para el análisis de conectividad por medio GCNN

Para poder llevar a cabo los estudios de conectividad es necesario contar tanto con un test clínico que contenga la información sobre el estado del cerebro del paciente como con una metodología que represente eficazmente la conexión entre las regiones cerebrales.

La teoría de grafos ofrece una buena manera de representar la conectividad del cerebro debido a su potencial para modelar las redes. Para ello, tanto los nodos como las aristas del grafo deben ser definidos en función del problema considerado. En [237], se realizó una revisión de los diferentes procedimientos para definir los nodos del grafo a partir de un EEG, según el tipo de estudio que se quiera realizar. Una de las metodologías comúnmente utilizadas consiste en relacionar los nodos del grafo con diferentes regiones del cerebro, de forma que las aristas determinan las conexiones entre ellas. Además, el uso de la teoría de grafos permite obtener la información de las conexiones de dos formas, visual y matricial. La representación visual del grafo puede ayudar al personal clínico en su práctica diaria, mientras que su representación matricial proporciona información a los investigadores como estructuras de datos con las que trabajar más cómoda y eficazmente en sus investigaciones sin perder la información contenida en el grafo [238].

Aunque la teoría de grafos ha sido considerada anteriormente en los estudios de conectividad, el proceso es todavía intrincado y no está extendido en la práctica clínica. Además, las pruebas utilizadas para obtener información del cerebro suelen ser caras e invasivas para el paciente. Aunque la electroencefalografía es una técnica inocua, estas señales son difíciles de analizar y no ha sido hasta los últimos años cuando se ha generalizado su uso en el análisis de la conectividad. La dificultad del análisis de estas señales junto con la escasez de pruebas realizadas, reducen la cantidad de datos disponibles, limitando este tipo de estudios y afectando a la investigación actual. Para lidiar con estos inconvenientes hemos desarrollado una librería de Python denominada EEGraph, que proporciona una representación del EEG mediante un grafo. Esta librería se ha construido con el objetivo principal de proporcionar una herramienta que facilite el desarrollo de estudios posteriores dentro de este campo.

Para evaluar la conectividad, EEGraph cuenta con las 12 medidas que hemos especificado, permitiendo abordar estudios relacionados tanto con el espectro de frecuencias como con el dominio temporal del EEG. Esta amplia variedad de medidas elegidas para evaluar la conectividad, junto con la flexibilidad que ofrece la librería para definir los parámetros implicados, proporcionan al usuario una potente herramienta que facilita la investigación en el área de la neurociencia. El resultado aportado por EEGraph consiste en una representación visual, a través de una imagen que facilita una rápida aplicación clínica, y una representación matricial, a través de estructuras de datos fácilmente manipulables para el análisis cuantitativo. Ambas representaciones proporcionan una posible aplicación en el ámbito clínico y una herramienta útil para el campo de la investigación en neurociencia.

Una vez desarrollada la herramienta, esta se utilizará para modelar los EEGs de los

pacientes con EP y poder realizar el estudio de conectividad. Para ello, contamos con los EEGs de pacientes antes y después de tomarse su medicación y con los EEGs de controles sanos que han sido seleccionados de modo que ambos grupos tuvieran las mismas características demográficas y fueran estadísticamente equivalentes. El análisis se llevará a cabo por medio de una GCNN, una red encargada de analizar los grafos obtenidos de los pacientes. Este modelo se encargará de extraer los patrones de conectividad por medio de un módulo de grafos, seguido de un módulo de redes convolucionales y finalizando con un módulo de clasificación para poder realizar el diagnóstico para un paciente.

#### 3.4.4.1. EEGraph

La librería EEGraph (cuyo código está disponible en GitHub<sup>3</sup>) es una herramienta que ofrece al usuario una forma rápida y eficiente de realizar diferentes estudios de conectividad a partir de EEGs. Utiliza la teoría de grafos para representar las conexiones cerebrales, identificando los nodos del grafo con los electrodos del casco de electroencefalografía y preservando su distribución espacial. Las conexiones entre los nodos, es decir, las aristas del grafo, están determinadas por las diferentes medidas de conectividad disponibles en la librería.

Actualmente, existen librerías de Python que o bien se centran en el procesamiento de los EEGs, o bien se dedican al manejo de grafos, pero, hasta donde sabemos, no existe ninguna librería que cubra ambos aspectos, es decir, el modelado de EEGs como grafos. Este hecho pone de manifiesto la necesidad de contar con una herramienta versátil y fácil de usar que conecte ambos enfoques.

Entre las librerías que realizan el procesamiento del EEG, encontramos *eeglib* [239], una herramienta dedicada a la extracción de características del EEG (a través de 8 métodos), además de incluir otros procedimientos para el análisis de la señal. Sin embargo, no proporciona ninguna representación visual del resultado. El paquete *Mne* [198] está diseñado para la exploración, visualización y análisis de datos neurofisiológicos, EEG incluido. En particular, soporta 6 medidas de conectividad y sus variaciones. Aunque esta librería permite una gran flexibilidad para la elección de ciertos parámetros, es bastante compleja y poco intuitiva, además de no construir el grafo asociado. Finalmente, encontramos la librería *SCoT* [201], dedicada a la descomposición de fuentes y a la estimación de la conectividad en EEGs a partir de medidas espectrales como la Coherencia, la Coherencia Parcial Dirigida, la DTF y sus variaciones. Esta librería contiene un número limitado de medidas de conectividad además de no ofrecer el modelado de EEGs mediante grafos.

<sup>3</sup><https://github.com/ufvceiec/EEGRAPH>

Entre las librerías de Python enfocadas al manejo de grafos, encontramos *NetworkX* [199], que permite la creación, manipulación, análisis y visualización de grafos, digrafos y multigrafos, junto con *spektral* [240] y *StellarGraph*<sup>4</sup>, que manejan estructuras de datos tipo grafo para su uso en modelos de DL y ML, respectivamente.

Por lo tanto, EEGGraph cubre el vacío existente en el modelado de EEG a través de grafos sirviendo de enlace entre estos dos campos. Además, se compone de 12 medidas de conectividad diferentes, además de proporcionar una gran flexibilidad para definir los parámetros relacionados con el modelado de grafos y medidas de conectividad en un entorno intuitivo para el usuario.

En el Apéndice F se facilita una guía de usuario acerca de la librería, especificando los parámetros que más adelante se utilizarán para el estudio de conectividad funcional de pacientes con EP, así como las especificaciones del software.

La salida que obtenemos de EEGGraph es doble. Por un lado, ofrece una representación visual del grafo mediante una imagen con los nodos y las conexiones, que depende de la medida de conectividad escogida. Por otro lado, EEGGraph proporciona una representación matricial del grafo, que contiene todos los valores de la medida de conectividad elegida por el usuario. Si el EEG se ha dividido en varias ventanas, EEGGraph ofrece ambas representaciones, visual (Fig. 3.3) y matricial (Fig. 3.4), del grafo para cada una de ellas, lo que permite al usuario evaluar la evolución temporal de las conexiones cerebrales.

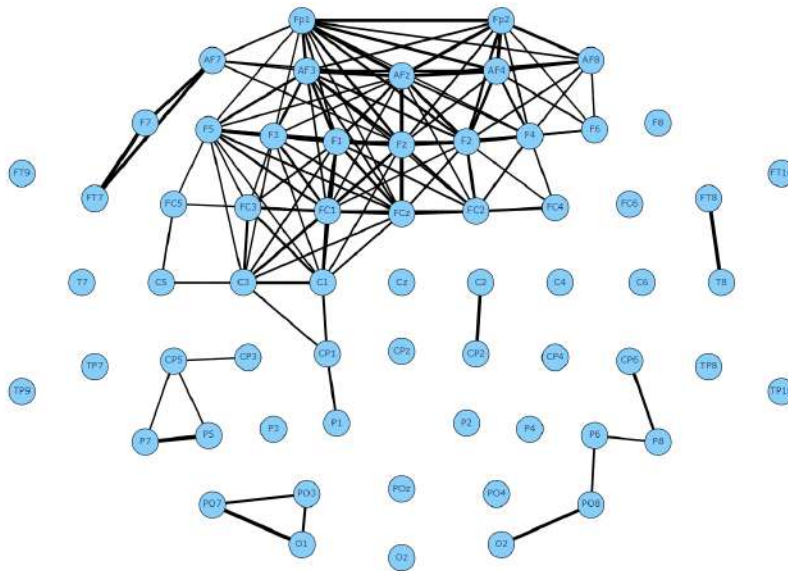


Figura 3.3: Representación visual del grafo obtenida mediante EEGGraph, para el ejemplo de un EEG aportado por la librería. Fuente: elaboración propia.

<sup>4</sup><https://stellargraph.readthedocs.io/en/stable>



```
[ [1.          0.90219588 0.75456562 ... 0.92090921 0.88247858 0.77651935 ]
  [0.90219588 1.          0.87692898 ... 0.82488    0.92291208 0.9472386  ]
  [0.75456562 0.87692898 1.          ... 0.62597206 0.72839699 0.82652213 ]
  ...
  [0.92090921 0.82488    0.62597206 ... 1.          0.9209615 0.65990146 ]
  [0.88247858 0.92291208 0.72839699 ... 0.9209615 1.          0.82249207 ]
  [0.77651935 0.9472386 0.82652213 ... 0.65990146 0.82249207 1.          ] ]
```

Figura 3.4: Representación matricial del grafo anterior obtenida mediante EEGraph. Fuente:elaboración propia.

### 3.4.5. Adaptación de técnicas de PLN para el análisis de los EEGs

Los modelos del campo del PLN pueden relacionar diferentes partes de una secuencia de entrada para producir el resultado deseado [241, 173, 174], dando lugar a la contextualización de un texto. Esta propiedad es la que buscamos al utilizar estos modelos en los EEGs de pacientes con EP, de manera que busquen la relación entre diferentes partes del EEG para encontrar las “reglas gramaticales” que definen la enfermedad. Para poder aplicar estas técnicas primero necesitamos conocer las bases del procesamiento requerido en el caso del texto para facilitar la comprensión del procesamiento del EEG.

El pre-procesamiento de texto en el campo del PLN se divide en varias etapas. En un primer paso, se descompone el texto en fragmentos más pequeños, o tokens, que pueden ser palabras o sílabas, mediante una técnica conocida como tokenización. Con los tokens podemos construir un vocabulario sobre el que trabajará el modelo. Suponiendo que la tokenización se produce a nivel de palabra, el siguiente paso es vectorizar cada uno de los tokens para obtener una representación numérica. Este proceso se lleva a cabo por medio de un embedding, proceso por el cual se genera un espacio vectorial cuyos elementos son la representación vectorial de cada palabra, situadas de modo que la distancia entre ellas sea una medida de la relación semántica entre dichas palabras. Por último, para mantener la secuencia de los datos, o dicho de otro modo, el orden de las palabras en el texto, se utiliza el “positional encoding”, que se encarga de etiquetar cada una de las palabras con su posición en el texto.

Para llevar a cabo el problema de diagnóstico de la EP, utilizamos uno de los modelos más conocidos en el área del PLN, un modelo BERT (por sus siglas en inglés “Bidirectional Encoder Representation from Transformers”) [174], para detectar los cambios del EEG de los pacientes en comparación con controles sanos. El modelo BERT basado en modelos Transformers, se ha adaptado para tomar como entrada los canales del EEG y poder desarrollar un modelo de clasificación.

Para poder tratar los EEGs con las técnicas del campo del PLN, hemos establecido una equivalencia texto-EEG. Esta equivalencia parte de la premisa de que ambos datos

son datos secuenciales, y por tanto podemos encontrar un modo de relacionar una frase con un canal del EEG. Consecuentemente, el EEG completo equivale a un texto compuesto de 64 frases que contiene información acerca del estado del cerebro del paciente. Para recrear las palabras del texto, hemos dividido el EEG en segmentos temporales de una longitud determinada, de modo que las señales del EEG se evalúan teniendo en cuenta lo que ocurre en todos los canales para un intervalo de tiempo dado. Por cada ventana del EEG se obtiene una matriz de dimensión  $64 \times N$ , donde  $N$  se corresponde con el número de puntos que contiene cada ventana, y 64 se refiere al número de electrodos del EEG. Mediante un mecanismo de ventana deslizante, definido por medio de dos parámetros encargados de crear las instancias, se recorre el EEG introduciendo cada ventana en el modelo. El primer parámetro se llama paso o “step” y controla cuánto se desplaza el inicio de una ventana respecto a un instante del EEG, que es el inicio de una ventana anterior. El segundo parámetro es la anchura y controla el número de valores entre el inicio y el final de la ventana. Tras evaluar los resultados ofrecidos por distintos parámetros, los que hemos escogido para realizar el estudio son: para el paso tomamos el 95 % del tamaño de ventana (esto es un solapamiento de un 5%) y para el ancho, un valor de 256 instancias. Un esquema de este proceso se muestra en la figura 3.5.

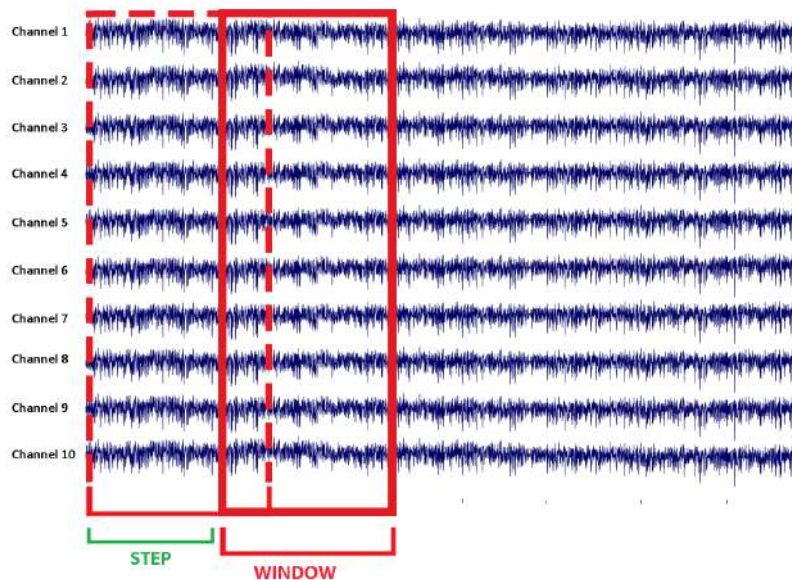


Figura 3.5: Tamaño de ventana y paso empleados en el modelo BERT para realizar la equivalencia texto-EEG. Fuente: elaboración propia.

Por tanto, el modelo toma como entrada la matriz anterior traspuesta, donde cada punto del EEG se corresponde con una palabra cuyo objetivo es el de realizar una cla-

sificación binaria (Parkinson, No-Parkinson) para cada ventana temporal del EEG. La probabilidad de pertenecer a cada una de las clases se obtiene por medio del valor medio de las probabilidades de cada uno de los fragmentos que componen el EEG del paciente.

No es necesario llevar a cabo el embedding, ya que una serie temporal tiene, por sí misma, una representación numérica, y por tanto, no es necesario codificarla para obtener una representación vectorial.

Para realizar el estudio se diseñaron dos experimentos basados en modelos BERT con la misma arquitectura, para facilitar la interpretación de los resultados y permitir la comparativa entre estudios. En uno de ellos asumimos que los canales periféricos añaden ruido a la hora de predecir si un paciente tiene EP [242], de modo que nos centramos en el procesamiento de los 28 canales concentrados en la región interior que recogen la información de las áreas motoras y sensitivas primarias. El otro experimento recibe el EEG completo con los 64 canales, de modo que utiliza toda la información registrada en el electroencefalograma. Para ambos experimentos se realizan dos estrategias de entrenamiento diferentes. En la primera de ellas se utilizan todas las pruebas de electroencefalografía (esto es Resting y pruebas motoras FT), y en la segunda, se eliminan las pruebas de reposo para realizar el entrenamiento únicamente con las pruebas asociadas a la activación motora de ambos hemisferios. Estos 4 experimentos se realizaron de forma independiente para evitar sesgos en los resultados.

### **3.4.6. Métodos de validación de resultados**

#### **3.4.6.1. Métricas**

Las métricas se utilizan en ML para proporcionar uno o varios valores que nos permitan evaluar el proceso de entrenamiento de los modelos y su capacidad para generalizar resultados. La elección de las métricas va asociada generalmente al tipo de problema que queramos abordar. Dado que nos encontramos en el campo de la medicina y buscamos una aplicación clínica, las métricas que debemos considerar tienen que estar orientadas a facilitar su interpretación y ayudar a la validación de los resultados dentro de este campo. Por ello utilizamos aquellas medidas que ofrecen información acerca de los fallos y aciertos obtenidos por el modelo: falsos positivos (FP por sus siglas en inglés “False Positive”), en nuestro contexto se corresponde con un control que ha sido diagnosticado con EP, falsos negativos (FN por sus siglas en inglés “False Negative”), en nuestro contexto un paciente con EP que ha sido diagnosticado como control, verdaderos positivos (TP por sus siglas en inglés “True Positive”), encargada de estimar aquellos pacientes con EP que han sido identificados correctamente por el modelo, y

verdaderos negativos (TN por sus siglas en inglés “True Negative”), indica los controles que son correctamente identificados por el modelo [243].

Aunque la métrica estándar para evaluar el aprendizaje es la exactitud, conocida en el campo de la IA como accuracy, que puede obtenerse a partir los valores TP, TN, FP y FN, como podemos ver en la ecuación 3.13, esta métrica no contiene la información de todo el proceso de entrenamiento ni permite evaluar los resultados y fallos del modelo desde diferentes enfoques. Por ello es necesario introducir otras métricas que complementen esa información. Podemos observar un esquema de las métricas introducidas en la Fig 3.6.

Matriz de confusión		Estimado				
		Negativo (N)	Positivo (P)			
Real	Negativo (N)	TN	FP	Especificidad <i>specificity</i>	% casos negativos detectados	TN/(TN+FP)
	Positivo (P)	FN	TP	Sensibilidad <i>recall</i>	% casos positivos detectados	TP/(TP+FN)
				Precisión <i>precision</i>	Exactitud <i>accuracy</i>	% predicciones correctas (TN+TP)/(TN+TP+FN+FP)
				% predicciones positivas OK TP/(TP+FP)		

Figura 3.6: Métricas consideradas para la evaluación de los resultados de los modelos junto con su definición. Fuente: elaboración propia.

Dentro de la gran variedad de métricas existentes, utilizaremos las cuatro métricas más utilizadas dentro del campo de la medicina. Estas son exactitud, especificidad, sensibilidad y precisión. Estas métricas se pueden calcular a partir de los valores anteriores. Su expresión junto con con una breve descripción acerca de su interpretación se proporciona a continuación:

- **Accuracy o exactitud.** Es la métrica utilizada por excelencia durante el proceso de entrenamiento, viene definida por:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}. \quad (3.13)$$

Es el número de predicciones correctas dividido por el número total de predicciones realizadas. Esta métrica se encarga de evaluar el conjunto de aciertos del modelo en función de la cantidad total de datos.

- **Sensitivity o sensibilidad.** También conocida en inglés como “recall”, “hit rate” o

tasa de TP, viene dada por la siguiente expresión:

$$TPR = \frac{TP}{TP + FN}. \quad (3.14)$$

Esta métrica proporciona información acerca de la tasa de verdaderos positivos que ha obtenido el modelo, lo que es muy útil para evaluar la cantidad de pacientes que no han sido diagnosticados por la enfermedad.

- **Specificity o especificidad.** También conocida en inglés como “selectivity” o tasa de TN, viene definida como:

$$TNR = \frac{TN}{TN + FP}. \quad (3.15)$$

Esta métrica proporciona información acerca de la tasa de verdaderos negativos que ha obtenido el modelo, permitiendo evaluar la cantidad de pacientes sanos que han sido diagnosticados con la enfermedad y, por tanto, esta métrica evita que personas sanas tomen la medicación y se sometan a las pruebas de diagnóstico de los pacientes con EP.

- **Precision o precisión.** También conocida en inglés como “positive predictive value” (PPV), viene dada por:

$$PPV = \frac{TP}{TP + FP}. \quad (3.16)$$

Esta métrica mide la relación entre el número de verdaderos positivos y el total de los predichos como verdaderos positivos y falsos positivos, lo cual es interesante en términos de costes económicos.

En los estudios llevados a cabo se proporcionan estas cuatro métricas para evaluar los resultados de los modelos y evaluar su capacidad como prueba diagnóstica de la enfermedad.

#### 3.4.6.2. Validación

El problema sesgo-varianza, que se encuentra tanto en estadística como en las técnicas de ML, se produce al intentar minimizar simultáneamente estas dos fuentes de error que impiden que los algoritmos de aprendizaje supervisado generalicen más allá de su conjunto de entrenamiento [244]. El sesgo de un estimador (o función de sesgo) es la diferencia entre el valor esperado de este estimador y el valor real del parámetro que se estima. Un alto valor en el sesgo puede hacer que el algoritmo no encuentre las

relaciones relevantes dentro del conjunto de datos, ya que se encarga de realizar suposiciones erróneas dando lugar al underfitting. La varianza es la esperanza de la desviación al cuadrado de una variable aleatoria con respecto a su media poblacional o a su media muestral, esto es,

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2, \quad (3.17)$$

donde  $n$  es el número de muestras,  $x_1, \dots, x_n$  las muestras consideradas para la estimación y  $\mu$  su media muestral. La varianza es un error derivado de la sensibilidad a pequeñas fluctuaciones en el dataset, un valor alto puede ser el resultado de que un algoritmo modele el ruido aleatorio de los datos de entrenamiento, dando lugar al overfitting.

Para asegurar que los resultados proporcionados por el modelo no están influenciados por el sesgo o varianza debido a la elección de los datos de entrenamiento, se realiza un proceso de validación que tiene como objetivo evaluar los resultados en cada una de las etapas del proceso de entrenamiento validación y prueba. Para asegurar la generalización de los resultados de cada una de las métricas implementadas, se realiza el método de validación cruzada. La validación cruzada (CV por sus siglas en inglés “Cross-Validation”) es una técnica utilizada para evaluar los resultados de un análisis estadístico y garantizar que son independientes de la partición entre datos de entrenamiento y prueba. Particularmente, realizaremos un  $k$ -fold CV, procedimiento por el cual se realiza una división de los datos en  $k$  subconjuntos o folds con  $k$  procesos de entrenamiento. En cada uno de esos procesos, el modelo selecciona un subconjunto de datos para validar los resultados y utiliza las  $k - 1$  restantes para el entrenamiento. En cada repetición la fold utilizada para validación va cambiando para recorrer el conjunto completo de datos y así obtener para cada ejecución los valores de las métricas implementadas. Un esquema del proceso de  $k$ -fold CV se muestra en la figura 3.7. Una vez realizado el método  $k$ -fold CV se tomará, para cada métrica, el valor medio junto con la desviación estándar, obtenidos durante las  $k$  ejecuciones del método.

La división realizada depende tanto del tamaño de la muestra como de la cantidad de datos que necesite el modelo para realizar los entrenamiento. En nuestro caso, hemos tomado una división de 5-fold para mantener una división 80 % de los datos en el entrenamiento y 20 % de los datos se utilizaron para validación.

Para evaluar la capacidad de generalización del modelo de forma más precisa, realizamos una primera división del dataset, extrayendo un conjunto pequeño de prueba (como muestra ciega), que se guardará para evaluar la capacidad de generalización del modelo, y que introduciremos una vez terminemos el proceso de entrenamiento y validación. Este conjunto de datos se extrae del dataset original una vez se han barajado los pacientes manteniendo el porcentaje de casos en cada clase. Este procedimiento

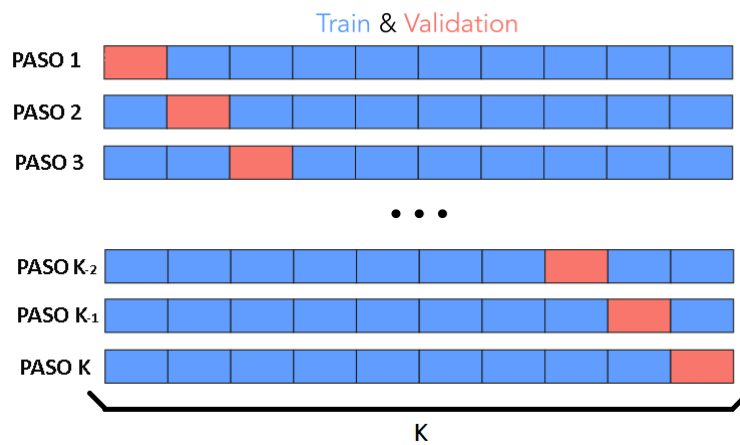


Figura 3.7: Ejemplo que muestra cómo se realiza la selección de subconjuntos durante la ejecución del método  $k$ -fold CV. Fuente: elaboración propia.

disminuye el sesgo producido por la elección de los datos. Por tanto, el proceso seguido en la validación fue:

- Barajar los datos para minimizar el sesgo y favorecer la generalización de los resultados del modelo.
- Extraer el conjunto de prueba manteniendo el porcentaje de pacientes de cada clase (esto es estratificando los datos).
- El resto del conjunto de datos fue utilizado en un 5-fold CV, estratificando los datos para cada una de las carpetas del método.
- Una vez terminado el entrenamiento junto con la validación de cada fold, se introdujo el conjunto ciego de prueba para evaluar la generalización de cada uno de los entrenamientos individuales.
- Se obtuvo el valor medio y la std para cada una de las métricas durante todo el proceso de validación cruzada.





## Capítulo 4

# Resultados

### 4.1. Análisis de series temporales aplicado a EEGs

En la Introducción validamos la *hipótesis 1* al encontrar numerosos estudios que afirman que el EEG proporciona información suficiente para identificar características intrínsecas de la EP. Por lo que comenzamos con la búsqueda del biomarcador por medio de la *hipótesis 2*.

Las deficiencias funcionales descritas en la EP se restauran en gran medida con la administración de levodopa [44], o con la estimulación cerebral profunda. Dado que ambos tratamientos están relacionados con cambios en la conectividad cerebral [245, 246], estos cambios se han considerado como características clave de la enfermedad y se han estudiado a fondo utilizando diferentes pruebas (EEG, MRI y MEG).

Para estudiar la conectividad funcional de los pacientes con EP y controles, dada la naturaleza de las señales del EEG, se han considerado varias medidas utilizadas en el análisis de series temporales. También se estudian las diferencias inter-hemisféricas, dado que la EP es una enfermedad predominantemente asimétrica, y existen algunos estudios sobre cambios significativos en la conectividad según el lado de la activación motora. Para estudiar la conectividad en pacientes con EP y evaluar los cambios producidos por la medicación utilizaremos el coeficiente de asortatividad [247], que es una medida indirecta de la resiliencia de la red, es decir, de la vulnerabilidad al daño. Esta propiedad es especialmente adecuada para el estudio de la EP, ya que su alteración funcional cortico-talámica en los ganglios basales apunta a una red funcional menos resiliente. También utilizaremos otras medidas de conectividad como el coeficiente de correlación, la coherencia y la entropía espectral.

Para llevar a cabo este estudio se formularon tres hipótesis. Primero, se estima que las diferencias de conectividad (asortatividad de la red) estarán presentes entre los con-

troles sanos y los pacientes con EP, especialmente durante la activación motora. En segundo lugar, dado que existe una lateralización funcional y estructural típicamente asimétrica característica de la EP, buscamos diferencias de conectividad inter-hemisféricas. Por tanto, la asortatividad de la conectividad inter-hemisférica también mostrará diferencias entre la EP y los controles. La tercera hipótesis fue que, dado que la acción de la dopamina determina cambios en la conectividad, esperamos encontrar cambios de asortatividad de la conectividad entre el EEG pre y post levodopa de los pacientes con EP.

Los pacientes incluidos en este estudio fueron 6 pacientes con EP, en estados pre y post LD, y 6 controles emparejados en edad, hemisferio dominante y género. Las características demográficas y clínicas de todos los participantes se muestran en la Tabla 4.1.

	Controles	EP	SIG(p)
N (Hombres)	6(5)	6(5)	1,000
Edad en años	68,83 ± 6,76	69,83 ± 7,36	0,811
Educación en años	15,17 ± 1,60	15,33 ± 5,46	0,945
HY	-	<i>HY1 = 2; HY2 = 1;</i> <i>HY3 = 3; Mo = 3</i>	-
UPDRS III	-	15,33 ± 9,63	-
Total dopamina por día (mg)	-	972,08 ± 565,31	-

Tabla 4.1: Datos demográficos y clínicos de los participantes incluidos en el estudio basado en el análisis de series temporales. Se muestra el valor medio y la desviación estándar de los datos demográficos y clínicos de los participantes. Fuente: [248].

Los resultados estadísticos de la Tabla 4.1 no muestran diferencias entre los grupos de pacientes con EP y controles en sexo ( $\chi^2(1) = 0,00$ ;  $p = 1,000$ ), edad ( $t(10) = -0,245$ ;  $p = 0,811$ ) y nivel educativo ( $t(6) = -0,072$ ;  $p = 0,945$ ).

Previo a aplicar el protocolo de limpieza de artefactos a las señales de EEG, se determinó la región de interés a analizar. Esta región se constituye por aquellos electrodos que rodean la corteza motora, de modo que eliminamos los canales periféricos que pueden contaminar la señal al registrar la actividad muscular. Los canales que fueron descartados son: Fp1, Fp2, AF8, F8, FT8, T8, TP8, P8, PO8, O2, Oz, O1, PO7, P7, TP7, T7, FT7, F7, AF7, FT9, PT9, PO9. La Fig. 4.1 muestra un esquema de los electrodos descartados (en azul) y los canales que se han considerado en el estudio.

Tras la elección de los electrodos, se realizó el proceso de limpieza de artefactos y

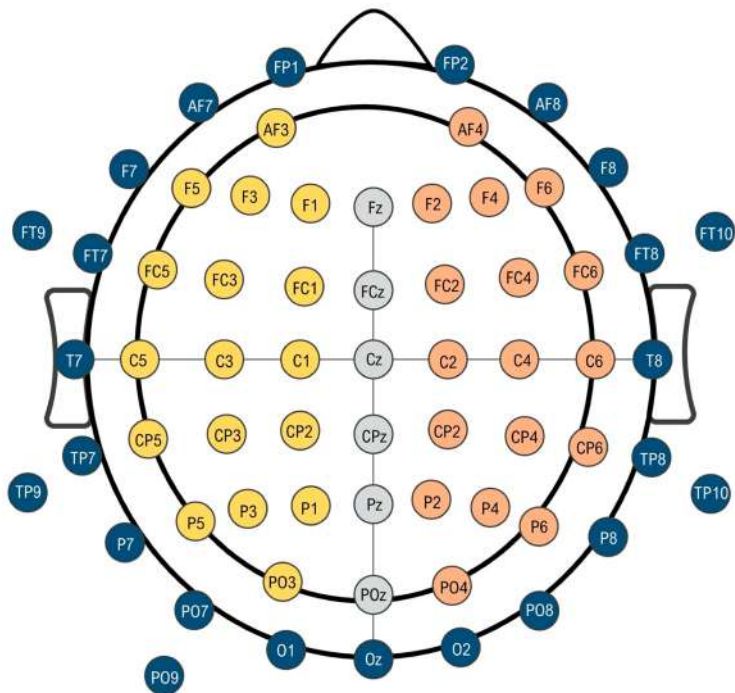


Figura 4.1: Representación esquemática del montaje de EEG de 64 canales. El círculo interior representa la región de interés en nuestras medidas. Los electrodos amarillos corresponden al hemisferio izquierdo (lado dominante) y los electrodos naranjas corresponden al hemisferio derecho (no dominante). Se han elegido los electrodos centrales (grises) como referencia en las medidas de asimetría. Los electrodos azules se han eliminado sistemáticamente del análisis. Fuente: [248].

reconstrucción de canales anómalos según el protocolo descrito en el Apéndice C.

#### 4.1.1. Aplicación de las medidas de análisis de series temporales

En la aplicación de las medidas de conectividad debemos tener en cuenta que algunas de ellas, como la entropía, dependen de la longitud de la señal, de modo que todos los EEGs fueron segmentados a la misma longitud de 30 segundos para favorecer la comparación entre los resultados de las medidas. Las señales están acotadas al intervalo (1 – 30) Hz mediante un filtro FIR de pasa banda con Ventana de Hamming, que se corresponde con las bandas delta, theta, alpha y beta.

El análisis de las series temporales se ha realizado asumiendo que las señales del EEG son estacionarias, de forma que las propiedades estadísticas no varían en el tiempo. Las medidas de conectividad calculadas fueron:

- **Coherencia.** Calculada como la densidad espectral cruzada por pares normalizada por el producto de las densidades autoespectrales de ambas señales. Se encarga

de medir la sincronización entre las bandas de frecuencia.

- **Entropía espectral.** Medida encargada de cuantificar la irregularidad de la señal operando en el espectro de frecuencias de la señal [249]. Esta medida permite independizar las contribuciones de cualquier rango de frecuencias [250].
- **Correlación de Pearson.** Indica el grado de similitud entre la información captada por los electrodos, cuantificando la cantidad de interdependencia entre ellos y, por tanto, su conectividad. Se aplica sobre los datos calculados de la entropía espectral y la coherencia para evaluar el grado de semejanza. Los resultados de la correlación no se presentarán como valores medios, sino como una distribución de valores individuales por pares (coeficiente de asortatividad).

Debemos de tener en cuenta que tanto la coherencia como la entropía espectral, proporcionan información complementaria y global sobre las señales del EEG. Finalmente, las entropías relativas entre los electrodos de ambos hemisferios se han obtenido como la Divergencia de Kullback-Leibler (KLD) [251] de la entropía espectral entre un electrodo con cualquier otro del hemisferio opuesto. El KLD de la entropía espectral, se ha utilizado como estimador de la divergencia inter-hemisférica. Aunque no es una medida directa de la asimetría entre hemisferios proporciona una idea acerca de la cantidad de información no compartida entre electrodos de ambos hemisferios. A este concepto lo llamamos independencia. La correlación entre electrodos, para un mismo paciente se obtiene aplicando esta función sobre los datos de la entropía espectral y la coherencia en el rango de frecuencias indicado.

Para comparar las distribuciones obtenidas, hemos realizado una prueba no paramétrica de Kruskal-Wallis (KW). Para evaluar la diferencia entre los valores de las medianas en los casos “*Pre\_LD*” vs. “*Controles*”, “*Post\_LD*” vs. “*Controles*”, “*Pre\_LD*” vs. “*Post\_LD*” utilizamos la prueba U de Mann-Withney y una prueba no paramétrica de suma de rangos de Wilcoxon para “*Pre\_LD*” vs. “*Post\_LD*”, con un nivel de significación de 0,05 y la hipótesis alternativa de valores desiguales. Por lo tanto, p-valores pequeños sugieren diferencias significativas entre las poblaciones.

No se ha realizado ningún promedio espacial en las propiedades mencionadas, de manera que todos los valores por pares de la matriz de conectividad se han incluido en nuestro análisis estadístico y se han resumido en box plots o diagramas de cajas. Los resultados de este estudio han sido publicados en el artículo “Time Series Analysis Applied to EEG Shows Increased Global Connectivity during Motor Activation Detected in PD Patients Compared to Controls” [248].

## 4.1.2. Resultados de las medidas de conectividad

### 4.1.2.1. Coeficiente de correlación

Comenzamos analizando la correlación de la entropía espectral. Los resultados del análisis pueden verse en la Fig. 4.2, donde se puede observar que todas las señales, en cada una de las pruebas registradas, parecen estar fuertemente correlacionadas.

La prueba de Kruskal Wallis indica que no hay diferencias significativas entre el valor medio del coeficiente de correlación de la entropía espectral entre los electrodos en la prueba de estado de reposo para los sujetos con EP en comparación con los controles ( $p = 0,09$ ). Sin embargo, sí encontramos diferencias significativas en las pruebas de activación motora de los pacientes con EP, donde se aprecia un aumento del coeficiente de correlación medio (mano derecha  $p < 0,001$ ; mano izquierda  $p < 0,001$ ), mientras que no hubo cambios significativos en los pacientes sanos (mano derecha  $p = 0,1$ ; mano izquierda  $p = 0,2$ ).

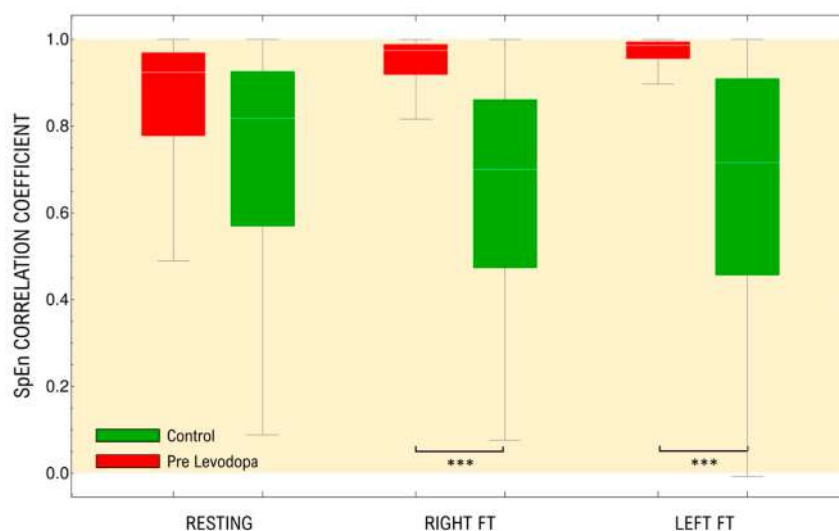


Figura 4.2: Distribuciones de los coeficientes de correlación por pares de la entropía espectral en el intervalo (1 – 30) Hz, para los controles (verde) y pacientes antes de la ingesta de su medicación (rojo) en los tres registros de electroencefalográficos. El área amarilla representa los valores con una correlación positiva. \*\*\* $p < 0,001$ . Fuente: [248].

### 4.1.2.2. Coherencia

En la Fig. 4.3 podemos observar que los canales del EEG tienen un fuerte acoplamiento en las bandas de baja frecuencia, especialmente en la banda delta (1 – 4 Hz). El aumento de la conectividad funcional cortico-cortical en estado de reposo en el rango alpha es una característica de la EP que aparece desde las primeras etapas clínicas.

Con la progresión de la enfermedad, las bandas de frecuencia vecinas se ven cada vez más implicadas, sugiriendo que los cambios en el acoplamiento funcional a lo largo de la evolución de la enfermedad pueden estar vinculados a la progresión topográfica de la patología sobre el cerebro [252].

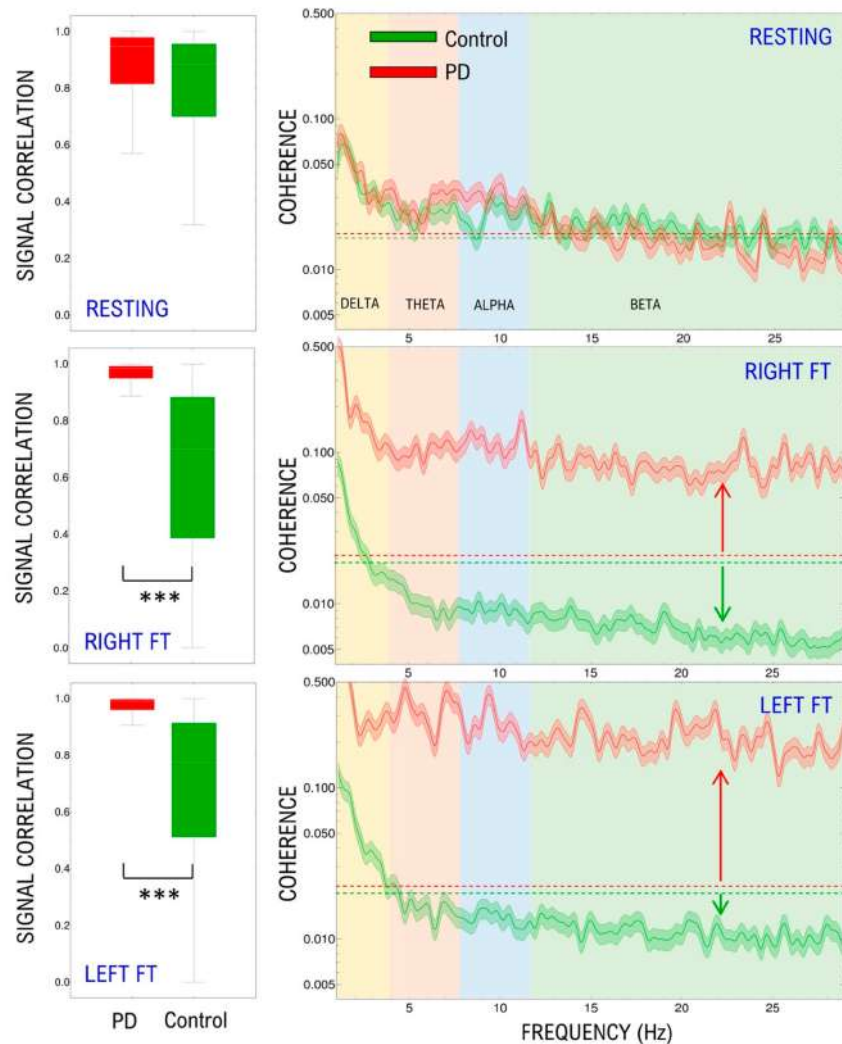


Figura 4.3: Coeficiente de correlación en el dominio temporal (paneles de la izquierda). Mediana y std de la coherencia espectral (paneles de la derecha). Los resultados se muestran para cada prueba registrada. Las líneas discontinuas en los paneles de la derecha indican la línea de base en estado de reposo, y las flechas reflejan el cambio producido por la activación motora.  $***p < 0,001$ . Fuente: [248].

En el estado de reposo no hay diferencias significativas entre controles y pacientes con EP ( $p = 0,15$ ), mientras que en las pruebas de activación motora, encontramos una disminución generalizada estadísticamente significativa, de la coherencia en los contro-

les, sin variaciones en ambas manos (mano derecha  $p = 0,01$ ; mano izquierda  $p = 0,03$ ), y un gran aumento generalizado estadísticamente significativo en el caso de los pacientes con EP, también sin diferencias entre manos (mano derecha  $p < 0,001$ ; mano izquierda  $p < 0,001$ ).

#### 4.1.2.3. Divergencia de Kullback-Leibler

La Divergencia de Kullback-Leibler (KLD) de la entropía espectral entre hemisferios cerebrales es una propiedad no simétrica. Por ello, hemos incluido tanto la divergencia directa como la inversa entre ambos hemisferios. Es decir, se ha calculado el KLD para cada canal de un hemisferio, con respecto a cada electrodo del otro hemisferio. Como se puede ver en la Fig. 4.4, los resultados de ambos hemisferios para los pacientes con EP y controles difieren en una cantidad basal de información en el estado de reposo. La divergencia parece ser mayor en los controles en comparación con los pacientes con EP, pero no es estadísticamente significativa ( $p = 0,15$ ). Los cambios de divergencia con la activación motora de la corteza cerebral fueron significativos tanto en los controles (mano derecha  $p = 0,01$ ; mano izquierda  $p = 0,015$ ) como en los pacientes con EP (mano derecha  $p = 0,02$ ; mano izquierda  $p = 0,002$ ). Llama la atención, la disminución de los valores de divergencia en los sujetos con EP mientras que los controles presentan un aumento de este valor. Al contrario que en el estado de reposo, cuando se compara la divergencia de ambos grupos durante las pruebas de activación motora, encontramos diferencias significativas (activación de la mano derecha en EP frente a los controles  $p < 0,001$ ; activación de la mano izquierda en EP frente a los controles  $p < 0,001$ ). No hay diferencias significativas al comparar las pruebas de FT para un mismo grupo (activación de la mano izquierda frente a derecha en controles  $p = 0,2$ ; activación de la mano izquierda frente a derecha en pacientes con EP  $p = 0,3$ ).

En el conjunto de pacientes que hemos considerado, se hace evidente la independencia parcial entre hemisferios cerebrales. Esta independencia, medida con la Divergencia de Kullback-Leibler, parece ser mayor en los controles en comparación con los pacientes con EP para la prueba de resting, pero no es estadísticamente significativa ( $p = 0,15$ ).

Sin embargo, existen diferencias significativas entre los cambios de independencia inter-hemisférica de ambos grupos para ambas activaciones motoras. Paralelamente a lo observado en la correlación de la coherencia y la entropía espectral, hay un aumento de la independencia en los controles y una disminución en los sujetos con EP, lo que indica que la conectividad entre ambos hemisferios puede haber aumentado durante las tareas motoras para los pacientes con EP. Esto requería una participación global de las redes biológicas y podría explicar un menor rendimiento de los pacientes enfermos en las actividades motoras. Este comportamiento fue general en ambas manos, sin mostrar



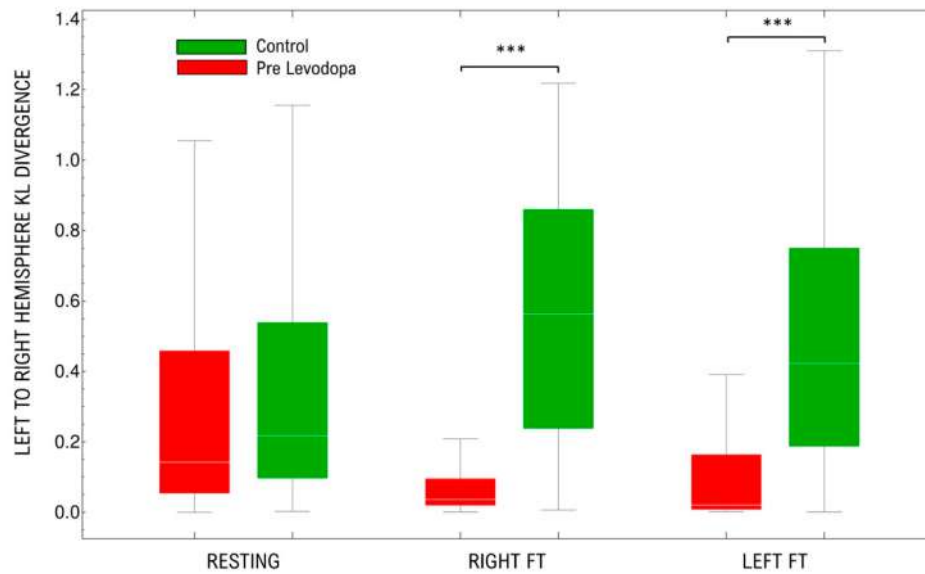


Figura 4.4: Distribución de las divergencias de Kullback-Leibler de la entropía espectral entre los electrodos de ambos hemisferios (izquierdo y derecho). Los resultados se muestran para las tres pruebas electroencefalográficas.  $***p < 0,001$ . Fuente: [248].

diferencias entre ellas. El aumento de la sincronía inter-hemisférica entre las cortezas sensoriomotoras se ha evidenciado previamente en modelos animales experimentales de EP [253] y también se ha asociado con disfunciones cognitivas en estudios de MEG [252].

En el conjunto de pacientes que hemos considerado, se hace evidente la independencia parcial entre hemisferios cerebrales. Esta independencia, medida con la Divergencia de Kullback-Leibler, parece ser mayor en los controles en comparación con los pacientes con EP para la prueba de resting, pero no es estadísticamente significativa ( $p = 0,15$ ).

Sin embargo, existen diferencias significativas entre los cambios de independencia inter-hemisférica de ambos grupos para ambas activaciones motoras. Paralelamente a lo observado en la correlación de la coherencia y la entropía espectral, hay un aumento de la independencia en los controles y una disminución en los sujetos con EP, lo que indica que la conectividad entre ambos hemisferios puede haber aumentado durante las tareas motoras para los pacientes con EP. Esto requería una participación global de las redes biológicas y podría explicar un menor rendimiento de los pacientes enfermos en las actividades motoras. Este comportamiento fue general en ambas manos, sin mostrar diferencias entre ellas. El aumento de la sincronía inter-hemisférica entre las cortezas sensoriomotoras se ha evidenciado previamente en modelos animales experimentales de EP [253] y también se ha asociado con disfunciones cognitivas en estudios de MEG [252].



#### 4.1.2.4. Efecto de la levodopa

También se analizaron los efectos de la levodopa en cada una de las medidas utilizadas, cuyos resultados pueden verse en la Figura 4.5.

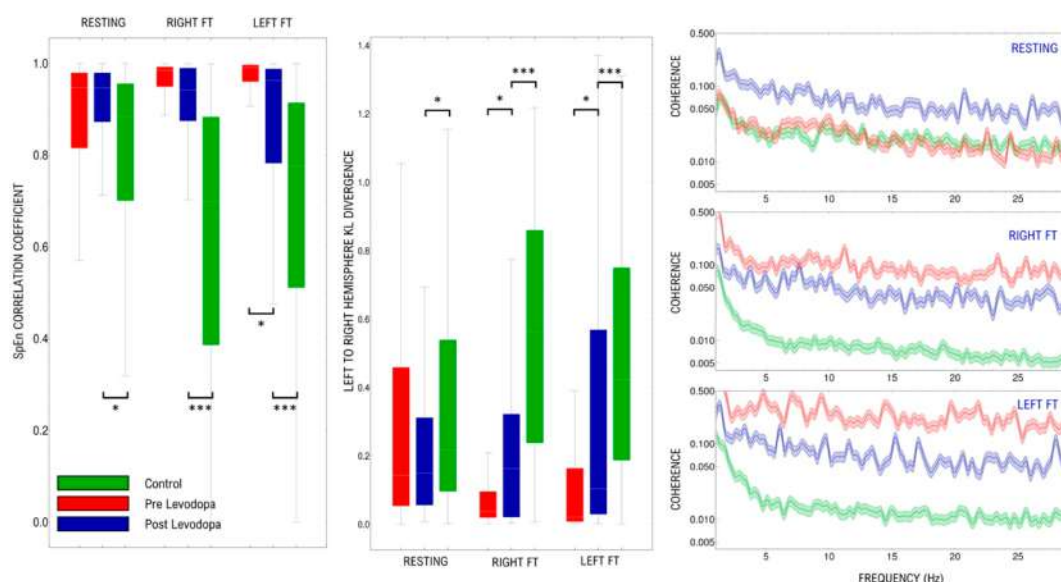


Figura 4.5: Comparación del efecto de la levodopa sobre: Coeficiente de Correlación de la Entropía Espectral (panel izquierdo), Divergencia de Kullback-Leibler (panel central) y Mediana y std de la coherencia de la señal (panel derecho), para el estado de reposo, las pruebas de FT. \* $p < 0,05$ , \*\* $p < 0,01$ , \*\*\* $p < 0,001$ . Fuente: [248].

El coeficiente de correlación de la entropía espectral es similar entre los sujetos con EP en estado de reposo antes y después de la administración de levodopa ( $p = 0,52$ ). Durante las pruebas de activación motora, la levodopa produce cambios significativos en el coeficiente de correlación en la activación de la mano izquierda ( $p = 0,016$ ) mientras que no produce cambios en el coeficiente de correlación en la mano derecha ( $p = 0,25$ ). Aunque en la mano izquierda la levodopa tiende a frenar el aumento del coeficiente de correlación observado en los sujetos con EP, sus valores siguen siendo significativamente diferentes a los de los controles ( $p < 0,001$ ).

Para la medida de conectividad de la coherencia, los valores obtenidos en el estado de reposo de los pacientes con EP después de la ingesta de la medicación, son significativamente más altos que los de los controles en todas las bandas ( $p = 0,02$ ). Durante la activación motora, la coherencia en ambas manos en los sujetos con EP tiende a permanecer similar al estado de reposo (mano derecha  $p = 0,18$ ; mano izquierda  $p = 0,25$ ), al contrario de los cambios observados en la activación motora pre levodopa (mano derecha  $p = 0,008$ ; mano izquierda  $p = 0,01$ ).

En cuanto a la divergencia, se puede observar que la administración de levodopa no produjo cambios en la divergencia inter-hemisférica en el estado de reposo de los pacientes con EP en comparación con el estado pre levodopa ( $p = 0,40$ ). Sin embargo, encontramos un incremento significativo en la divergencia inter-hemisférica durante la activación motora en los pacientes con EP, debido a la administración del fármaco (derecha  $p = 0,06$  e izquierda  $p = 0,015$ ) sin diferencias significativas cuando se comparan ambas manos ( $p = 0,41$ ).

## 4.2. Revisión bibliográfica de las técnicas de ML aplicadas a los EEGs de pacientes con EP

Como se ha comentado al presentar los objetivos específicos, se ha decidido incluir como resultados de esta tesis las revisiones bibliográficas llevadas a cabo relacionadas con el uso de técnicas de Machine Learning/Deep learning en el estudio de la EP. Dado que los trabajos originales de esta investigación encaminados al análisis de los EEGs como series temporales con técnicas propias de este campo no fueron todo lo concluyentes que se esperaba, fue necesario establecer una nueva hoja de ruta que explorara otras opciones. Estas opciones debían de presentar una mayor potencia en el tratamiento de señales complejas y ofrecer la suficiente novedad para hacer de esta tesis un trabajo original, publicable y que contribuyera al avance del estado del arte.

Estas revisiones permitirán verificar la *hipótesis 3* además de proporcionar un punto de partida a nuestra investigación, abordando aquellos campos que todavía no han sido explorados mediante el análisis de EEGs con técnicas de DL en la EP.

### 4.2.1. Revisión sistemática desde un punto de vista clínico

La revisión enfocada desde el punto de vista médico, ha sido publicada como “Machine Learning Approaches for Detecting Parkinson’s Disease from EEG Analysis: A Systematic Review” [234]. El diagrama PRISMA que se muestra en la Fig. 4.6, refleja los resultados obtenidos para cada uno de los pasos descritos en la metodología. Inicialmente el proceso de búsqueda en las bases de datos proporcionó 230 resultados (49 de Web of Science, 29 de PUBMED, 84 de Scopus, 25 de MEDLINE, 3 de CINAHL y 40 de Science Direct), 65 de los cuales eran duplicados, por lo que fueron eliminados como primer paso dentro del proceso de selección, dando lugar a 165 resultados. Los estudios no catalogados como artículos de investigación fueron rechazados (es decir, 36 actas y artículos de congresos, 17 capítulos de libros y 4 pósteres y editoriales) por no ser revisados por pares. En consecuencia, 57 estudios fueron eliminados en este paso,

quedando un total de 108 artículos, que fueron sometidos al proceso de elegibilidad, donde se aplicaron los criterios de inclusión y exclusión descritos en la metodología. Como resultado de esta fase, 9 artículos fueron excluidos por no utilizar técnicas de ML, 24 artículos no centraron su estudio en la EP, 27 artículos no utilizaron registros de EEG, 3 artículos consideraron estudios en animales, 1 artículo tuvo temática farmacológica, 24 artículos fueron revisiones con un propósito diferente, 3 artículos realizaron estudios sobre registros de EEG del sueño, 6 artículos se basaron en cambios de EEG provocados por estímulos exógenos y 2 artículos tenían descripciones incompletas sobre la metodología utilizada. La suma de todos ellos dio como resultado un total de 99 exclusiones, dejando un total de 9 artículos de investigación que fueron incluidos en esta revisión.

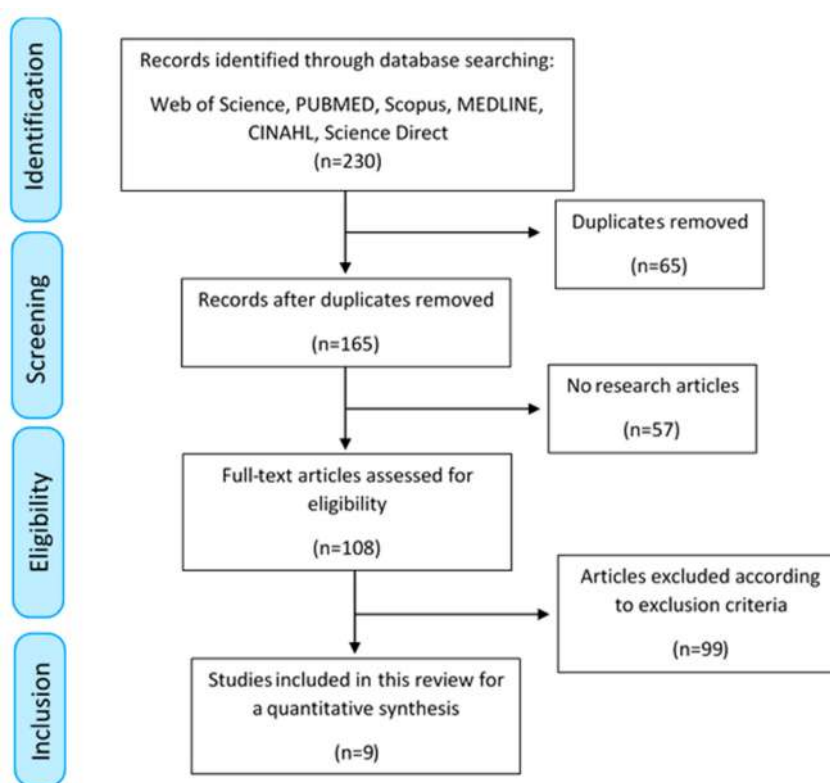


Figura 4.6: Diagrama PRISMA de la revisión bibliográfica realizada desde el punto de vista clínico. Fuente: [234].

#### 4.2.1.1. Análisis de la calidad de los artículos

Para evaluar la calidad de las publicaciones obtenidas en la revisión, se consideraron los ítems de la lista de verificación que se muestra en la Tabla D.1 del Apéndice D, para comparar el contenido de las publicaciones. El primer evaluador proporcionó un valor promedio de  $9,56 \pm 1,89$  sobre 12 para los 9 artículos, mientras que el segundo evaluador

determinó una evaluación promedio de  $8,89 \pm 1,97$  sobre 12. Para valorar la concordancia entre ambas evaluaciones, se calculó el valor kappa ( $\kappa$ ), que tiene en cuenta el efecto del azar sobre la concordancia observada, obteniendo un valor de  $\kappa = 0,67$ . Este resultado indica un nivel de acuerdo moderado-alto entre los evaluadores. Para facilitar el análisis acerca del cumplimiento de los ítems de la lista de verificación, la Fig. 4.7 muestra un gráfico que contiene el número de artículos que satisfacen cada uno de los ítems por cada uno de los evaluadores. Ambos evaluadores coinciden en que los ítems menos cumplidos fueron el 11 y el 12, que se relacionan con las limitaciones del modelo y los resultados inesperados, respectivamente. El hecho de que la mayoría de los artículos no incluyeran este tipo de análisis puede deberse a que en ocasiones se desconocen los alcances o limitaciones en el momento de la publicación y sólo se hacen evidentes con el paso de los años y con el desarrollo de nuevos algoritmos.

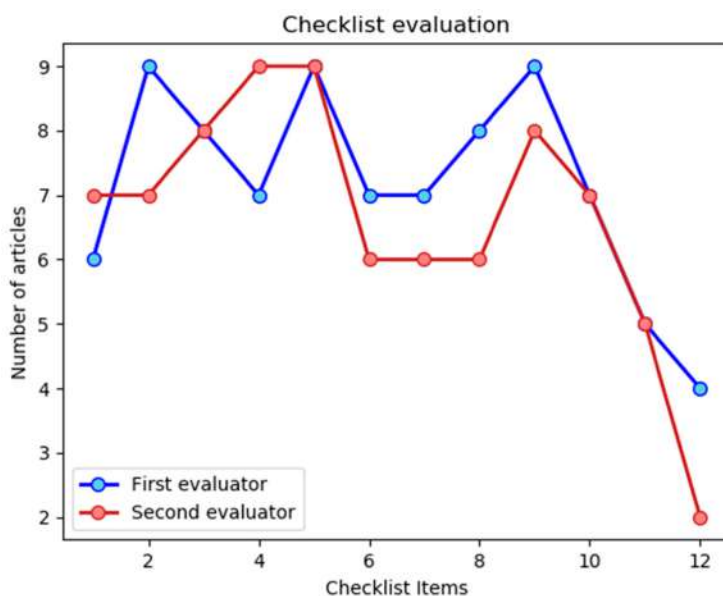


Figura 4.7: Número de artículos seleccionados que satisfacen cada uno de los ítems de la lista de verificación presentada en la Tabla D.1. Fuente: [234].

El contenido de los artículos se extrajo en función de los criterios expuestos, que abordan tanto un enfoque clínico y computacional. La información se recoge en las tablas D.2 y D.3 que encontramos en el Apéndice D, proporcionando un análisis cualitativo de los puntos recogidos en la lista de verificación mostrada en la Tabla D.1. Los aspectos que se consideraron incluyen:

- **Análisis de la calidad del conjunto de datos.** Mediante el estudio del número de pacientes reclutados, el tipo de registro electroencefalográfico realizado y sus parámetros.

- **Análisis del pre-procesamiento de dato.** Mediante el protocolo utilizado de limpieza del EEG y las características extraídas.
- **Características del modelo utilizado.** Especificando si utilizaron uno o más modelos, los parámetros de la arquitectura del modelo y los métodos de entrenamiento y validación llevados a cabo.

Se incluye una columna adicional con los resultados más relevantes para cada uno de los artículos seleccionados, facilitando el análisis de los pares modelo-parámetros más representativos para cada estudio.

La información resumida en las tablas [D.2](#) y [D.3](#) del Apéndice [D](#), facilita la comparación entre los diferentes artículos y entre las propiedades de cada estudio. En cuanto al objetivo de los artículos seleccionados, 8 de ellos se centraron en el problema de clasificación, es decir, buscaban diagnosticar la EP distinguiendo entre pacientes y controles. El artículo restante clasificó el grado de deterioro cognitivo de la EP.

El equilibrio entre el número de pacientes con EP y controles es crucial cuando se utilizan técnicas de ML, ya que los datos desbalanceados pueden generar errores en la predicción. Se puede comprobar que esta era una práctica común en los artículos revisados, ya que 7 de los 8 artículos que clasifican la EP consideraban un conjunto de datos balanceado. En cuanto al número de pacientes incluidos en los estudios, predominaron los estudios con menos de 50 pacientes en cada categoría, con un valor medio de  $28,20 \pm 11,53$  para el grupo de pacientes con EP y un valor medio de  $27,20 \pm 7,83$  para el grupo de los controles. Los artículos no indicaron si el número de pacientes era adecuado para el problema de clasificación. Además, aunque en todos los artículos no se especificaba el valor medio de la edad de ambos grupos, todos ellos consideraron pacientes con EP de edades comprendidas entre los 45-70 años, cuyo valor medio oscilaba en torno a los 60 años, correspondiendo con la edad de inicio de la enfermedad. En cuanto a los controles, se eligieron de forma que presentaran las mismas características demográficas que el grupo de pacientes con EP. Cabe señalar que sólo 4 de los artículos seleccionados indicaron si los pacientes habían tomado su dosis de levodopa (3 estudios realizaron el EEG en estado ON y 1 en estado OFF).

En cuanto al grado de progresión de la enfermedad, encontramos en la Tabla [D.2](#) una ausencia generalizada de información. Específicamente, 6 artículos indican el estado de los pacientes según la escala HY, 4 de ellos considerados en los estadios HY:1-3, y 2 de ellos consideraron pacientes en estadios tempranos de la enfermedad (HY 1 y 1,5). Ninguno de los artículos resultantes incluyó pacientes en fases más avanzadas, esto puede ser una limitación para evaluar la capacidad de extrapolar los resultados o evaluar la progresión de la enfermedad. De igual forma, sólo 3 artículos mostraron el estado de

los pacientes según la UPDRS, con un valor promedio de  $34,43 \pm 6,43$ . La duración de la enfermedad se especifica en 4 artículos, con un valor medio de  $6,38 \pm 1,35$  años.

Con respecto a los parámetros de registro del EEG, se puede apreciar que el número de canales varió entre los diferentes estudios. En 2 artículos se realizaron los registros con un EEG de alta densidad de electrodos (esto es, más de 100 canales), mientras que en 5 artículos el registro se llevó a cabo con un EEG de baja densidad de electrodos (menor o igual a 20) con un valor promedio de  $16,2 \pm 2,72$  electrodos. Los 2 artículos restantes utilizaron grabaciones de EEG con sólo 2 canales, pero consideraron una técnica que combinaba características de EEG y EMG (electromiografía). Cabe señalar que estos artículos estaban relacionados con el mismo estudio y realizado por el mismo grupo de investigación. El tiempo de registro del EEG también fue variable entre los artículos seleccionados, mostrando nuevamente valores heterogéneos, aunque el registro de 5 min de duración fue el más frecuente apareciendo en 4 de los artículos incluidos.

Con respecto al pre-procesamiento, es posible distinguir entre el protocolo de limpieza de los EEGs (que se muestra en la Tabla D.2) y la extracción de características del conjunto de datos (que se muestra en la Tabla D.3). Por un lado, el pre-procesamiento o limpieza del EEG, varió de un artículo a otro, principalmente debido a la falta de un protocolo estándar de limpieza de artefactos, dificultando la evaluación de la calidad del dataset. En particular, 3 de los artículos realizaron el pre-procesamiento del EEG mediante la eliminación de artefactos de la señal, 3 artículos minimizaron el ruido de las señales de EEG por medio de filtros y los 3 artículos restantes no especificaron el proceso de limpieza, lo que nos conduce a pensar que no llevaron a cabo alteraciones en las señales del EEG. Por otro lado, asociado con la extracción de características del dataset, destaca la gran heterogeneidad entre los artículos. Sin embargo, todas ellas fueron extraídas del espectro de frecuencias, y sólo hubo un artículo en el que no se realizó ningún pre-procesamiento de datos.

#### 4.2.1.2. Tipos de modelos considerados

Centrándonos en el análisis de los modelos empleados, una de las características más destacables en los artículos seleccionados tiene que ver con la variedad de modelos utilizados. De los 9 artículos que componen esta revisión, se utilizaron un total de 11 técnicas de ML diferentes para llevar a cabo los objetivos propuestos, siendo el principal objetivo el diagnóstico de la EP. La Figura 4.8 muestra un gráfico circular con los diferentes modelos y el número de veces que fueron considerados en los artículos. Los modelos incluidos son: SVM (por sus siglas en inglés, "Support Vector Machine"), KNN (por sus siglas en inglés, "K-Nearest Neighbors"), DT (por sus siglas en inglés, "Decision Tree"), CNN, MLP (por sus siglas en inglés, "Multilayer Perceptron"), RF (por sus siglas

en inglés, “Random Forest”), RNN (por sus siglas en inglés, “Recurrent Neural Network”), DFA (por sus siglas en inglés, “Discriminant Function Analysis”), FKNN (por sus siglas en inglés, “Fuzzy K-Nearest Neighbors”), NB (por sus siglas en inglés, “Naïve Bayes”) y PNN (por sus siglas en inglés, “Probabilistic Neural Network”).

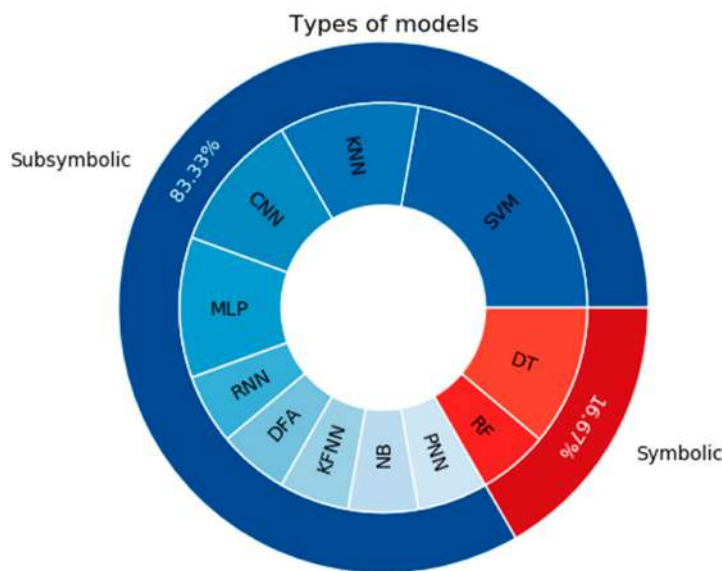


Figura 4.8: Diagrama circular que indica los tipos de modelos usados. En rojo se muestran aquellos con procesamiento simbólico y en azul los modelos con procesamiento subsimbólico. Fuente: [234].

En la Fig. 4.8 se puede observar que el número de técnicas empleadas excede al número total de artículos. Esto se debe a que 4 de los artículos seleccionados compararon los resultados ofrecidos por diferentes modelos, mientras que 5 de ellos consideraron un único modelo. En cuanto al tipo de procesamiento de los modelos, podemos apreciar que tanto DT como RF pertenecen al grupo de modelos simbólicos, predominando aquellos modelos que pertenecen al grupo de procesamiento subsimbólico.

Como técnica individual, SVM fue la más utilizada, además de proporcionar los mejores resultados para la clasificación de los pacientes con EP frente a los controles. Por otro lado, cabe destacar que las ANN también jugaron un papel importante en los artículos revisados ya que estas técnicas fueron utilizadas 6 veces a través de los modelos CNN, MLP, RNN y PNN. Se debe notar que uno de los artículos consideraba dos modelos diferentes asociados a RNN, con capas LSTM y GRU respectivamente, como se muestra en la Tabla D.3, aunque esto no se ha tenido en cuenta ni en la Fig. 4.8, ni en el cálculo anterior.

Teniendo en cuenta que los modelos más utilizados fueron SVM y KNN y la métrica más utilizada fue la exactitud o accuracy, se realizó un estudio comparativo entre ambos



modelos mediante un Forest-plot. Para realizar este análisis se necesita que varios estudios utilicen estos dos modelos, y sólo dos artículos cumplieron esta condición: [254] y [255]. La Figura 4.9 refleja el meta-análisis con los resultados del accuracy de ambos modelos, mostrando las medidas de efecto asociadas a la Diferencia de Medias Estandarizada (SMD por sus siglas en inglés, "Standarized Mean Difference"), calculadas usando modelos de efecto fijo y la  $g$  de Hedges ponderada según la inversa de la varianza.

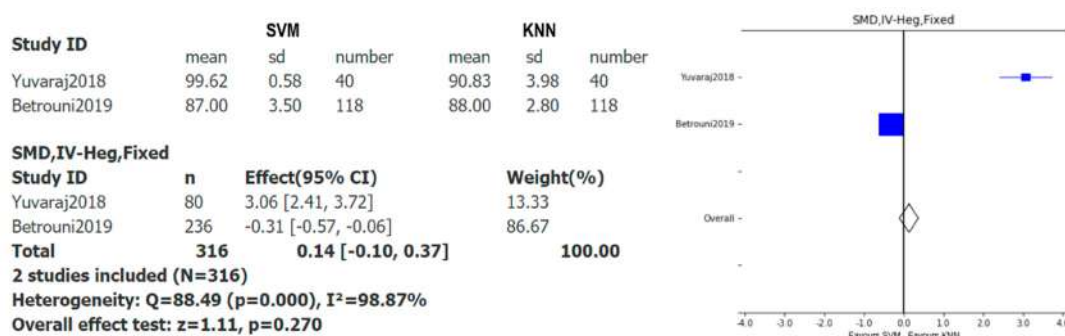


Figura 4.9: Forest-plot con la Diferencia de Medias Estandarizada de la exactitud o accuracy en ambos modelos, SVM and KNN, calculada usando la  $g$  de Hedges ponderada según la inversa de la varianza. Fuente: [234].

La Fig. 4.9 muestra la diferencia (estandarizada) de los valores medios de los datos de SVM y KNN, por lo que favorece los resultados con valores más bajos de accuracy. En el caso del primer artículo el valor de accuracy se obtiene para KNN y en el segundo artículo para SVM. Dado que la diferencia entre los resultados de ambos modelos es mayor en el primer artículo, el intervalo de confianza (IC) se encuentra más lejos de cero, mientras que en el caso del segundo artículo, como los valores medios de los resultados son cercanos, el IC es cercano a cero. Entre los resultados mostrados en la Figura 4.9 resalta la diferencia entre el número de sujetos considerados en cada estudio y cómo este valor influye en el meta-análisis. Llama especialmente la atención cómo el tamaño de la muestra influye tanto en el IC como en el peso de cada uno de los casos. De hecho, cuanto mayor es el número de pacientes, menor es la amplitud del IC y mayor el peso asociado. Dado que ninguno de los intervalos de confianza cruza la "línea de no efecto", la diferencia entre los resultados de SVM y KNN es estadísticamente significativa en ambos estudios.

Como resultado general, el meta-análisis muestra que no hay un beneficio estadísticamente significativo entre elegir un modelo sobre el otro, ya que el rombo mostrado en la Fig 4.9 cruza la "línea de no efecto". Hay que tener en cuenta que únicamente se comparan dos estudios con una gran heterogeneidad en los datos, de modo que los resultados presentados en el meta-análisis son menos concluyentes. Consecuentemente,



se necesitan más estudios que consideren ambos modelos de ML simultáneamente para proporcionar una conclusión objetiva más confiable. Finalmente, vale la pena enfatizar que, aunque las técnicas de ML están influenciadas por la cantidad de datos introducidos en el modelo, esto no implica que los modelos con más datos siempre den mejores resultados, pero es crucial que el conjunto de entrenamiento sea lo suficientemente grande para llevar a cabo el estudio.

#### 4.2.1.3. Tipo de registro del EEG

En cuanto al tipo de registro electroencefalográfico, se puede ver en la Tabla D.2 del Apéndice D que los artículos seleccionados presentaron 2 tipos de registro. Por tanto, hemos realizado una división en dos categorías teniendo en cuenta las pruebas realizadas. Por un lado, el grupo de “EEGs en estado de reposo”, que se corresponde con los artículos [255, 256, 257, 258] y [259, 260, 254]. Por otro lado, el grupo de “EEGs de activación motora”, que se corresponde con los artículos [261, 262]. Para cada grupo se realizan pruebas de Resting y pruebas de extensión y flexión de muñeca, respectivamente. Para cada uno de estos artículos, el modelo considerado, los resultados de clasificación obtenidos, las características introducidas y el tipo de limpieza EEG realizada se resumen en la Tabla 4.2.

El grupo de los EEGs en estado de reposo contiene 7 artículos, que consideraron diferentes protocolos de registro y número de canales en la grabación de los EEGs. Los artículos [257] y [254] se basaron en el mismo estudio, pero utilizaron diferentes características y modelos. El protocolo que predominó dentro de este grupo consistió en registrar el EEG en estado de reposo con los ojos cerrados, utilizado en 4 artículos. Tampoco hubo un criterio estándar para la duración del registro, cuyo valor promedio fue de  $6,37 \pm 3,10$  min, con una moda de 5 min, considerada en 4 de los 7 artículos que forman este grupo. El número de canales de EEG también fue heterogéneo, prevaleciendo una baja densidad de electrodos, considerada en el 71,43% de los estudios con registros de 14 y 20 electrodos. Los restantes artículos consideraron EEGs con una alta densidad de electrodos (esto es, más de 100 canales).

El pre-procesamiento de datos también puede dividirse en dos categorías. Por un lado, el pre-procesamiento o limpieza del EEG, y por otro lado, el pre-procesamiento de datos o extracción de características. En cuanto a la limpieza del EEG, tal y como se muestra en la Tabla 4.2, se puede concluir que no hubo un protocolo de limpieza estándar, ya que 3 de los estudios incluidos eliminaron los artefactos de las señales electroencefalográficas, 3 realizaron un pre-procesamiento basado en filtros para reducir el ruido en las señales (sin la eliminación de artefactos), y el protocolo empleado no se especificó en [260]. En cuanto a la extracción de características, sólo en [257] se

Ref	Año	Resultados de accuracy	Características	Artefactos
[255]	2019	84 % - SVM; 88 % - KNN	Potencia relativa con baja densidad de electrodos	Libre de artefactos
[256]	2017	78 % - RF	Selecciona las características más importantes	Libre de artefactos
[257]	2018	88,25 % - CNN	No considera	Filtro de ruido
[258]	2019	79 % - CNN; 81 % - RNN	Espectrogramas	Filtro de ruido
[261]	2019	EEG - 62 %, EMG - 73 %, y ambas - 98,8 %	EEG: parámetros no-lineales. EMG: parámetros estadísticos	No específica
[262]	2019	ANN 100 % con el algoritmo Quasi-Newton (Trainlm)	EEG: parámetros no-lineales. EMG: parámetros estadísticos	No específica
[259]	2018	94,34 % - SVM	Espectro de potencia de las bandas y regiones	Libre de artefactos
[260]	2020	95,24 % - DFA	Coherencia en la banda beta	No específica
[254]	2018	99,62 % - SVM	Características del espectro de orden superior (HOS)	Filtro de ruido

Tabla 4.2: Resumen de los resultados, características introducidas a los modelos y filtrado de las señales basados en las Tablas D.2 y D.3 del Apéndice D de los artículos seleccionados. Se incluye el año de publicación de cada estudio. Fuente: [234].

introdujo como entrada al modelo la señal del EEG para realizar un análisis morfológico, en los artículos restantes se calcularon diferentes características espectrales.

Los modelos que proporcionaron mejores resultados de clasificación (accuracy superior al 90 %) para pacientes con EP frente a controles fueron SVM y DFA, considerados en los artículos [259, 260, 254]. Hay que tener en cuenta que [260] registró el EEG de pacientes con EP sin ingesta de levodopa. Estos artículos introdujeron diferentes características del espectro de frecuencia como entrada al modelo y realizaron diferentes protocolos de limpieza de las señales del EEG. Debemos notar que, aunque [255] también utilizó SVM, su objetivo no fue la clasificación de pacientes con EP frente a controles, sino el estudio de la progresión de la enfermedad. Por lo tanto, los resultados obtenidos para ese caso no son comparables con los resultados obtenidos en los restantes artículos.

Para el grupo de los EEGs en estado de reposo, la evaluación obtenida de los criterios de calidad según la lista de verificación de la Tabla 4.2 proporcionó un valor promedio de  $10,43 \pm 1,05$  sobre 12 según el primer evaluador y un valor de  $9,71 \pm 1,28$  sobre 12 según el segundo evaluador. El valor kappa calculado para los ítems de este grupo fue de 0,71, indicando un nivel de acuerdo alto entre los evaluadores. Además, se puede apreciar que los valores obtenidos en las puntuaciones son mayores que considerando

el conjunto completo de artículos que constituye esta revisión.

El grupo de los EEGs de activación motora lo componen únicamente 2 artículos basados en el mismo estudio, [261, 262]. Por tanto, en ambos casos se registraron 2 canales de EEG y EMG durante 30 min en una prueba de activación motora en la que se extendía y flexionaba la muñeca. Se calcularon los mismos parámetros no lineales, aunque las características introducidas en el modelo cambiaron en cada artículo. No se especificó el pre-procesamiento de los EEGs. En ambos casos, se utilizaron ANNs pero para diferentes propósitos. En [261] se realizaron 3 estudios para seleccionar los parámetros de entrada al modelo que proporcionaban mejores resultados, mientras que en [262] se estudiaron 6 técnicas diferentes para los mismos parámetros de entrada con el objetivo de seleccionar el modelo con mejor rendimiento. En este caso, las características de entrada fueron una combinación de EEG y EMG, coincidiendo con las características que proporcionaron mejores resultados en [261].

Finalmente, de acuerdo con la lista de verificación mostrada en [235], estos artículos recibieron una evaluación de  $6,5 \pm 0,5$  sobre 12 por parte del primer evaluador y  $6 \pm 1,0$  sobre 12 por parte del segundo evaluador. El valor kappa para este grupo fue de 0,58. En este caso, se puede apreciar que el valor  $\kappa$  resultante es ligeramente inferior a los obtenidos para el grupo de los EEGs en reposo y para el conjunto global de artículos, lo que indica una menor concordancia entre los evaluadores respecto a los casos anteriores.

#### 4.2.2. Revisión sistemática desde el punto de vista de las ciencias computacionales

La revisión sistemática desde el punto de vista de las ciencias computacionales ha sido publicada y puede verse en [236]. Como se muestra en el diagrama de flujo PRISMA en la Fig. 4.10, el proceso de búsqueda en las bases de datos proporcionó un total de 358 resultados, de los cuales 132 fueron duplicados y, por lo tanto, fueron eliminados. Los 226 resultados restantes se sometieron al proceso de selección, donde se rechazaron 25 libros académicos, 1 resumen y 1 póster. Como consecuencia, 27 estudios fueron eliminados, quedando un total de 199 artículos para la fase de elegibilidad, en la que se aplicaron los criterios de exclusión descritos en la Tabla 3.3. Como resultado de esta fase, se encontró que 24 artículos no utilizaron técnicas de ML, 27 artículos no enfocaron su estudio en la EP, 29 artículos no utilizaron técnicas de electroencefalografía, 2 artículos consideraron estudios en animales, 7 artículos fueron farmacológicos, 43 artículos fueron revisiones con una finalidad diferente, y 6 estudios consideraron EEGs invasivos. Como se detalla en el diagrama PRISMA de la Fig. 4.10, la suma de todas las exclusiones dio como resultado un total de 138 artículos rechazados, quedando 61 estudios que fueron considerados para la fase de inclusión. Finalmente, uno de los estudios considerados fue

retractado y otro no se pudo encontrar, por lo que ambos fueron excluidos del resultado final. Por lo tanto, el proceso dio lugar a 59 artículos que fueron los considerados para realizar el estudio.

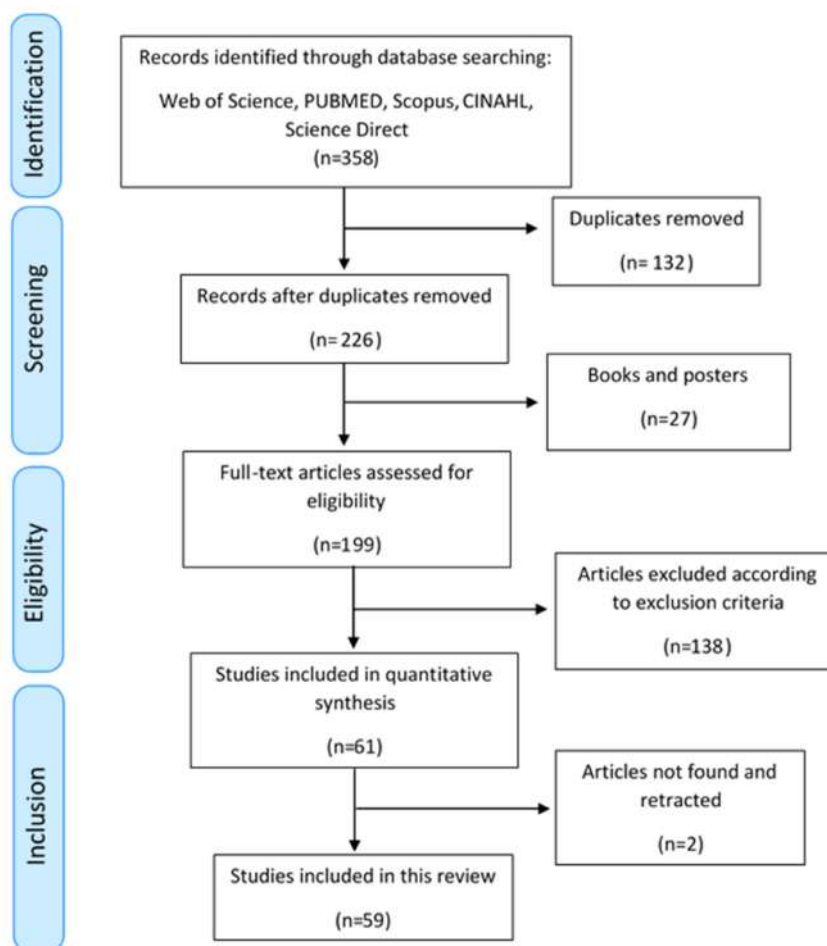


Figura 4.10: Diagrama PRISMA de la revisión bibliográfica desde el punto de vista de las ciencias computacionales. Fuente: [236].

#### 4.2.2.1. Análisis estadístico

Las técnicas de ML han tenido un creciente desarrollo en los últimos años, aumentando su aplicación en diferentes áreas. Este crecimiento, particularizado en la aplicación de técnicas de ML sobre los EEGs asociados a la EP, fue captado por los artículos seleccionados en esta revisión, como se muestra en la imagen A de la Figura 4.11, que muestra un aumento del número de publicaciones en los últimos 5 años, y por tanto un creciente interés en este tema, mientras que en los años anteriores, el desarrollo fue intermitente. Se puede apreciar una disminución en la producción del número de estudios en el año

2022 respecto al año anterior, aunque hay que tener en cuenta que el año 2022 no se incluyó en su totalidad en la revisión, limitado por la fecha de búsqueda (14 de febrero) y, por lo tanto, los resultados de este año no son comparables con los resultados de años anteriores. Se debe notar que, en la imagen A de la Fig. 4.11, se ha considerado la última fecha de publicación.

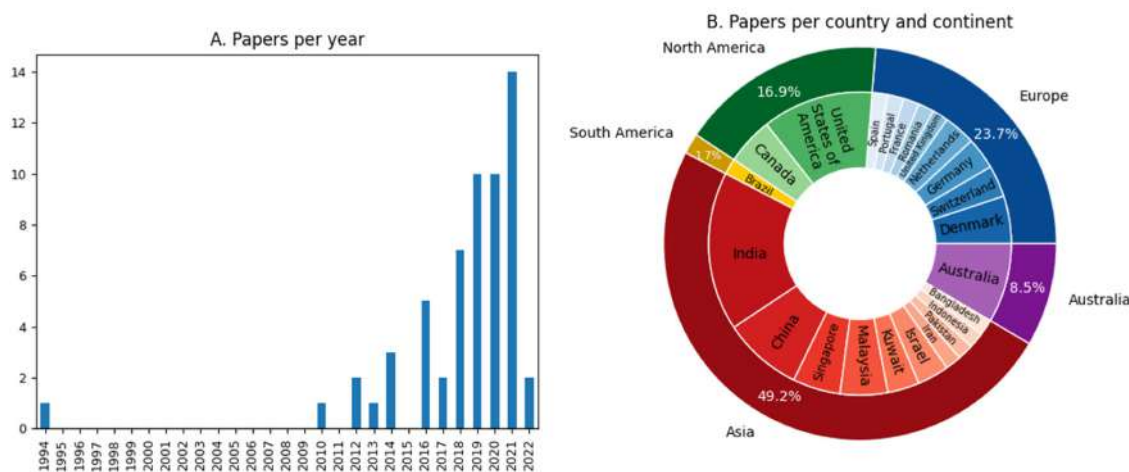


Figura 4.11: A: Diagrama de barras acerca del número de artículos publicado por año teniendo en cuenta la última fecha de publicación. B: Gráfico circular con la distribución de los artículos seleccionados según el país y continente asociado a la primera filiación del primer autor. Fuente: [236].

La imagen B de la Figura 4.11 muestra la distribución de los artículos según el primer país de filiación del primer autor, así como el continente al que pertenece dicho país. Como podemos observar, el continente que más contribuyó al desarrollo de nuevas investigaciones de estudios de la EP mediante EEGs con técnicas de ML fue Asia (49,2%), seguido de Europa (23,7%). Aunque las contribuciones de Norte América (16,9%) y Australia (8,5%) fueron menores en proporción, vale la pena señalar el interés científico mundial por este tema en los últimos años. En cuanto al país, se puede apreciar que tanto EE.UU. como India aportaron el mayor número de publicaciones dentro de sus respectivos continentes en proporciones similares. Tanto China como Australia tuvieron una contribución similar, mientras que en Europa la distribución fue homogénea.

Para analizar el contenido de los artículos seleccionados en esta revisión y extraer la información más relevante, fue necesario fijar unas pautas comunes a todos los estudios. En la Tabla E.1 del Apéndice E se resumió la información extraída de los artículos según tres puntos clave elegidos por su importancia dentro del campo de ML. Los puntos considerados fueron:

- **Evaluación de la calidad del dataset.** Se evaluó si los artículos consideraban clases balanceadas, si las muestras eran estadísticamente equivalentes por medio del número de sujetos en cada grupo y los datos demográficos de los pacientes. También se evaluaron otros parámetros clínicos relacionados con la progresión de la EP, como la escala de Hoehn-Yahr, la UPDRS, los años de enfermedad y si el registro se llevó a cabo bajo los efectos de la medicación dopaminérgica. Estos parámetros proporcionaron información sobre la calidad del conjunto de datos, favoreciendo la evaluación del rendimiento del modelo. Finalmente, se valoraron los parámetros relacionados con el registro de las señales electroencefalográficas, como el número de canales, la duración de la prueba y el tipo de prueba realizada, proporcionando información acerca de la calidad del proceso de registro de las señales de EEG. Estos puntos coinciden con las columnas “Participants”, “Stage PD” y parte de la columna “EEG Pre-Processing” de la Tabla E.1 del Apéndice E.
- **Pre-procesamiento de datos.** En este punto se analizó el protocolo de limpieza de los EEGs y la extracción de características. El protocolo de limpieza del EEG es un proceso que a veces no se especifica o no se realiza debido a la falta de un estándar, ya que la gran variedad de técnicas que se pueden considerar producen diferentes modificaciones en las señales del EEG. Por lo tanto, se evaluó la limpieza del EEG en cada artículo para comprobar el impacto de este pre-procesamiento en el rendimiento de los modelos utilizados. Con respecto a la extracción de características de las señales electroencefalográficas, se debe tener en cuenta que las características introducidas en los modelos juegan un papel trascendental en las técnicas de ML y en los resultados obtenidos, de modo que fueron recopiladas para determinar aquellas utilizadas con mayor frecuencia y aquellas para las que se obtuvieron mejores resultados. La información asociada a este punto se especificó en parte de la columna de “EEG Pre-Processing” y la columna “Features” de la Tabla E.1 del Apéndice E.
- **Evaluación de los modelos utilizados.** Se examinó el tipo de modelo utilizado junto con su arquitectura y sus métodos de entrenamiento y validación. Este conjunto de parámetros se consideró para evaluar qué modelos obtuvieron mejores resultados según el objetivo del artículo. Además, el análisis del proceso de validación llevado a cabo por cada artículo permitió evaluar la calidad de los resultados y proporcionó una valoración más objetiva del alcance de los resultados predictivos del modelo. Dentro de la Tabla E.1 del Apéndice E, esta información corresponde a las columnas “Models”, “Model Parameters”, y “Validation”.

Los aspectos expuestos anteriormente combinan los puntos de vista clínico y computacional, brindando en conjunto, un análisis de los tres pasos fundamentales que deben considerarse en un problema de ML. Para facilitar el análisis posterior, la Tabla E.1 incluye una fila en el encabezado, indicando la correspondencia de cada uno de estos puntos con las columnas introducidas. Se añadió una columna adicional indicando el objetivo de cada estudio, y otra columna con los resultados más relevantes de cada modelo junto con las métricas utilizadas.

Se proporciona un análisis global para cada uno de los tres puntos clave introducidos: calidad del conjunto de datos, preprocesamiento de datos y evaluación del modelo. En cuanto al objetivo de los artículos seleccionados, se pudo observar que 30 de los estudios incluidos en esta revisión cubrían el problema de la clasificación entre pacientes con EP y controles, mientras que los artículos restantes consideraban temas diversos: 6 detectaban alteraciones en la marcha, 4 clasificaban el deterioro cognitivo, 2 seleccionaban las características para clasificar el deterioro cognitivo y pacientes con EP, y 1 distinguía entre pacientes con medicación frente a pacientes sin medicación. Los 16 artículos restantes no se incluyeron en la Tabla E.1 del Apéndice E a pesar de superar los criterios de exclusión, ya que su objetivo no estaba completamente relacionado con el estudio de la EP con EEG. Dentro de este grupo de artículos, 5 clasificaron las emociones [263, 264, 265, 266, 267], 4 identificaron las alteraciones del sueño [268, 269, 270, 271], 5 utilizaron la ECP o la neuroestimulación [272, 273, 274, 275, 276], uno combinó las características del EEG y del EMG [277] y uno clasificó las tareas mentales [278]. En conclusión, se puede apreciar que el problema de diagnóstico de la EP es el objetivo principal de este tipo de estudios. La evaluación posterior se realizó sobre los 43 artículos incluidos en la Tabla E.1.

#### 4.2.2.2. Evaluación del conjunto de datos

La evaluación de este punto contiene el análisis del número de sujetos que contenía cada clase, los parámetros clínicos asociados con la progresión de la enfermedad y los parámetros de registro del EEG.

De la Tabla E.1 del Apéndice E se obtuvo que el 69,77 % de los estudios seleccionados (30 artículos) utilizaron clases balanceadas, es decir, las clases contenían un número similar de sujetos. Un conjunto de entrenamiento con clases desequilibradas puede generar errores de predicción y conducir a una mala generalización de los resultados del modelo. El criterio escogido para evaluar el equilibrio entre los sujetos de cada clase fue que la diferencia entre cada par de clases fuera mayor al 25 % de la clase mayoritaria. Según este criterio, el 16,28 % de los estudios no consideraron clases balanceadas, y los artículos [279, 280, 281, 282, 283, 284] sólo consideraron una única clase. Cabe



señalar que en aquellos estudios que utilizaron una mezcla de pacientes con diferentes enfermedades, se tuvo en cuenta el número total de sujetos por clase. El número medio de sujetos en los estudios que consideraron clases balanceadas fue de  $23,10 \pm 16,74$  sujetos en cada clase (este resultado se calculó para los pacientes con EP). Se debe notar que, dentro de este conjunto de datos, sólo [262] supera los 50 pacientes en una de las clases, mientras que la moda fue de 20 sujetos por clase (considerado en 9 artículos). Para comprobar que ambos grupos podían ser comparados estadísticamente, era necesario comprobar que los sujetos exhibían las mismas características demográficas. De los 30 artículos que consideraron clases equilibradas, sólo 18 especificaron la edad media de los pacientes incluidos en el estudio, cuya media fue de  $63,37 \pm 4,22$  años, que coincide aproximadamente, con la edad de inicio habitual de la EP. De los restantes artículos pertenecientes a este grupo, 8 indicaron el rango de edad de los sujetos, contenido en el intervalo (40-80) años, y los estudios [285, 286, 287, 288, 289, 290] no especificaron la edad de los participantes.

Conocer el estado de la enfermedad en el que se encuentra el paciente es necesario para evaluar la calidad del dataset. Estos parámetros pueden influir en el rendimiento del modelo y, por lo tanto, afectar al problema de clasificación, ya que los pacientes con EP en estadios más avanzados de la enfermedad que fueron monitoreados sin su toma habitual de medicación dopaminérgica, pueden ser más fáciles de identificar que un grupo de pacientes con EP en las primeras etapas de la enfermedad que tomaron su medicación. Sin embargo, no todos los artículos contenían la información asociada al estado de la enfermedad de los pacientes. En la Tabla E.1 del Apéndice E se puede encontrar que el 32,56 % de los artículos seleccionados, no aportaron información sobre el estado de la enfermedad o la medicación, el 13,95 % de los estudios sólo indicaron la medicación, el 6,98 % sólo especificaron la afectación a través de la escala HY y el 11,63 % de los artículos proporcionaron información parcial. Consecuentemente, sólo el 34,88 % de los artículos seleccionados contenían toda la información relacionada con el estado clínico de los pacientes y su medicación. En cuanto a la medicación dopaminérgica, se observó que el 51,16 % de los artículos no indicaron el estado de la medicación durante el registro del EEG, el 13,95 % registró el EEG en estado ON, el 16,28 % registró el EEG en estado OFF y el 18,60 % registró el EEG en los estados ON y OFF.

Con relación al proceso de grabación del EEG, los parámetros especificados en la Tabla E.1 fueron: el número de canales del EEG, la duración del registro del EEG y la prueba realizada. El número de electrodos se especificó en todos los artículos con un valor medio de  $43,34 \pm 62,18$  electrodos. Como se puede apreciar, hubo una gran heterogeneidad en la densidad de electrodos, considerando una moda de 14 canales en 8 de los estudios incluidos. Llama la atención que los estudios [279, 262, 261, 283, 284, 291]



utilizaron EEGs con menos de 5 canales. Sin embargo, se debe notar que 3 de ellos detectaron alteraciones de la marcha, 2 diagnosticaron la EP en combinación con características de EMG, y [291] utilizó 1 canal con estímulos emocionales. En cuanto a la duración del registro del EEG, la moda fue de 5 min. En cuanto a la metodología empleada, la tendencia era dividir las señales en ventanas de longitud uniforme considerando cada segmento de forma independiente. En consecuencia, los registros más grandes proporcionaron datasets con mayor número de muestras, permitiendo a algunos de los estudios, eliminar aquellos segmentos del EEG que tenían defectos o podían perturbar el análisis. Finalmente, se consideraron diferentes tipos de pruebas electroencefalográficas, como pruebas de estado de reposo, estimulación por medio de emociones o pruebas de activación motora, que registran diferentes aspectos del cerebro del paciente y, por lo tanto, no son comparables. En los artículos seleccionados para esta revisión predominaron las pruebas en estado de reposo, consideradas en 22 artículos, seguidas de las tareas “Timed Up and Go” (presentes en 5 estudios), y las pruebas que requieren una respuesta física a un estímulo (“Oddball” y “Visual Go/No-Go” consideradas en 4 estudios).

#### 4.2.2.3. Pre-procesamiento de datos

Dentro del pre-procesamiento de datos, se pueden distinguir dos puntos clave: el protocolo de limpieza del EEG y la extracción de características. De la información de la Tabla E.1 del Apéndice E, se puede observar que el proceso de limpieza fue muy heterogéneo entre los artículos seleccionados, que puede ser consecuencia de la falta de un protocolo estándar de limpieza de EEGs. De hecho, el 27,91 % de los estudios no especificaba el proceso de limpieza llevado a cabo, el 23,26 % realizaba un pequeño pre-procesamiento en los datos del EEG y el 48,84 % de los artículos utilizaba señales de EEG libres de artefactos. Este aspecto dificulta la evaluación del dataset y la valoración de cómo este pre-procesamiento afecta a las señales del EEG, ya que las alteraciones en las señales pueden modificar aspectos esenciales y conducir a un diagnóstico falso. En el Capítulo 5 de discusión se muestra una evaluación más precisa de estos aspectos.

Los artículos también mostraron una gran heterogeneidad en las características extraídas de las señales de EEG, predominando las características espectrales (consideradas en el 88,37 % de los artículos). Sólo [292, 257, 293, 294] utilizaron segmentos de señales del EEG como datos de entrada al modelo. Aunque los estudios utilizaron una gran variedad de características espectrales, el procedimiento más común consistió en la descomposición en bandas de frecuencia.

#### 4.2.2.4. Evaluación de los modelos

Como se muestra la Tabla E.1 del Apéndice E, una de las características más destacadas de los artículos seleccionados tiene que ver con la variedad de modelos utilizados. Podemos apreciar que el número de modelos implementados superó al número de artículos seleccionados debido a que, en 18 artículos (correspondiendo al 41,86%), se realizaron estudios comparativos entre varios modelos, mientras que los restantes artículos utilizaron un único modelo. Dentro de este último grupo, el modelo más utilizado fue el MLP, considerado en 5 artículos, seguido de CNN utilizado en 4 estudios, y SVM y RF, considerados en 3 artículos cada uno.

La imagen A de la Figura 4.12 muestra un diagrama de barras que contiene la lista de los modelos utilizados en los artículos incluidos en esta revisión junto con el número de veces que aparecieron, además de identificar aquellos modelos que utilizaron procesamiento simbólico (señalados en rojo) y procesamiento subsimbólico (señalados en azul). Para este análisis, se tuvo en cuenta que J48 es un modelo basado en DT, VGG-16 es un modelo basado en CNNs, y que SGD, “adaboost”, “bagging” y “vote method” son técnicas de entrenamiento y optimización y por lo tanto no pertenecen al grupo de modelos utilizados. Además, SL se incorporó al grupo de LR. El grupo ANN contiene tanto las ANN no especificadas (los artículos correspondientes dieron a entender que se trataban de MLP) como BP-NN y FF-NN. Como era de esperar, dado que el dataset estaba compuesto por series temporales, predominaron aquellos modelos con procesamiento subsimbólico, considerados el 82,61% de las veces, mientras que RF, DT y RT fueron los únicos modelos utilizados dentro del grupo de procesamiento simbólico. Para enfatizar estas diferencias, cabe señalar que mientras que el 17,39% de los artículos consideraron modelos con procesamiento simbólico, sólo 2 artículos los utilizaron de forma exclusiva. En cuanto a los modelos individuales más utilizados, tal y como podemos ver en la imagen A de la Fig. 4.12, encontramos a SVM y ANN, considerados por 14 estudios, seguidos de KNN, utilizado en 10 estudios, y los modelos RF y CNN, que aparecieron en 9 estudios.

En la imagen B de la Figura 4.12 destaca la importancia de las técnicas de ML neuronales, que constituyó el 42,39% de los modelos utilizados, aunque los modelos individuales más utilizados fueron SVM, una técnica de ML no-neuronal, y ANN, una técnica de ML neuronal, en la misma proporción. Las redes neuronales profundas son redes con arquitecturas más complejas que permiten un mayor nivel de abstracción a expensas de una mayor potencia computacional, por lo que han ido adquiriendo cada vez mayor importancia y por lo tanto, se esperaba que su desarrollo se concentrara en los últimos años. De hecho, desde 2019 encontramos 28 artículos, de los cuales descartamos [295, 296] por no aportar suficiente información sobre el modelo utilizado. De

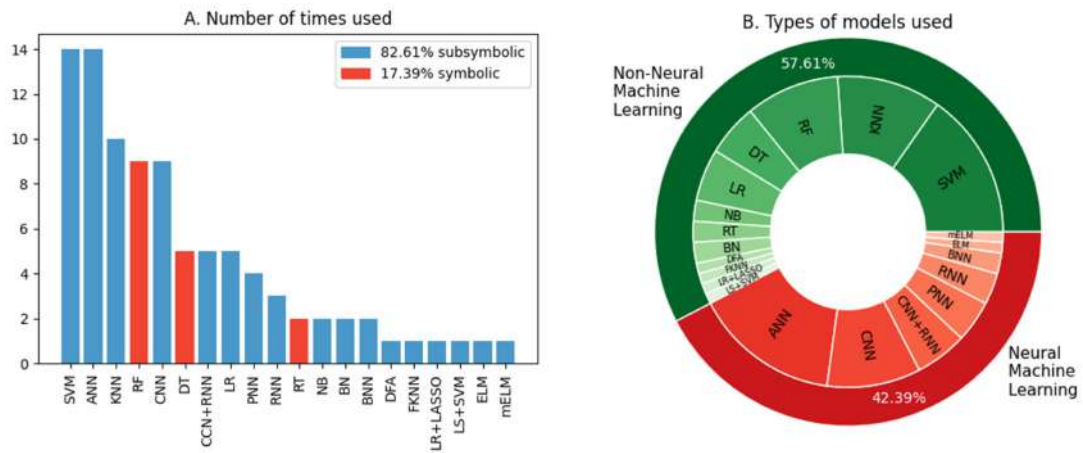


Figura 4.12: A: Modelos considerados en los artículos que componen la revisión junto con el número de apariciones indicando su pertenencia al grupo de procesamiento simbólico o subsimbólico. B: Diagrama circular del tipo de modelo utilizado, especificando las técnicas de ML no-neuronales y aquellas que pertenecen al subgrupo de ANN. Fuente: [236].

los 26 artículos restantes, el 57,69% utilizó modelos de ML neuronales, predominando el uso de las CNNs, presente en el 34,62% de esos estudios.

En cuanto a la arquitectura de los modelos empleados, cabe señalar que, si bien se especifica para la mayoría de ellos, el grado de especificidad de esta información tiene una gran variabilidad. De hecho, para los modelos más complejos, parámetros como el número de unidades en cada capa, la función de activación utilizada, el optimizador implementado, el valor de la tasa de aprendizaje o la función de pérdida elegida, se especificaron en algunos de los artículos, además de no ser uniforme en todos los modelos considerados. También hay que tener en cuenta la ausencia de una tabla de referencia con la que podamos realizar comparaciones, dificultando la comparación entre los estudios. Ambos puntos impidieron proporcionar una evaluación precisa de la arquitectura de los modelos. Entre los parámetros más destacados encontramos el uso del optimizador Adam, la función de activación ReLU y softmax para la última capa. En la fase de entrenamiento y validación, encontramos que 24 de los artículos seleccionados (55,81%) utilizaron técnicas de validación cruzada (2 de los cuales fueron LOO), y 16 de los artículos consideraron una división del dataset para conjuntos de entrenamiento, validación y prueba, según los porcentajes fijados en cada estudio. En particular, para este último grupo de artículos, la moda fue utilizar en torno al 70% de los datos para el entrenamiento, considerado en 6 de ellos. Los artículos restantes [297, 298, 299] no especificaron la metodología utilizada.

Para evaluar los resultados obtenidos, los artículos consideraron diferentes métricas,

siendo el accuracy la más utilizada apareciendo en el 90,70 % de los estudios, seguida de sensitivity, utilizada en el 69,77 %, y specificity, utilizada en el 46,51 % de los artículos considerados en la Tabla E.1. Para los 31 artículos que cubrían el problema diagnóstico de la EP (se ha incluido el estudio que realizaba la extracción de características en el problema de clasificación de la EP), se seleccionaron aquellos estudios con resultados superiores al 90 % en al menos dos métricas, estos fueron: [286, 294, 290, 291] usando ANNs con valores superiores al 97 % para sensitivity y precision, [300, 301, 302] considerando CNNs con valores superiores al 99 % para accuracy, precision y sensitivity ([300, 301] se basaron en el mismo estudio), [303, 292] utilizando CNN+RNN con valores superiores al 93 % en accuracy, sensitivity y precision, [254, 259] utilizando SVM con valores de accuracy y sensitivity superiores al 93 %, y [260] utilizando DFA con valores de accuracy, specificity y sensitivity superiores al 94 %. Para estos artículos se evaluó el método de validación realizado para obtener información acerca de la capacidad del modelo para generalizar los resultados en pruebas ciegas de test. Se encontró que el método predominante fue la validación cruzada para 9 de los 12 artículos, mientras que [294, 302, 303] dividieron los datos en entrenamiento, validación y prueba. Los modelos con mejores resultados en el problema de clasificación de pacientes con EP y controles fueron CNN, y el grupo de modelos ANN.

### 4.2.3. Resultados que contribuyen al desarrollo de esta tesis

Los objetivos de los artículos que constituyen las revisiones bibliográficas pueden dividirse principalmente en tres grupos: diagnóstico de la EP, evaluación del deterioro cognitivo y alteraciones en la marcha. El problema más abordado es el diagnóstico de la EP, presente en el 69,77 % de los artículos. En cuanto a la identificación del nivel de deterioro cognitivo de los pacientes con EP, todos los artículos que abordan este estudio han sido publicados en los años 2019 y 2021, indicando la novedad del tema y marcando la evolución en el estudio y diagnóstico de la EP.

Para los artículos que abordan el diagnóstico de la EP se evaluaron las técnicas de DL utilizadas y cómo se han aplicado para analizar los EEGs de pacientes con EP. Este estudio se realizó por medio de tres puntos clave dentro de las técnicas de ML:

- **Calidad del conjunto de datos.**

El uso de clases equilibradas durante el entrenamiento de modelos de ML y DL, es importante para garantizar resultados fiables, ya que en caso de desequilibrio se favorece la clase mayoritaria dando lugar a métricas de precisión sesgadas. El 69,77 % de los estudios trabajaron con conjuntos de datos balanceados con un promedio de  $23,10 \pm 16,74$  sujetos por clase, proporcionando una medida acerca

del número adecuado de sujetos para un problema de clasificación. Los parámetros clínicos como la edad, la medicación y los años de enfermedad pueden influir en el rendimiento del modelo, y sólo el 34,88 % de los artículos especificaron la información relacionada con los datos demográficos. En cuanto a los parámetros de registro del EEG encontramos una gran heterogeneidad, pudiendo afectar a la calidad de las señales, lo que resalta la ausencia de un protocolo estándar para realizar el registro de esta prueba. Predominaron los montajes con 14 electrodos junto con las pruebas de reposo. Esto puede deberse a la influencia de estudios de neuroimagen sobre el estado de reposo en enfermedades neurológicas como la EP.

■ **Pre-procesamiento de datos.**

Ninguna de las revisiones realizadas proporcionó un criterio unificado en cuanto al proceso de limpieza de los EEGs. El 27,91 % de los estudios no especificaron el proceso de limpieza del EEG por lo que asumimos la ausencia de pre-procesamiento, el 23,26 % realizó un pequeño pre-procesamiento en los datos (basado en la aplicación de filtros sin la eliminación artefactos) y el 48,84 % de los artículos utilizaron señales de EEGs libres de artefactos. Los resultados proporcionados por un mismo modelo para estos tres tipos de limpieza mostraron que el rendimiento fue independiente del proceso realizado, indicando que las técnicas de DL son robustas al ruido presente en los EEGs permitiendo eliminar el proceso de limpieza junto con las limitaciones implicadas. También encontramos una gran heterogeneidad en cuanto a las características extraídas del EEG, aunque predominó el uso de características espectrales en un 88,37 %. Esto puede deberse a que las características espectrales brindan información sobre variaciones en las bandas del EEG, y las alteraciones en estas bandas proporcionan más información clínica que un análisis morfológico de la señal, especialmente en la EP, donde no se observan alteraciones visuales en las señales.

■ **Evaluación de los modelos utilizados.**

La elección del modelo está influenciada tanto por el tamaño del conjunto de datos como por la naturaleza de los mismos, dado que nuestro dataset se compone de series temporales, siendo cada una de ellas un canal del EEG, es de esperar que los modelos con procesamiento subsimbólico sean los más utilizados. Entre ellos predominan las técnicas no-neuronales, ya que requieren una menor cantidad de datos para su entrenamiento y debemos tener en cuenta que la adquisición de datos clínicos es compleja. Cosecuentemente, el modelo más utilizado ha sido SVM aunque también destaca el uso de las ANN, siendo las técnicas neurona-

les más utilizadas en los últimos años. El grado de descripción de la arquitectura de los modelos es enormemente heterogénea, lo que dificulta la reproducibilidad del experimento. El método de validación predominante fue el k-fold cross-validation, considerado por el 55,81 % de los artículos con  $k = 5, 10$ , ya que minimiza el sesgo producido por la elección de los datos. Finalmente, el rendimiento de los modelos se evaluó por medio de la métrica accuracy, utilizada en un 90,70 %, seguida de la métrica de sensibilidad, considerada en el 69,77 % de los estudios. Cabe señalar que en los artículos más recientes, la evaluación de los resultados se realizó mediante varias métricas, lo que proporcionó más información acerca del rendimiento del modelo.

La metodología que implementaremos tendrá en cuenta estas características, que serán analizadas para evaluar de forma objetiva los resultados de nuestro modelo y su capacidad de generalización. Utilizaremos medidas temporales y espectrales, pero no realizaremos un proceso de limpieza de artefactos, aunque en algún caso podemos probar únicamente con algún tipo de filtrado para acotar el espectro al intervalo que vayamos a analizar.

En cuanto a los modelos implementados, dentro del objetivo de diagnóstico o identificación de la enfermedad, no encontramos estudios relacionados con la conectividad funcional. Por ello, proponemos un análisis basado en la teoría de grafos y un modelo GCNN para analizar las diferencias de conectividad existentes entre sujetos sanos y pacientes con EP. Además, desde un punto de vista más computacional, observamos la ausencia de modelos de PLN para el análisis de las señales del EEG. Dado el carácter secuencial de estas señales podemos realizar una analogía con el texto para utilizar estas técnicas sobre las señales en bruto buscando relaciones entre los datos que permitan identificar la EP.

### 4.3. Conectividad funcional en pacientes con EP

Las alteraciones en los patrones de conectividad cerebral pueden estar asociadas a diferentes patologías o trastornos neurológicos además de representar alteraciones farmacológicas. Por ello, en la búsqueda de un biomarcador de la enfermedad, estudiaremos las diferentes relaciones entre las conexiones de los pacientes con EP en ausencia de estimulación dopaminérgica (esto es, en estado OFF), y tras la toma de su medicación (esto es, estado ON) con respecto a los controles. Esperamos que las diferencias de conectividad existentes entre los estados de activación dopaminérgica, identifiquen características esenciales en la enfermedad que trasciendan los efectos de la medicación. Este estudio también puede ofrecer una evaluación acerca de cómo la LD afecta a

la conectividad y si este estado se aproxima al de un paciente sano. Por ello, se realizará una comparación de los estados dopaminérgicos.

Para realizar este tipo de estudio basado en la *hipótesis 4* de esta tesis, se ha utilizado la teoría de grafos para modelar los EEGs, de modo que facilite la obtención de información relacionada con los cambios en las conexiones y permita establecer relaciones entre las diferentes regiones cerebrales. Sin embargo, el proceso de modelado puede ser largo además de no estar optimizado, ya que no existe una herramienta específica que permita a los investigadores representar los EEGs como grafos para sus estudios. Por ello, desarrollamos la librería EEGraph, que ha sido anteriormente detallada, encargada de modelar un EEG por medio de un grafo bajo las variables definidas por el usuario. Con esta herramienta enfrentaremos el estudio de conectividad por medio de EEGs en busca de un biomarcador de la enfermedad.

Los participantes considerados en este estudio fueron 24 controles y 24 pacientes con EP cuyos datos demográficos se especifican en la Tabla 4.3.

	Controles	EP
N (Hombres)	24(18)	24(11)
Edad en años	62,54 ± 9,21	62,54 ± 11,27
HY	-	2,02 ± 0,76
UPDRS	-	20,50 ± 9,90
Años Enfermedad	-	6,83 ± 4,71
Total dopamina por día (mg)	-	655,96 ± 415,00

Tabla 4.3: Datos demográficos y clínicos de los participantes incluidos en el estudio de conectividad cerebral basado en el modelado de EEGs como grafos. Se muestra el valor medio y la desviación estándar de los datos demográficos y clínicos de los participantes. Fuente: elaboración propia.

Con respecto a los datos demográficos de los participantes expuestos en la Tabla 4.3, podemos observar que se ha buscado la equidad en cuanto a género. Sin embargo, dado que la afectación de la EP predomina en hombres, los participantes masculinos incluidos en el grupo de los controles tienen una mayor representación. También se ha tratado de que ambos conjuntos fueran equivalentes en edad, ya que esta es una característica que influye en las señales del EEG. Para corroborar estos criterios, se ha realizado un estudio de significancias comparando ambas muestras mediante una distribución t-Student, en la cual se obtuvo un p-valor de 0,68 para la edad y un valor de 0,04 para el género.

Para llevar a cabo este estudio, se desarrollaron un total de 4 modelos que se especifican a continuación:

- **Modelo 1: Controles-Pre.** Se realiza una clasificación binaria con los EEGs de controles y pacientes con EP en estado OFF, esto es sin afectación por levodopa.

Con ello buscamos que el modelo identifique las características intrínsecas a la enfermedad para poder clasificar los EEGs.

- **Modelo 2: Controles-Post.** Se realiza una clasificación binaria con los EEGs de controles y pacientes con EP en estado ON, esto es, en el periodo de máxima actuación de la medicación dopaminérgica. El propósito de este estudio es el de identificar a los pacientes con EP cuyo comportamiento está influenciado por la medicación. De modo que comparando estos resultados con los obtenidos por el *Modelo 1*, podremos evaluar las diferencias en la clasificación producida por la levodopa.
- **Modelo 3: Controles-Both.** Se realiza una clasificación binaria con los EEGs de controles y pacientes con EP. En el caso de los pacientes con EP, la mitad se encuentran en estado ON y la otra mitad en estado OFF, sumando un total de 24 participantes distintos. Para determinar el conjunto de pacientes con EP, se ha tenido en cuenta la individualidad de los sujetos, es decir, no se ha tomado el mismo paciente en diferentes estados de activación dopaminérgica. Con ello buscamos comprobar que el modelo es capaz de clasificar a los pacientes con EP independientemente de la medicación, para ello el modelo debe identificar componentes intrínsecas a la enfermedad, ya que los EEGs pueden o no estar influenciados por la medicación.
- **Modelo 4: Pre-Post.** Se realiza una clasificación binaria con los EEGs de pacientes con EP en estados pre y post LD. El objetivo de este estudio es identificar las variaciones en la conectividad derivadas de la medicación.

La Figura 4.13 muestra las posibles combinaciones entre los grupos existentes de los participantes.

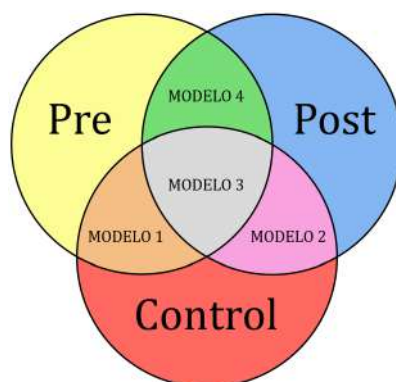


Figura 4.13: Se muestran las posibles combinaciones entre los conjuntos de datos incluidos en el estudio, estos son, pacientes en estado ON (post LD), pacientes en estado OFF (pre LD) y controles. Fuente: elaboración propia.



En la Tabla 4.4 se presenta un resumen con los objetivos de los modelos que incluyen los EEGs de los controles, y en la Tabla 4.5 se presenta el resumen del objetivo del modelo que evalúa la actuación de la dopamina.

	Pre LD	Post LD	Both
Controles	Identificar las diferencias de conectividad entre controles y pacientes con EP en ausencia de estimulación dopaminérgica	Identificar las diferencias de conectividad entre controles y pacientes con EP bajo la influencia de la medicación	Identificar las diferencias de conectividad entre controles y pacientes con EP independientemente de la influencia dopaminérgica

Tabla 4.4: Se presenta el objetivo de cada uno de los modelos especificados que contienen los EEGs de los controles. Fuente: elaboración propia.

	Post LD
Pre LD	Identificar las diferencias de conectividad producidas por los diferentes estados de activación dopaminérgico en pacientes con EP

Tabla 4.5: Se presenta el objetivo del modelo que utiliza los EEGs de pacientes con EP en los diferentes estados de activación dopaminérgica. Fuente: elaboración propia.

Los EEGs de los participantes incluidos se modelaron como grafos por medio de la librería desarrollada EEGraph, teniendo en cuenta medidas de conectividad relacionadas con el dominio espectral y temporal de las señales. Con respecto al dominio temporal, utilizamos la correlación de Pearson y la entropía de Shannon. Con respecto al dominio espectral, utilizamos la coherencia y la entropía espectral. En el caso de las medidas relacionadas con el dominio de las frecuencias, se obtuvieron los grafos de las 5 bandas especificadas en esta tesis. Por tanto, para cada uno de los modelos especificados, se realizaron 4 experimentos, cada uno de ellos asociados a una medida de conectividad diferente.

- **Experimento a: Correlación de Pearson.** Se obtienen los grafos cuyas conexiones vienen determinadas por la correlación de Pearson entre pares de electrodos.
- **Experimento b: Entropía de Shannon.** Se obtienen los grafos cuyas conexiones vienen determinadas por la entropía de Shannon para cada electrodo que contiene el casco de electroencefalografía.

- **Experimento c: Coherencia.** Se obtienen los grafos cuyas aristas vienen definidas por la coherencia entre pares de electrodos para cada una de las bandas especificadas (delta, theta, alpha, beta y gamma). Cada una de estas bandas se utilizará de forma individual en los entrenamientos, esto es, se realizará un entrenamiento diferente para cada banda. Además, se realizará un experimento adicional en el que se incluyan los grafos de todas las bandas espectrales.
- **Experimento d: Entropía espectral.** Se obtienen los grafos cuyas aristas vienen definidas por la entropía espectral de cada canal del EEG para cada una de las bandas especificadas (delta, theta, alpha, beta y gamma). Cada una de estas bandas se utilizará de forma individual en los entrenamientos, y se realizará un experimento adicional en el que se incluyan los grafos de todas las bandas espectrales.

La Tabla 4.6 muestra un resumen de la cantidad de ejecuciones que se realizarán por cada uno de los experimentos especificados.

Correlación de Pearson	Entropía de Shannon	Coherencia	Entropía Espectral
1. Coeficiente de correlación de Pearson	1. Valor de la entropía de Shannon	1. delta 2. theta 3. alpha 4. beta 5. gamma 6. all	1. delta 2. theta 3. alpha 4. beta 5. gamma 6. all

Tabla 4.6: Se presenta, para cada medida de conectividad, la cantidad de procesos que se llevarán a cabo en la experimentación. Fuente: elaboración propia.

Por tanto, para cada modelo se realizarán 4 experimentos, sumando un total de 14 ejecuciones por modelo. El código de identificación de cada ejecución será el empleado en las definiciones expuestas, esto es, un código numérico (1, 2, 3, o 4) para los modelos, y un código simbólico (*a*, *b*, *c*, o *d*) para los experimentos con las diferentes medidas de conectividad. De modo que para especificar el proceso Controles-Both con la entropía de Shannon utilizaremos el código 3.b. En el caso de las medidas espectrales se especificará la banda utilizada. En el caso de realizar la ejecución con todas las bandas, se identificará con “all”. De modo que para especificar el proceso Controles-Pre con la coherencia en la banda beta utilizaremos el código 1.3.beta. Esta notación se mantendrá a lo largo de este estudio.

Para evaluar la generalización de los resultados de estos procesos más allá de la validación y prueba empleados por el modelo, se introducirá un conjunto externo de prueba compuesto por pacientes con EP tras la toma diaria de su medicación pero no necesari-

riamente en el estado ON. Los datos demográficos de estos pacientes se muestran en la Tabla 4.7.

	EP	p-valor
N (Hombres)	13(8)	0,38
Edad en años	61,69 ± 11,30	0,92
HY	1,54 ± 0,54	0,07
UPDRS	16,69 ± 10,55	0,42
Años Enfermedad	7,08 ± 5,65	0,82
Total dopamina por día (mg)	691,62 ± 340,84	0,62

Tabla 4.7: Se muestra el valor medio y la desviación estándar de los datos demográficos y clínicos de los pacientes con EP incluidos en el conjunto de prueba para el estudio de conectividad basado en el modelado de EEGs como grafos. También se muestra el p valor obtenido por medio de una distribución t-Student con respecto a los participantes considerados para el conjunto de entrenamiento. Fuente: elaboración propia.

La Tabla 4.7 muestra los valores obtenidos por una distribución t-Student, encargada de evaluar la semejanza entre dos ditribuciones de datos. En este caso, hemos comparado el grupo de test con los datos clínicos de los pacientes con EP que han sido introducidos en el proceso de entrenamiento. Los resultados muestran que hay una gran semejanza entre los grupos de edad y años de enfermedad, aunque las variables asociadas al estado del paciente muestran diferencias significativas.

#### 4.3.1. Diseño y arquitectura de la GCNN

Se realizaron pruebas con diferentes hiper-parámetros (estos son, aquellos parámetros que pueden ser modificados por el usuario) para identificar la combinación con la que el modelo proporciona mejores resultados. Entre los hiper-parámetros considerados encontramos:

- Tamaño de ventana.** Los resultados son muy sensibles al tamaño de ventana o intervalo de tiempo al que se segmenta el EEG. Intervalos más largos contienen más información acerca del estado del cerebro pero se pierde la secuencia temporal del EEG. Con intervalos pequeños, mantenemos el carácter dinámico de la serie pero disminuimos la cantidad de información que contiene cada intervalo, pudiendo no ser suficiente para llevar a acabo el aprendizaje del modelo. Los valores que han sido considerados son 2, 3 y 4 segundos, siendo el valor óptimo un tamaño de ventana de 2 segundos.
- Learning rate.** Los modelos de DL son muy sensibles a la tasa de aprendizaje del

optimizador, por lo que este es un hiper-parámetro que ajustamos para favorecer el aprendizaje del modelo. los valores considerados son  $10^{-2}$ ,  $10^{-3}$ , y  $10^{-4}$ , siendo el valor óptimo el de  $10^{-4}$ .

- **Batch size.** El tamaño del lote introducido para realizar el proceso de entrenamiento, también puede ser un factor influyente en los resultados del modelo. Se ha probado con 12, 25 y 50 grafos por lote, siendo 12 el valor óptimo.

Aunque las técnicas de DL han demostrado ser robustas al ruido de los EEGs, hemos realizado una prueba con EEGs en bruto y EEGs filtrados con un filtro pasa-baja a 49 Hz mediante la librería *mne*, evitando así el pico de la corriente eléctrica en España a 50 Hz [304]. El intervalo especificado, contiene la actividad cerebral asociada con las bandas espectrales y por tanto, alteraciones en estas bandas permiten realizar una interpretación clínica. La prueba con EEGs sin filtrar se ha realizado con los valores óptimos de learning rate y batch size, variando el tamaño de ventana y comparando dichos resultados. Dado que los resultados obtenidos con los EEGs filtrados fueron mejores en cada uno de los procesos de entrenamiento/validación/prueba para la métrica de accuracy, la experimentación se realizó con los EEGs filtrados para la correlación de Pearson.

Los grafos utilizados en este estudio han sido generados por la librería *EEGraph* para las 4 medidas de conectividad especificadas (correlación de Pearson, entropía de Shannon, coherencia y entropía espectral). Los EEGs se han segmentado en ventanas de 2 segundo, de modo que por cada participante, tendremos un total de 15 grafos para cada EEG asociado a una prueba de FT y 30 grafos para la prueba de resting.

La arquitectura del modelo utilizado consta de dos módulos, uno encargado de la parte de análisis de grafos y otro encargado de la parte de clasificación. Ambas partes se especifican a continuación:

- **GCNSupervisedGraphClassification.** Encargado de analizar la información contenida en los grafos por medio de su representación matricial y con las características asociadas a cada nodo del grafo. La entrada es, por tanto, una matriz cuadrada de dimensión 64 (coincidiendo con el número de canales del EEG), con función de activación ReLU y dropout de 0,2. El dropout es un método de regularización que busca evitar el sobre-entrenamiento y mejorar la capacidad de generalización de la red. Esta técnica que se encarga de desconectar al azar algunas neuronas de una capa durante el entrenamiento del modelo, obligando a estas capas a aprender el mismo concepto de distintas maneras. En este caso 0,2 se corresponde con un 20 % de las neuronas de dicha capa.
- **MLP.** Seguido al módulo anterior encontramos dos capas totalmente conectadas,

una con 32 neuronas seguida de otra con 16 neuronas, ambas con función de activación ReLU. Finalmente, la capa de salida es una capa totalmente conectada de una neurona con función de activación sigmoide.

Para introducir las características a cada nodo del grafo, estos son los valores de conectividad, utilizaremos la matriz de conectividad o “connectivity\_matrix” que proporciona EEGraph, en la que podemos encontrar todos los valores asociados a cada una de las medida escogidas. En el caso de las medidas que ofrecen información por pares de canales, cada fila de esta matriz está asociada a un electrodo y contiene el valor de las conexiones de dicho canal con el resto. En el caso de las medidas que obtengan los valores de los canales individuales, el valor se encontrará en la diagonal. Por tanto asociamos a cada nodo del grafo su fila correspondiente como características.

Dado que este modelo es sensible a las clases desbalanceadas, hemos considerado el mismo número de sujetos en cada clase. El optimizador utilizado fue Adam con un learning rate de  $10^{-4}$  y *binary\_crossentropy* como función de pérdida. Las funciones de activación en el módulo de predicción fueron de ReLU para las capas densas y sigmoide para la capa de salida del clasificador.

El método de validación empleado fue un 5-fold CV, tras separar previamente un 4% (que se corresponde con 2 controles y 2 pacientes con EP) de los datos para el conjunto ciego de prueba que se evaluó tras finalizar el entrenamiento de cada fold. La división en cada subconjunto se realizó aleatoriamente por paciente para evitar correlaciones entre los datos de entrenamiento, validación y prueba de grafos asociados al mismo sujeto. Con el modelo entrenado, se evaluó el conjunto externo con otros pacientes con EP en diferentes estados de activación dopaminérgica, para evaluar la generalización de los resultados obtenidos. Se calculó la media y la std para las ejecuciones del CV en todos los procesos para las métricas de accuracy, sensitivity, specificity y precision, por medio de los valores de TN, TP, FN y FP. La interpretación de estas medidas se expone a continuación:

- **Accuracy.** Indica la tasa total de aciertos del modelo.
- **Sensitivity.** Indica la tasa de pacientes con EP acertados, esto es, entre todos los pacientes con EP, cuántos ha acertado el modelo. Es una medida de sensibilidad a la enfermedad.
- **Specificity.** Indica la tasa de controles acertados, esto es, entre todos los controles, cuántos ha acertado el modelo.
- **Precision.** Indica la cantidad de pacientes con EP que ha acertado el modelo frente al total de pacientes que el modelo ha indicado que tienen EP.

### 4.3.2. Resultados de los estudios de conectividad relacionados con controles

#### 4.3.2.1. Medidas definidas en el dominio temporal

Las medidas de conectividad consideradas en este estudio definidas en el dominio temporal son, la correlación de Pearson y la entropía de Shannon. En la Tabla 4.9 se muestran los resultados obtenidos para las 4 métricas consideradas (accuracy, sensitivity, specificity y precision), en los procesos 1.a, 2.a y 3.a, para el conjunto ciego de prueba. En el caso de los modelos 1 y 2, este conjunto lo componen los EEGs de 2 pacientes con EP y 2 controles, mientras que el conjunto de prueba para el modelo 3 contiene 2 controles y 24 pacientes con EP (12 en estado pre LD y 12 en estado post LD que no se incluyeron en el conjunto de entrenamiento). Los valores de accuracy para los procesos de entrenamiento y validación se introducen para consolidar los resultados del proceso de prueba y corroborar la ausencia de overfitting en los entrenamientos de las distintas ejecuciones. Estos resultados se muestran en la Tabla 4.8.

Correlación de Pearson			
	Entrenamiento	Validación	Prueba
Control-Pre	92,08 ± 0,35	85,34 ± 9,32	56,51 ± 4,57
Control-Post	92,62 ± 0,49	85,06 ± 5,81	93,17 ± 4,85
Control-Both	92,27 ± 0,66	82,12 ± 8,53	95,24 ± 2,08

Tabla 4.8: Se muestra el valor medio y la desviación estándar del proceso de validación, para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 1.a, 2.a y 3.a. Fuente: elaboración propia.

En la Tabla 4.8 podemos observar, que para el caso del modelo 1.a, el resultado del proceso de prueba dista mucho de los obtenidos por los procesos de entrenamiento y validación, indicando que la muestra de EEGs extraída para ese conjunto de test (aún siendo aleatoria) no comparte tantas características comunes con los conjuntos anteriores. Sin embargo, la selección de sujetos de test se ha mantenido en todas las ejecuciones realizadas, lo que nos permite evaluar qué medida de conectividad permite una mayor generalización de los resultados.

Con respecto a la Tabla 4.9, los modelos 2.a y 3.a presentan resultados más homogéneos en todas las métricas, con valores muy próximos entre sí, salvo en el caso de la especificidad en el modelo Control-Both, cuyo valor es significativamente inferior con una desviación estándar que supera el 35% del valor medio. Dado que esta métrica se relaciona con la tasa de acierto de los controles, esto nos lleva a pensar que los valores obtenidos están condicionados por los pacientes sanos que conforman el grupo de test.

<b>Correlación de Pearson</b>				
	Accuracy	Sensitivity	Specificity	Precision
Control-Pre	56,51 ± 4,57	14,92 ± 7,15	98,10 ± 2,61	86,50 ± 18,67
Control-Post	93,17 ± 4,85	87,94 ± 7,98	98,41 ± 1,94	98,11 ± 2,46
Control-Both	95,24 ± 2,08	97,70 ± 0,78	65,71 ± 23,80	97,19 ± 1,87

Tabla 4.9: Resultados de las métricas accuracy, sensitivity, specificity y precision para la medida de la correlación de Pearson en el proceso de prueba de los modelos 1.a, 2.a y 3.a. Los resultados muestran el valor medio y la desviación estándar durante el proceso de validación cruzada. Fuente: elaboración propia.

La diferencia en los valores de sensitivity del modelo 1.a muestran un claro sesgo producido por los datos de los pacientes con EP del conjunto de prueba. Por tanto, comparando los resultados de estos tres modelos, el que ofrece rendimiento homogéneo en todas las métricas, es el modelo 2.a, aunque el modelo 3.a muestra valores en general más altos con una menor std.

Los resultados asociados a la entropía de Shannon para la métrica de accuracy en los procesos de entrenamiento, validación y prueba, se muestran en la Tabla 4.10, particularmente, para los modelos 1.b, 2.b y 3.b.

<b>Entropía de Shannon</b>			
	Entrenamiento	Validación	Prueba
Control-Pre	96,36 ± 0,46	85,74 ± 8,52	65,40 ± 12,76
Control-Post	97,07 ± 0,51	81,09 ± 8,88	96,83 ± 1,86
Control-Both	96,86 ± 0,57	85,84 ± 4,58	89,39 ± 1,67

Tabla 4.10: Se muestra el valor medio y la desviación estándar para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 1.b, 2.b y 3.b. Fuente: elaboración propia.

En la Tabla 4.10 observamos nuevamente que el conjunto de prueba del modelo 1.b dista mucho de los resultados de validación, por lo que corroboramos que los pacientes escogidos en para el conjunto ciego no son una muestra representativa de los pacientes introducidos en los conjuntos de entrenamiento y validación de la red. Comparando los resultados de los procesos para el modelo 2.b, podemos observar un indicio de overfitting, al encontrar bastante diferencia entre los valores medios de los resultados de entrenamiento y validación. Sin embargo, dado que los valores de test (que han sido escogidos aleatoriamente) han mostrado ser mejores que los de validación con una menor desviación estándar, descartamos un posible sobre-ajuste del modelo.

<b>Entropía de Shannon</b>				
	<b>Accuracy</b>	<b>Sensitivity</b>	<b>Specificity</b>	<b>Precision</b>
Control-Pre	65,40 ± 12,76	34,29 ± 29,43	96,51 ± 3,95	94,30 ± 5,32
Control-Post	96,83 ± 1,86	97,46 ± 2,66	96,19 ± 1,81	96,25 ± 1,78
Control-Both	89,39 ± 1,67	89,23 ± 1,92	91,27 ± 5,24	99,20 ± 0,47

Tabla 4.11: Resultados de las métricas accuracy, sensitivity, specificity y precision para la medida de la entropía de Shannon en el proceso de prueba para los modelos 1.b, 2.b y 3.b. Se muestra el valor medio y la desviación estándar durante el proceso de validación cruzada. Fuente: elaboración propia.

En la Tabla 4.11 encontramos que los resultados del modelo 1 (Control-Pre) vuelven a mostrarse sesgados por los pacientes con EP introducidos en el conjunto de prueba. Para esta medida de conectividad, el modelo 2.b (Control-Post) presentó valores más altos en la mayoría de las métricas además de presentar valores más precisos, esto es, menor desviación estándar en los resultados.

Comparando los valores obtenidos por ambas medidas de conectividad, en la Tabla 4.9 para la correlación de Pearson y la Tabla 4.11 para la entropía de Shannon, encontramos que los mejores resultados se obtienen para el modelo 2.b. Debemos de tener en cuenta que, aunque los resultados de test para los modelos 1 y 2 son muy diferentes, los valores de validación no muestran diferencias significativas entre estos modelos. Por tanto, los resultados de estos procesos pueden estar condicionados por la elección del conjunto de pacientes. Se debe notar, que la entropía de Shannon ha mostrado una mayor capacidad de generalizar los resultados al distinguir mejor los pacientes con EP en el modelo 1.b, y los controles en el modelo 3.b con respecto a la correlación de Pearson.

#### 4.3.2.2. Medidas definidas en el dominio espectral

A continuación se presentan los resultados de las medidas de conectividad definidas en el dominio espectral. La Tabla 4.12 contiene los resultados para la métrica de accuracy, en los procesos de entrenamiento, validación y prueba. La Tabla 4.13 contiene los resultados de las métricas para los modelos 1.c y 1.d, esto es, para el modelo Control-Pre con los experimentos de coherencia y entropía espectral.

Recordemos que los resultados para el conjunto de test estaban sesgados por los pacientes con EP. Sin embargo, en la Tabla 4.12, podemos observar que para las medidas de conectividad definidas en el dominio espectral, no hay tantas diferencias entre los resultados de validación y prueba en las bandas espectrales. Únicamente encontramos diferencias para el caso 1.c.gamma. Consecuentemente, podemos decir que las medidas espectrales permiten una mayor generalización de los resultados. También llama la



<b>Control-Pre</b>			
	Entrenamiento	Validación	Prueba
<b>Coherencia</b>			
Delta	94,40 ± 0,86	86,36 ± 9,69	82,70 ± 6,60
Theta	91,59 ± 0,53	86,04 ± 13,05	69,68 ± 9,16
Alpha	88,26 ± 0,76	85,48 ± 17,18	75,08 ± 11,52
Beta	83,17 ± 1,59	76,49 ± 16,69	71,75 ± 7,97
Gamma	82,95 ± 1,50	72,24 ± 18,35	48,89 ± 20,65
All	89,80 ± 0,83	85,56 ± 12,16	80,95 ± 4,54
<b>Entropía espectral</b>			
Delta	94,96 ± 1,27	80,90 ± 6,79	73,97 ± 3,94
Theta	92,76 ± 1,14	80,16 ± 6,01	72,38 ± 3,24
Alpha	92,93 ± 1,69	80,63 ± 8,09	76,03 ± 3,39
Beta	90,57 ± 1,91	75,99 ± 3,86	70,63 ± 2,45
Gamma	90,31 ± 1,39	73,11 ± 5,70	74,13 ± 1,33
All	64,15 ± 0,56	57,06 ± 2,94	57,90 ± 0,89

Tabla 4.12: Se muestra el valor medio y la desviación estándar del proceso de validación, para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 1.c y 1.d. Fuente: elaboración propia.

<b>Control-Pre</b>				
	Accuracy	Sensitivity	Specificity	Precision
<b>Coherencia</b>				
Delta	82,70 ± 6,60	67,62 ± 12,63	97,78 ± 4,14	97,09 ± 5,31
Theta	69,68 ± 9,16	43,81 ± 14,44	95,56 ± 9,94	93,00 ± 15,65
Alpha	75,08 ± 11,52	50,16 ± 23,05	100 ± 0,00	100 ± 0,00
Beta	71,75 ± 7,97	45,08 ± 17,55	98,41 ± 3,55	97,73 ± 5,08
Gamma	48,89 ± 20,65	33,02 ± 9,48	64,76 ± 46,66	70,04 ± 35,40
All	80,95 ± 4,54	62,10 ± 9,30	99,81 ± 0,43	99,73 ± 0,60
<b>Entropía espectral</b>				
Delta	73,97 ± 3,94	56,19 ± 6,41	91,75 ± 2,84	87,13 ± 4,44
Theta	72,38 ± 3,24	59,68 ± 5,68	85,08 ± 2,13	79,94 ± 3,09
Alpha	76,03 ± 3,39	60,63 ± 7,31	91,43 ± 2,41	87,67 ± 2,99
Beta	70,63 ± 2,45	65,71 ± 2,88	75,56 ± 6,11	73,21 ± 4,89
Gamma	74,13 ± 1,33	69,84 ± 5,50	78,41 ± 5,09	76,64 ± 3,05
All	57,90 ± 0,89	53,08 ± 6,75	62,73 ± 7,99	59,06 ± 2,55

Tabla 4.13: Resultados de las métricas accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el experimento de controles vs. pacientes pre LD. Los resultados muestran el valor medio y la desviación estándar durante el proceso de validación cruzada para cada métrica del conjunto de prueba. Fuente: elaboración propia.

atención que los procesos de validación y prueba de la entropía espectral, tienen menor std que para la coherencia aunque los valores asociados son menores. En cuanto al proceso de aprendizaje de la red, podemos apreciar que el caso de la entropía espectral, modelo 1.d, presenta indicios de sobre-ajuste, claramente visibles en el caso de las bandas asociadas a altas frecuencias (beta y gamma).

Los resultados de las 4 métricas del conjunto de prueba para los modelo 1.c y 1.d se muestran en la Tabla 4.13. Observando los resultados proporcionados por la sensitivity, podemos apreciar el sesgo producido por los EEGs de los pacientes con EP en el conjunto de prueba para ambas medidas de conectividad. Particularmente para la coherencia, podemos apreciar que la banda delta junto con el experimento que introduce todas las bandas al modelo, son los más significativos para identificar la enfermedad. En el caso de la entropía espectral, encontramos menos fluctuaciones entre las bandas con una menor std en todos los procesos ejecutados. En este caso, las bandas beta y gamma son las más significativas para identificar la enfermedad pero las que peor tasa de aciertos tienen en los controles. Llama la atención la diferencia entre los resultados proporcionados por los modelos 1.c.all y 1.d.all, ya que en el caso de la entropía espectral este experimento no realizó un proceso de entrenamiento óptimo a pesar de tener una mayor cantidad de datos. Esto puede deberse a que la red no ha sido capaz de extraer patrones comunes a todas las bandas introducidas para identificar la EP.

<b>Control-Post</b>			
	Entrenamiento	Validación	Prueba
<b>Coherencia</b>			
Delta	94,79 ± 0,99	90,62 ± 8,51	95,08 ± 5,54
Theta	92,04 ± 1,40	87,26 ± 7,45	96,19 ± 4,57
Alpha	88,64 ± 2,05	83,82 ± 16,15	92,54 ± 13,42
Beta	84,88 ± 1,88	78,83 ± 13,45	68,73 ± 20,62
Gamma	84,21 ± 1,44	78,94 ± 9,15	66,35 ± 19,57
All	90,62 ± 1,03	82,61 ± 9,11	84,54 ± 7,03
<b>Entropía espectral</b>			
Delta	94,66 ± 1,33	80,01 ± 7,32	81,90 ± 1,89
Theta	92,44 ± 1,25	78,00 ± 8,88	85,24 ± 3,35
Alpha	93,07 ± 1,69	78,93 ± 9,30	85,08 ± 3,61
Beta	90,71 ± 1,96	74,24 ± 9,04	76,03 ± 3,65
Gamma	90,95 ± 1,97	72,70 ± 8,39	71,27 ± 4,74
All	64,65 ± 1,34	57,18 ± 4,65	57,33 ± 0,89

Tabla 4.14: Se muestra el valor medio y la desviación estándar del proceso de validación, para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 2.c y 2.d. Fuente: elaboración propia.

En la Tabla 4.14 se muestran los resultados del accuracy en los procesos de entrenamiento, validación y prueba para los modelos 2.c y 2.d, esto es, controles frente a pacientes con EP tras la toma de la medicación. Los resultados de los procesos de entrenamiento y validación vuelven a mostrar una tendencia al overfitting en el caso de la entropía espectral, dando a entender que esta medida requiere de una mayor cantidad de datos para realizar un entrenamiento óptimo. Sin embargo, estos resultados muestran una menor std que en el caso de la coherencia. Volvemos a observar una menor std en los valores asociados a la entropía espectral comparados con los obtenidos por la coherencia, pero con una presencia de sobre-entrenamiento. Sin embargo, el modelo ha mostrado poder generalizar los resultados, con valores de test próximos a los obtenidos por la validación.

<b>Control-Post</b>				
	Accuracy	Sensitivity	Specificity	Precision
<b>Coherencia</b>				
Delta	$95,08 \pm 5,54$	$95,24 \pm 3,37$	$94,92 \pm 10,49$	$95,67 \pm 8,79$
Theta	$96,19 \pm 4,57$	$99,37 \pm 1,42$	$93,02 \pm 9,63$	$94,06 \pm 8,17$
Alpha	$92,54 \pm 13,42$	$97,46 \pm 5,68$	$87,62 \pm 27,68$	$92,35 \pm 17,10$
Beta	$68,73 \pm 20,62$	$70,79 \pm 17,91$	$66,67 \pm 44,60$	$78,02 \pm 26,08$
Gamma	$66,35 \pm 19,57$	$82,86 \pm 12,77$	$49,84 \pm 46,96$	$70,96 \pm 26,70$
All	$84,54 \pm 7,03$	$69,08 \pm 14,07$	$100 \pm 0,00$	$100 \pm 0,00$
<b>Entropía espectral</b>				
Delta	$81,90 \pm 1,89$	$70,79 \pm 2,41$	$93,02 \pm 3,65$	$91,20 \pm 4,09$
Theta	$85,24 \pm 3,35$	$79,68 \pm 9,68$	$90,79 \pm 4,26$	$90,00 \pm 3,50$
Alpha	$85,08 \pm 3,61$	$77,46 \pm 5,19$	$92,70 \pm 3,48$	$91,42 \pm 3,95$
Beta	$76,03 \pm 3,65$	$73,65 \pm 6,79$	$78,41 \pm 8,59$	$77,98 \pm 6,71$
Gamma	$71,27 \pm 4,74$	$56,83 \pm 7,97$	$85,71 \pm 2,51$	$79,70 \pm 4,37$
All	$57,33 \pm 0,89$	$51,30 \pm 8,00$	$63,37 \pm 7,72$	$58,52 \pm 1,66$

Tabla 4.15: Resultados de las métricas accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el experimento de controles vs. pacientes post LD. Los resultados muestran el valor medio y la desviación estándar durante el proceso de validación cruzada para cada métrica en el conjunto de prueba. Fuente: elaboración propia.

En la Tabla 4.15 se incluyen los resultados para las 4 métricas seleccionadas de los modelos 2.c y 2.d para el conjunto de test. Los valores del modelo 2.c, muestran una gran std en las bandas alpha, beta y gamma, indicando grandes fluctuaciones en las predicciones de la red, lo que en algunos casos, invalida los resultados de las métricas. Si observamos los resultados de sensibilidad, encontramos que las bandas más significativas para identificar la EP son delta, theta y alpha, siendo la banda theta la que ofrece un mejor rendimiento. En el caso de la entropía espectral, los resultados presen-

tan menos fluctuaciones entre las bandas, además de mostrar una menor std en todas las ejecuciones implementadas con respecto a los resultados obtenidos por la coherencia. Debemos notar que en este caso, las bandas más significativas para la identificación de la enfermedad son theta y alpha, aunque esta medida de conectividad muestra una mejor tasa de aciertos en los controles que en los pacientes con EP. Particularmente, llama la atención los resultados del modelo 2.c.all, ya que el modelo muestra un gran rendimiento en la identificación de controles pero los valores de sensibilidad son mucho menores que en los obtenidos por otras medidas de conectividad.

Finalmente, en la Tabla 4.16 se muestran los resultados de los procesos de entrenamiento, validación y prueba para la métrica de accuracy en los modelos 3.c y 3.d. Debemos notar que estos experimentos muestran una menor std que en los casos anteriores, destacando los resultados de las bandas delta, theta y all para la coherencia. La entropía espectral vuelve a presentar valores más bajos de rendimiento pero con menor std que en el caso de la coherencia. Debemos notar que, aunque nuevamente tenemos la presencia de sobre-ajuste en el caso de la entropía espectral, ya que los participantes incluidos en el grupo de test (2 controles y 24 pacientes con EP) no parecen estar condicionados por el overfitting, ya que la red ha sido capaz de generalizar los resultados con bastante eficiencia.

<b>Control-Both</b>			
	Entrenamiento	Validación	Prueba
<b>Coherencia</b>			
Delta	95,60 ± 0,56	85,66 ± 8,80	93,32 ± 2,31
Theta	91,61 ± 0,46	88,68 ± 9,21	94,80 ± 3,65
Alpha	88,34 ± 1,16	81,07 ± 10,17	85,07 ± 11,10
Beta	83,08 ± 2,18	73,13 ± 14,58	85,69 ± 11,10
Gamma	83,24 ± 1,28	77,68 ± 8,59	86,28 ± 7,59
All	90,09 ± 0,69	84,73 ± 3,37	95,21 ± 3,03
<b>Entropía espectral</b>			
Delta	95,57 ± 0,79	79,60 ± 3,59	81,60 ± 2,38
Theta	94,99 ± 1,58	82,13 ± 4,96	86,20 ± 2,31
Alpha	95,68 ± 0,75	80,96 ± 2,65	85,09 ± 1,31
Beta	93,97 ± 1,09	78,89 ± 5,09	81,07 ± 3,08
Gamma	92,27 ± 0,71	73,43 ± 3,69	76,00 ± 2,06
All	65,16 ± 0,67	58,03 ± 2,21	55,21 ± 6,10

Tabla 4.16: Se muestra el valor medio y la desviación estándar de cada una de las folds del CV, para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 3.c y 3.d. Fuente: elaboración propia.

La Tabla 4.17 muestra los resultados del conjunto de test para las cuatro métricas consideradas. Estos resultados son coherentes con los obtenidos para los casos ante-

riores. La entropía espectral presenta valores más homogéneos en todas las bandas con una menor std comparada con los valores de la coherencia. Además, en ambos casos, los valores de bajas frecuencias (delta, theta y alpha) parecen ser los más representativos para identificar la enfermedad con independencia del estado de activación dopaminérgico. También llama la atención que la entropía espectral tiene un mayor rendimiento en la identificación de controles.

<b>Control-Both</b>				
	Accuracy	Sensitivity	Specificity	Precision
<b>Coherencia</b>				
Delta	93,32 ± 2,31	93,56 ± 2,90	90,48 ± 8,03	99,18 ± 0,68
Theta	94,80 ± 3,65	95,77 ± 3,94	83,17 ± 8,57	98,56 ± 0,76
Alpha	85,07 ± 3,78	84,40 ± 4,55	93,02 ± 11,99	99,36 ± 1,09
Beta	85,69 ± 11,10	86,61 ± 13,09	74,60 ± 13,02	97,78 ± 1,06
Gamma	86,28 ± 7,59	87,49 ± 9,48	71,75 ± 17,65	97,54 ± 1,47
All	95,21 ± 3,03	96,64 ± 3,60	77,97 ± 6,76	98,15 ± 0,52
<b>Entropía espectral</b>				
Delta	81,60 ± 2,38	80,69 ± 2,65	92,54 ± 1,98	99,24 ± 0,19
Theta	86,20 ± 2,31	85,28 ± 2,56	97,31 ± 2,99	99,74 ± 0,28
Alpha	85,09 ± 1,31	84,51 ± 1,55	92,06 ± 2,19	99,23 ± 0,20
Beta	81,07 ± 3,08	79,91 ± 3,54	95,08 ± 5,08	99,50 ± 0,51
Gamma	76,00 ± 2,06	75,83 ± 2,55	77,94 ± 4,33	97,65 ± 0,38
All	55,21 ± 6,10	53,99 ± 7,38	69,84 ± 9,62	95,65 ± 0,72

Tabla 4.17: Resultados de cada uno de los experimentos realizados para las métricas de accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el experimento de controles vs. pacientes en los estados ON y OFF. Los resultados muestran el valor medio y la desviación estándar de cada una de las folds del CV para cada métrica del conjunto de prueba. Fuente: elaboración propia.

Comparando los resultados obtenidos por las diferentes medidas de conectividad, encontramos que estas medidas espectrales han mostrado una mayor capacidad para generalizar los resultados que las medidas definidas en el dominio temporal. Los resultados proporcionados por las medidas definidas en el dominio de las frecuencias indican que las bandas asociadas a bajas frecuencias, son más representativas para identificar la EP. Aunque la coherencia proporciona resultados con valores más altos en las métricas escogidas, tienen una mayor std que los resultados que proporciona la entropía espectral. Sin embargo, se debe notar, que en el caso del modelo 3, que coincide con el modelo que contiene más datos en el conjunto de test, los valores de desviación estándar han disminuido significativamente para la coherencia. En cuanto a los procesos de entrenamiento, aunque los valores de la entropía espectral presenten indicios de

overfitting, el modelo ha mostrado capacidad de generalización de los resultados en los conjuntos de test, indicando que una mayor cantidad de datos de entrenamiento podría producir resultados más precisos con valores más altos en las métricas consideradas.

### 4.3.3. Resultados del estudio de conectividad de los estados dopaminérgicos

#### 4.3.3.1. Medidas definidas en el dominio temporal

A continuación presentamos los resultados del modelo 4, encargado de clasificar a los pacientes con EP en estado ON de los pacientes con EP en estado OFF. En la Tabla 4.18 se muestran los resultados obtenidos para los procesos de entrenamiento, validación y prueba. El conjunto de prueba contiene 2 pacientes en estado ON y dos pacientes en estado OFF. Por ello, evaluaremos los resultados de validación comparándolos con los resultados de test.

	Pre-Post		
	Entrenamiento	Validación	Prueba
Correlación de Pearson	68,85 ± 1,93	44,11 ± 5,91	39,37 ± 14,31
Entropía de Shannon	82,91 ± 0,89	44,06 ± 9,00	48,10 ± 8,37

Tabla 4.18: Se muestra el valor medio y la desviación estándar para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 4.a y 4.b. Fuente: elaboración propia.

Los resultados de entrenamiento y validación de la Tabla 4.18, muestran un proceso de entrenamiento fallido debido a la presencia de overfitting. Los resultados de validación (cerca del 50 %) indican que el modelo no ha sido capaz de extraer las características necesarias de los datos de entrenamiento para poder llevar a cabo la clasificación. Consecuentemente, los resultados obtenidos por el conjunto de test no son representativos, ya que el modelo ha efectuado la clasificación al azar. No obstante los resultados para el conjunto de prueba de las métricas consideradas se proporcionan en la Tabla 4.19.

	Pre-Post			
	Accuracy	Sensitivity	Specificity	Precision
Correlación de Pearson	39,37 ± 14,31	29,52 ± 28,91	49,21 ± 42,90	40,22 ± 18,40
Entropía de Shannon	48,10 ± 8,37	25,40 ± 18,31	70,79 ± 22,96	47,64 ± 22,94

Tabla 4.19: Resultados de las métricas accuracy, sensitivity, specificity y precision del proceso de prueba para los modelos 4.a y 4.b. Se muestra el valor medio y la desviación estándar obtenidas durante el proceso de validación. Fuente: elaboración propia.

La Tabla 4.19 muestra que para ambas medidas de conectividad, los resultados de sensitivity (relacionado con la tasa de pacientes en estado ON acertados) tienen valores más bajos, y los resultados de specificity (relacionado con la tasa de pacientes en estado OFF acertados) proporcionan los valores más altos. Sin embargo, estos resultados no son representativos, ya que la red no llevó a cabo un proceso de aprendizaje adecuado. Debemos de tener en cuenta que las condiciones de entrenamiento y arquitectura de la red se han mantenido en todos los experimentos realizados. Por tanto, las diferencias en el entrenamiento de este modelo son debidas a las características de los datos, esto es, hay menos diferencias de conectividad entre estados de activación dopaminérgica que entre un paciente con EP y un control. Consecuentemente, es necesario una mayor cantidad de datos para que el modelo identifique las alteraciones en la conectividad producidas por la medicación.

#### 4.3.3.2. Medidas definidas en el dominio espectral

A continuación se presentan los resultados de las medidas de conectividad definidas en el dominio espectral, estas son la coherencia y la entropía espectral. La Tabla 4.20 contiene los resultados de los procesos de entrenamiento, validación y prueba para ambas medidas de conectividad en todas las bandas consideradas.

	Pre-Post		
	Entrenamiento	Validación	Prueba
<b>Coherencia</b>			
Delta	68,60 ± 2,67	30,53 ± 7,94	51,90 ± 5,49
Theta	68,62 ± 3,86	30,04 ± 6,28	49,68 ± 12,00
Alpha	69,30 ± 3,01	32,52 ± 9,91	58,73 ± 9,36
Beta	70,98 ± 3,58	35,95 ± 14,23	62,70 ± 8,09
Gamma	70,22 ± 3,97	33,92 ± 13,84	60,63 ± 5,71
All	72,48 ± 2,87	28,33 ± 8,27	52,03 ± 4,08
<b>Entropía espectral</b>			
Delta	77,75 ± 0,52	45,75 ± 5,86	50,95 ± 2,05
Theta	76,67 ± 0,96	43,90 ± 3,16	52,54 ± 4,81
Alpha	77,75 ± 0,39	49,05 ± 6,99	43,65 ± 4,63
Beta	77,24 ± 1,29	44,09 ± 2,72	51,59 ± 5,41
Gamma	79,19 ± 2,19	45,80 ± 5,06	51,75 ± 1,72
All	58,96 ± 0,56	48,21 ± 2,60	49,59 ± 0,90

Tabla 4.20: Se muestra el valor medio y la desviación estándar del proceso de validación cruzada, para la métrica de accuracy en los procesos de entrenamiento validación y prueba para los modelos 4.c y 4.d. Fuente: elaboración propia.

Como podemos observar en la Tabla 4.20, en ambos casos, los procesos de entrenamiento no fueron óptimos, ya que encontramos diferencias significativas entre los procesos de entrenamiento y validación, además de presentar valores cercanos a procesos aleatorios. Aunque la Tabla 4.21 muestre los resultados obtenidos por el resto de las métricas para el conjunto de test, estos no son representativos. Dado que las características de estos experimentos coinciden con las otras experimentaciones, volvemos a concluir que las diferencias producidas en el entrenamiento son debidas a los datos, indicando una mayor semejanza entre pacientes con EP en diferentes estados de activación dopaminérgica, que entre pacientes con EP y controles.

	Pre-Post			
	Accuracy	Sensitivity	Specificity	Precision
<b>Coherencia</b>				
Delta	51,90 ± 5,49	38,73 ± 35,79	65,08 ± 40,89	62,96 ± 25,02
Theta	49,68 ± 12,00	33,02 ± 33,30	66,35 ± 40,83	60,76 ± 38,41
Alpha	58,73 ± 9,36	44,44 ± 16,46	73,02 ± 23,60	70,98 ± 22,68
Beta	62,70 ± 8,09	69,52 ± 42,41	55,87 ± 36,81	71,82 ± 19,55
Gamma	60,63 ± 5,71	70,48 ± 43,58	50,79 ± 33,12	67,31 ± 18,39
All	52,03 ± 4,08	70,79 ± 30,96	33,27 ± 26,88	50,83 ± 3,00
<b>Entropía espectral</b>				
Delta	50,95 ± 2,05	53,65 ± 32,96	48,25 ± 32,91	51,64 ± 4,33
Theta	52,54 ± 4,81	56,51 ± 21,43	48,57 ± 29,24	54,63 ± 7,65
Alpha	43,65 ± 4,63	32,06 ± 21,51	55,24 ± 16,63	39,28 ± 10,12
Beta	51,59 ± 5,41	48,57 ± 15,74	54,60 ± 21,28	53,18 ± 8,51
Gamma	51,75 ± 1,72	40,63 ± 16,63	62,86 ± 17,98	52,89 ± 2,98
All	49,59 ± 0,90	40,38 ± 21,78	58,79 ± 21,27	49,15 ± 1,56

Tabla 4.21: Resultados de las métricas accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el experimento de pacientes en estado ON vs. pacientes en estado OFF. Los resultados muestran el valor medio y la desviación estándar durante el proceso de validación cruzada para el conjunto de prueba. Fuente: elaboración propia.

#### 4.3.4. Generalización de los resultados

Dado que la muestra del conjunto de prueba para los modelos 1 y 2 estaba muy condicionada por la elección de los pacientes, ya que se componía únicamente de 2 controles y 2 pacientes con EP, hemos considerado un nuevo conjunto de test para evaluar la capacidad de generalización de los resultados. Este conjunto de prueba lo componen 13 pacientes con EP bajo la influencia de la medicación, pero no necesariamente en el estado ON y 2 controles seleccionados al azar. Los resultados de las métricas obtenidas para este conjunto externo, para la correlación de Pearson se muestra en la Tabla 4.22.



<b>Correlación de Pearson</b>				
	Accuracy	Sensitivity	Specificity	Precision
Control-Pre	93,08 ± 0,95	95,29 ± 1,19	78,73 ± 0,78	96,68 ± 0,09
Control-Post	92,04 ± 2,75	93,53 ± 4,29	82,38 ± 8,55	97,25 ± 1,32
Control-Both	77,90 ± 1,20	79,42 ± 2,28	65,71 ± 23,80	95,06 ± 2,99

Tabla 4.22: Resultados de las métricas accuracy, sensitivity, specificity y precision para el conjunto de test externo para los modelos 1.a, 2.a y 3.a. Se muestra el valor medio y la desviación estándar durante el procedo de validación cruzada. Fuente: elaboración propia.

Los resultados de la sensibilidad, relacionada con la tasa de aciertos de los pacientes con EP, ha aumentado considerablemente con respecto a los experimentos anteriores, siendo el modelo 1.a el que ofrece mejor rendimiento en la identificación de los pacientes con EP. Sin embargo, los valores de especificidad, relacionados con la tasa de acierto de controles, han disminuido. Dado los resultados anteriores, podemos concluir que los valores de la especificidad están sesgados por la escasez de controles en el conjunto de test externo.

Los resultados de las entropía de Shannon se muestran en la Tabla 4.23. Los resultados de la especificidad muestran que la entropía de Shannon permite una mayor generalización de los resultados. Llama la atención que para esta medida, el modelo 2.b ha mostrado tener un mejor rendimiento para identificar a los pacientes con EP. Sin embargo, comparando los resultados de ambas medidas de conectividad encontramos que los valores de sensibilidad han empeorado considerablemente con respecto a los obtenidos por medio de la correlación de Pearson. Debemos notar que al contar con más datos se ha reducido la std de la medida de la correlación, siendo ahora semejantes en ambas medidas de conectividad.

<b>Entropía de Shannon</b>				
	Accuracy	Sensitivity	Specificity	Precision
Control-Pre	85,63 ± 1,63	84,98 ± 2,09	89,84 ± 5,05	98,21 ± 0,85
Control-Post	87,17 ± 2,55	87,20 ± 3,39	86,98 ± 5,74	97,78 ± 0,91
Control-Both	75,59 ± 2,00	73,63 ± 2,26	91,27 ± 5,24	98,55 ± 0,85

Tabla 4.23: Resultados de las métricas accuracy, sensitivity, specificity y precision para el conjunto de test externo para los modelos 1.b, 2.b y 3.b. Se muestra el valor medio y la desviación estándar durante el procedo de validación cruzada. Fuente: elaboración propia.

En cuanto a las medidas espectrales de los modelos 1.c y 1.d presentadas en la Tabla 4.24, encontramos que la entropía espectral presenta una menor std que la coherencia y ha mostrado tener un gran rendimiento en la tarea de identificación de controles.

Debemos notar que los resultados asociados a los experimentos de de bajas frecuencias (delta y theta) han mostrado ser más representativos para identificar la enfermedad.

<b>Control-Pre</b>				
	Accuracy	Sensitivity	Specificity	Precision
<b>Coherencia</b>				
Delta	89,29 ± 2,72	88,86 ± 3,80	92,06 ± 7,65	98,70 ± 1,24
Theta	85,57 ± 7,18	84,79 ± 9,22	90,63 ± 10,74	98,47 ± 1,67
Alpha	76,91 ± 9,21	74,73 ± 12,65	91,11 ± 15,03	98,59 ± 2,20
Beta	76,34 ± 14,68	76,07 ± 20,17	78,10 ± 23,19	96,66 ± 3,22
Gamma	84,04 ± 12,75	85,13 ± 17,02	76,98 ± 19,79	96,54 ± 2,81
All	85,85 ± 9,66	85,63 ± 12,50	87,33 ± 12,38	98,00 ± 1,80
<b>Entropía espectral</b>				
Delta	77,71 ± 2,37	75,38 ± 2,90	92,86 ± 2,56	98,57 ± 0,49
Theta	78,10 ± 2,87	75,21 ± 3,65	96,83 ± 3,05	99,38 ± 0,58
Alpha	79,03 ± 1,13	77,22 ± 1,56	90,79 ± 2,73	98,21 ± 0,49
Beta	77,50 ± 2,37	74,73 ± 3,30	95,56 ± 4,51	99,13 ± 0,87
Gamma	74,90 ± 3,33	72,04 ± 3,85	93,49 ± 3,53	98,63 ± 0,72
All	55,58 ± 4,68	53,26 ± 6,71	70,67 ± 8,64	92,34 ± 1,10

Tabla 4.24: Resultados de las métricas accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el conjunto de test externo para los modelos 1.c y 1.d. Los resultados muestran el valor medio y la desviación. Fuente: elaboración propia.

La Tabla 4.25 también muestra grandes fluctuaciones en las predicciones de la coherencia con respecto a las de la entropía espectral. Sin embargo, los valores de std obtenidos para la coherencia de este modelo son menores que en el caso del modelo 1.c. También podemos notar que la banda delta sigue siendo la más representativa para identificar la enfermedad, aunque no es el caso de la entropía espectral, cuya banda más representativa es la alpha.

Finalmente, los resultados de los modelos 3.c y 3.d se muestran en la Tabla 4.26. Para la medida de la coherencia volvemos a observar grandes fluctuaciones en las predicciones asociadas con las bandas alpha, beta y gamma y valores más homogéneos en los resultados de la entropía espectral. En este caso, los valores de sensibilidad para la entropía espectral indican un bajo rendimiento en la tarea de identificación de características de la EP, al mostrar valores cercanos a los obtenidos por un proceso aleatorio. Sin embargo, muestra un gran rendimiento en la tarea de identificación de controles, ya que los valores obtenidos para la especificidad son superiores al 90 % en las bandas delta, theta, alpha y beta.

Comparando los resultados de todas las medidas espectrales observamos que el modelo 3 ha identificado peor a los pacientes con EP. Dado que este modelo ha sido en-

<b>Control-Post</b>				
	<b>Accuracy</b>	<b>Sensitivity</b>	<b>Specificity</b>	<b>Precision</b>
<b>Coherencia</b>				
Delta	89,63 ± 4,11	90,28 ± 5,95	85,40 ± 8,35	97,67 ± 1,31
Theta	82,75 ± 7,97	81,78 ± 10,88	89,05 ± 11,94	98,24 ± 1,83
Alpha	80,63 ± 7,58	78,39 ± 9,91	95,24 ± 7,94	99,24 ± 1,20
Beta	71,49 ± 9,50	67,89 ± 11,57	94,92 ± 5,71	98,98 ± 1,03
Gamma	87,70 ± 8,25	89,16 ± 11,29	78,25 ± 12,02	96,63 ± 1,80
All	73,71 ± 5,57	69,90 ± 6,45	98,48 ± 0,61	99,67 ± 0,13
<b>Entropía espectral</b>				
Delta	78,60 ± 2,07	76,65 ± 2,56	91,27 ± 3,62	98,29 ± 0,68
Theta	80,38 ± 1,15	77,75 ± 1,73	97,46 ± 2,68	99,51 ± 0,50
Alpha	81,97 ± 1,26	80,51 ± 1,88	91,43 ± 3,56	98,40 ± 0,62
Beta	80,59 ± 1,36	78,36 ± 1,93	95,08 ± 4,33	99,06 ± 0,81
Gamma	77,40 ± 1,51	74,87 ± 2,23	93,81 ± 4,58	98,77 ± 0,87
All	58,05 ± 3,84	56,42 ± 5,01	68,63 ± 4,54	92,15 ± 0,60

Tabla 4.25: Resultados de las métricas accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el conjunto de test externo para los modelos 2.c y 2.d. Los resultados muestran el valor medio y la desviación. Fuente: elaboración propia.

<b>Control-Both</b>				
	<b>Accuracy</b>	<b>Sensitivity</b>	<b>Specificity</b>	<b>Precision</b>
<b>Coherencia</b>				
Delta	73,37 ± 3,30	71,23 ± 4,62	90,47 ± 8,03	98,45 ± 1,28
Theta	77,94 ± 1,27	77,28 ± 2,35	83,17 ± 8,57	97,40 ± 1,28
Alpha	63,09 ± 9,17	59,35 ± 11,75	93,02 ± 11,99	98,91 ± 1,72
Beta	72,84 ± 13,46	72,62 ± 16,64	74,60 ± 13,02	96,21 ± 1,80
Gamma	72,91 ± 10,68	73,06 ± 13,93	71,75 ± 17,65	95,86 ± 2,38
All	79,38 ± 0,40	79,55 ± 1,17	77,97 ± 6,76	96,68 ± 0,93
<b>Entropía espectral</b>				
Delta	66,46 ± 1,60	63,19 ± 1,92	92,54 ± 1,98	98,55 ± 0,36
Theta	69,66 ± 2,93	66,21 ± 3,53	97,30 ± 2,99	99,51 ± 0,52
Alpha	68,32 ± 0,65	65,36 ± 0,94	92,06 ± 2,19	98,51 ± 0,38
Beta	68,29 ± 2,90	64,94 ± 3,51	95,08 ± 5,08	99,09 ± 0,94
Gamma	67,60 ± 1,96	66,31 ± 2,67	77,94 ± 4,33	96,03 ± 0,61
All	52,73 ± 5,40	50,59 ± 7,26	69,84 ± 9,62	93,22 ± 1,03

Tabla 4.26: Resultados de las métricas accuracy, sensitivity, specificity y precision para la coherencia y la entropía espectral en el conjunto de test externo para los modelos 3.c y 3.d. Los resultados muestran el valor medio y la desviación durante el proceso de validación cruzada. Fuente: elaboración propia.

trenado con pacientes en diferentes estados dopaminérgicos, nos lleva a pensar que las diferencias de conectividad producidas por la medicación, han dificultado la extracción de elementos característicos de la EP, y por ello encontramos una disminución del rendimiento del modelo. Llama la atención que en todos los experimentos que han utilizado la coherencia, las bandas que han mostrado ser más representativas para identificar la enfermedad son las bandas asociadas a bajas frecuencias. En particular, el modelo que ha mostramos mayor precisión para identificar la EP ha sido el proceso 2.c.delta.

Al comparar los resultados de todas las medidas de conectividad, encontramos que la medida que ha proporcionado mejores resultados de exactitud para clasificar pacientes con EP de controles es la correlación de Pearson con un valor de  $93,08 \pm 0,95$ . Particularmente, esta medida ofrece un gran rendimiento en la identificación de pacientes con EP pero tiene una tasa de aciertos de controles menor que otras medidas.

#### **4.4. Aplicación de las técnicas de PLN para el diagnóstico de la EP**

Los modelos del campo del PLN se desarrollaron con el fin de analizar datos no estructurados, en particular, su aplicación estaba enfocada al análisis de texto. Sin embargo, dado que cada una de las señales que compone el EEG es una serie temporal, siendo esta un tipo de dato secuencial, la hipótesis de la que parte este estudio para aplicar este tipo de técnicas (*hipótesis 5*), es que existe una equivalencia entre texto y EEG, de modo que podemos implementar estos métodos con resultados satisfactorios. Con esta analogía, interpretamos cada señal del EEG como una frase de un texto para relacionar fragmentos del EEG de igual modo que estos modelos relacionan palabras de un texto.

Para desarrollar un biomarcador de la enfermedad mediante un modelo basado en BERT, uno de los modelos más representativos dentro del campo del PLN, debemos implementar un módulo encargado de realizar una clasificación. Debemos notar, que estos modelos requieren de una gran cantidad de datos para realizar el entrenamiento y para obtener resultados que puedan generalizarse a otros datos. De modo que los participantes incluidos en este estudio son 80 pacientes con EP 24 controles, cuyos datos demográficos y clínicos se encuentran en la Tabla [4.27](#).

Teniendo en cuenta las 3 pruebas electroencefalográficas que se realiza a cada sujeto, esto es una prueba de reposos y dos de activación motora (cada una asociada a la activación de un hemisferio), contamos con un total de 240 pruebas para los pacientes con EP y con un total de 72 pruebas para los controles. Aunque esta cantidad de datos puede parecer pequeña para entrenar un modelo basado en la arquitectura BERT, al-

	Controles	Pacientes con EP
N(Hombres)	24(19)	80(48)
Edad en años	58,12 ± 6,91	63,89 ± 9.21
Duración de la enfermedad en años	-	7,21 ± 4,54
HY	-	2,99 ± 1,35

Tabla 4.27: Datos demográficos y clínicos de los participantes incluidos en el estudio de diagnóstico de la EP mediante un modelo BERT. Se muestra el valor medio y la desviación estándar de cada una de las medidas consideradas. Fuente: [305].

gunos estudios han demostrado su buen rendimiento con conjuntos de datos pequeños. Entre ellos encontramos [306], que obtiene precisiones superiores al 80 % con conjuntos de datos de 10 muestras por clase, y [307], que obtiene métricas cercanas al 80 %, con un conjunto de datos de unas 100 instancias por clase.

Como mencionamos en el Capítulo 3, realizaremos 4 experimentos que nos permiten, por un lado, evaluar la cantidad de información que contienen los electrodos, esto es, realizando el entrenamiento con un conjunto de 64 electrodos y con un conjunto de 28 electrodos en el que prescindimos de los canales periféricos. Por otro lado, se evalúa la información obtenida por las pruebas electroencefalográficas mediante 2 estrategias de entrenamiento, una de ellas restringida a las pruebas de activación motora y la otra añadiendo la prueba de reposo. Los modelos asociados a los 4 experimentos, que se obtienen de la combinación de estas opciones, tienen la misma arquitectura para favorecer la comparación de resultados. Los experimentos junto con su especificación se resume a continuación:

- *64\_all*: utiliza las 64 señales del EEG incluyendo todas las pruebas registradas a los pacientes.
- *64\_FT*: utiliza las 64 señales del EEG incluyendo las pruebas motoras (FTD y FTI) de los participantes.
- *28\_all*: utiliza 28 señales del EEG con todas las pruebas registradas a los pacientes.
- *28\_FT*: utiliza las 28 señales del EEG con las pruebas motoras (FTD y FTI) registradas a los participantes.

Aunque la arquitectura sea la misma en los 4 experimentos, debemos notar que el número de parámetros del modelo disminuye al considerar una menor cantidad de canales de entrada (estos son los casos de 28 electrodos) debido a la dimensionalidad de las matrices consideradas. Los resultados de este estudio pueden verse en [305].

#### 4.4.1. Clasificador de la EP basado en BERT

La arquitectura del modelo se divide en dos módulos. El primero consiste en un modelo BERT sin capa de embedding (ya que nuestros datos tienen una representación numérica), encargado de relacionar las diferentes instancias. Después, encontramos un módulo de clasificación, encargado de asignar a cada dato de entrada una clase, siendo estas “PD” y “No\_PD”.

- **Módulo BERT.** Tiene 6 codificadores Transformer con 64 neuronas y 4 cabezas de atención. Entre cada Transformer se establece una capa feed-forward con 1.536 neuronas, con función Gaussiana GELU (por sus siglas en inglés “Gaussian Error Linear Units”) y un dropout de 0,3.
- **Módulo de clasificación.** Es un perceptrón multicapa con una capa de entrada de 64 neuronas, una capa oculta de 268 neuronas, seguida de una salida SoftMax sobre dos clases que representan las posibles etiquetas en el problema de diagnóstico de la EP. La arquitectura del modelo de 64 canales se puede visualizar en la Fig. 4.14. El resultado del modelo es una clasificación binaria que se encarga de identificar para cada segmento del EEG si tiene o no EP.

En cuanto a la segmentación del dataset, este se divide en entrenamiento, validación y prueba. Los subconjuntos de entrenamiento y validación se encuentran en la etapa de entrenamiento del modelo, mientras que el subconjunto de prueba (ejemplos nunca vistos por el modelo durante el entrenamiento) se utiliza para evaluar la capacidad de generalización de los resultados. Para ello, tras barajar los datos, separamos el 20% del dataset para el conjunto de prueba. Con el resto de datos se utilizó un 5-fold CV, separando el dataset restante en 80% entrenamiento y 20% validación. Los diferentes subconjuntos se eligieron de forma aleatoria en cuanto a individuos, pero siempre manteniendo el porcentaje de pacientes y controles. La división en conjuntos de entrenamiento/validación/prueba se llevó a cabo garantizando la independencia del paciente, y luego se realizó la división en ventanas.

También debemos notar que las clases entre pacientes con EP y controles no están balanceadas. Aunque los modelos basados en BERT o Transformer pueden trabajar con clases no equilibradas [308], esta doble validación (por medio del CV y conjunto de prueba) permite eliminar el sesgo producido por la elección de los datos e identificar los fallos durante el proceso de entrenamiento.

En cuanto a los tiempos de entrenamiento, el experimento que aborda los 28 canales se entrenó durante cinco épocas en 3 días, 12 horas y 29 minutos incluyendo todas las pruebas de los EEGs y 1 día y 12 horas eliminando las pruebas de reposo. En el caso

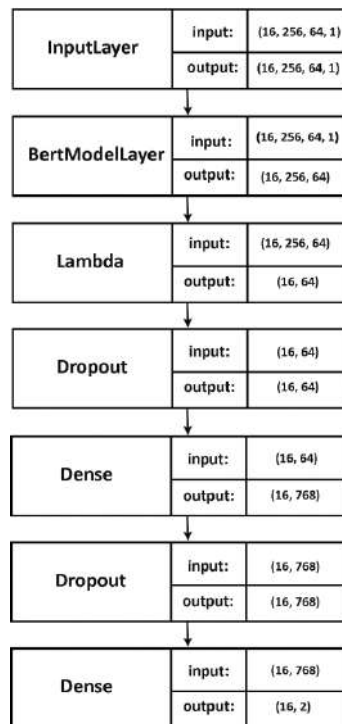


Figura 4.14: Representa la arquitectura del modelo basado en BERT para el diagnóstico de la EP. Para cada capa se muestra la dimensión del tensor de entrada y salida indicando la dirección en la que se producen los procesos. Fuente: [305].

de los 64 canales, se entrenó durante cinco épocas en 3 días, 16 horas y 48 minutos en el primer caso y 1 día y 12 horas en el segundo. Los parámetros de entrenamiento del modelo fueron de 583.842 para el experimento con 28 canales, y de 1.374.978 para el experimento de 64 canales. Debemos notar que la eliminación de la capa embedding reduce considerablemente el tamaño de los modelos y contribuye a una reducción del tiempo de entrenamiento.

#### 4.4.2. Resultados en el diagnóstico de la EP

Los resultados ofrecidos por el modelo se evalúan utilizando las 4 métricas relacionadas con el campo de la medicina y que han sido especificadas anteriormente (accuracy, specificity, sensitivity y precision), favoreciendo la evaluación e interpretación de los resultados. En este estudio, se considera un falso positivo a una persona sana diagnosticada erróneamente con EP, y un falso negativo a un paciente con EP diagnosticada como sana.

Después de entrenar los modelos de 28 canales con ambos conjuntos de entrenamiento (con y sin EEGs en estado de reposo) durante cinco épocas, obtuvimos los re-

sultados que se muestran en la Tabla 4.28. En esta tabla se muestran los resultados de las cuatro métricas para los procesos de entrenamiento, validación y prueba para cada uno de los experimentos de 28 canales.

Model	Training without resting states			Training with resting states		
	Train	Validation	Test	Train	Validation	Test
Accuracy	75.98 ± 6.59	65.67 ± 5.16	70.16 ± 10.01	72.30 ± 4.71	68.75 ± 8.66	71.30 ± 5.58
Specificity	52.05 ± 30.01	61.89 ± 27.47	30.53 ± 25.06	21.65 ± 24.92	19.92 ± 20.64	18.08 ± 20.64
Sensitivity	87.47 ± 10.05	82.73 ± 11.52	90.00 ± 10.00	96.58 ± 44.13	92.20 ± 44.16	98.00 ± 44.72
Precision	79.22 ± 20.57	71.20 ± 12.65	72.15 ± 19.44	72.05 ± 17.27	70.65 ± 19.19	70.47 ± 17.61

Tabla 4.28: Evaluación de los resultados de los modelos de 28 canales con ambas estrategias de entrenamiento, esto es, por un lado, con todas las pruebas electroencefalográficas y, por otro lado, únicamente con las pruebas asociadas a la activación motora. Fuente: [305].

En el experimento de 28 canales considerando todos los EEGs, podemos ver que llama la atención los resultados de dos métricas. Por un lado, la sensibilidad, encargada de proporcionar la tasa de verdaderos positivos, con valores superiores al 90 % en todas las etapas de experimentación (es decir, entrenamiento, validación y prueba). Sin embargo, presenta valores de desviación estándar muy elevados, en torno al 44 % en todos los casos. Por otro lado, la especificidad, encargada de proporcionar la tasa de verdaderos negativos, tiene valores muy bajos, en torno al 20 %.

Cuando se consideran únicamente los EEGs relacionados con las pruebas de activación motora, podemos ver que la variabilidad de los resultados para la sensibilidad y especificidad es menor aunque sigue siendo significativamente alta. Tenemos que tener en cuenta que en esta experimentación tenemos menos datos que en el caso de introducir las pruebas de reposo, aunque se mantiene la diferencia entre las clases. Este resultado afecta especialmente a la métrica de especificidad, como podemos ver en la Tabla 4.28, que además de tener un valor significativamente bajo, su desviación estándar presenta valores elevados, en torno al 30 %. Cuando la prueba de reposo no se utiliza, no observamos diferencias significativas entre los resultados de las métricas de precisión y exactitud.

Ahora, entrenamos un modelo de 64 canales con los mismos datos y condiciones de entrenamiento, esto es durante cinco épocas. La Tabla 4.29 recopila la información de las cuatro métricas para los procesos de entrenamiento, validación y prueba.



Model	Training without resting states			Training with resting states		
	Train	Validation	Test	Train	Validation	Test
Accuracy	93.45 ± 2.82	84.67 ± 3.79	86.41 ± 7.08	76.04 ± 5.85	68.41 ± 10.08	67.93 ± 5.59
Specificity	93.08 ± 1.66	84.94 ± 11.52	96.00 ± 8.94	54.11 ± 9.47	44.26 ± 22.28	36.00 ± 8.93
Sensitivity	93.64 ± 4.21	84.56 ± 9.00	81.63 ± 11.58	86.54 ± 1.63	79.98 ± 9.33	83.96 ± 10.99
Precision	96.59 ± 26.93	92.15 ± 23.90	97.61 ± 23.72	79.77 ± 19.69	75.00 ± 19.37	72.35 ± 5.59

Tabla 4.29: Evaluación de los resultados de los modelos de 64 canales con ambas estrategias de entrenamiento, esto es, por un lado, con todas las pruebas electroencefalográficas y, por otro lado, únicamente con las pruebas asociadas a la activación motora. Fuente: [305].

Para el experimento que considera únicamente los EEGs relacionados con activaciones motoras, podemos observar que tanto la exactitud como la precisión tienen valores altos en las diferentes etapas de la experimentación. Sin embargo, cabe señalar que durante el proceso de validación cruzada, para los FP, hay una iteración que contiene valores muy diferentes al resto de iteraciones y por tanto estos resultados alteran las medidas de las métricas que contienen este valor.

En el caso del experimento que incluyen todos los EEGs, sólo los valores de FN han mostrado cierta variabilidad en las iteraciones del CV, aunque mucho menos que en los experimentos anteriores de 28 canales. Esto se refleja en los valores obtenidos de las métricas y su desviación estándar, donde aún prevalecen valores bajos de la métrica de especificidad con una alta variabilidad entre los procesos de entrenamiento.

## 4.5. Contribuciones derivadas de esta tesis

### 4.5.1. Artículos científicos

Los resultados de esta tesis han desembocado en 6 artículos científicos, de los cuales 5 han sido publicados y uno se encuentra en desarrollo. Los artículos son:

- “Exploring functional connectivity in Parkinson’s EEGs via graph theory across different dopaminergic-activation stages”. Maitin, A.M.; García-Tejedor, Á.J.; Arroyo, A.; Romero, J.P. Computer methods and programs in biomedicine. Submitted JCR (Computer science, interdisciplinary applications) 2022: Q1.

- “EEGGraph: An open-source Python library for modeling electroencephalograms using graphs”. Maitin, A.M.; Nogales, A.; Chazarra, P.; García-Tejedor, Á.J. *Neurocomputing*, 519, 127-134 (2023). DOI: [10.1016/j.neucom.2022.11.050](https://doi.org/10.1016/j.neucom.2022.11.050). JCR (Artificial Intelligence) 2022: Q2.
- “BERT learns from electroencephalograms about Parkinson’s Disease: Transformer-based models for aid diagnosis”. Nogales, A., García-Tejedor, Á. J., Maitín, A. M., Pérez-Morales, A., Del Castillo, M. D., and Romero, J. P. *IEEE Access*, 10, 101672 (2022). DOI: [10.1109/ACCESS.2022.3201843](https://doi.org/10.1109/ACCESS.2022.3201843). JCR (Computer Science) 2022: Q2.
- “Survey of Machine Learning Techniques in the Analysis of EEG Signals for Parkinson’s Disease: A Systematic Review”. Maitin, A.M.; Romero Muñoz, J.P.; García-Tejedor, Á.J. *Applied Sciences*, 12(14), 6967 (2022). DOI: [10.3390/app12146967](https://doi.org/10.3390/app12146967). JCR (Engineering, Multidisciplinary - Scie) 2022: Q2.
- “Time Series Analysis Applied to EEG Shows Increased Global Connectivity during Motor Activation Detected in PD Patients Compared to Controls”. Maitín, A., Perez-zan, R., Herráez-Aguilar, D., Serrano, I.J., Castillo, M.D., Arroyo, A., Andreo, J. and Romero, J.P. *Applied Sciences* 11(1), 15 (2021). DOI: [10.3390/app11010015](https://doi.org/10.3390/app11010015). JCR (Engineering, Multidisciplinary - Scie) 2021: Q2.
- “Machine Learning Approaches for Detecting Parkinson’s Disease from EEG Analysis: A Systematic Review”. Maitín, A. M., García-Tejedor, A. J. and Romero Muñoz, J. P. *Applied Sciences* 10(23), 8662-9683 (2020). DOI: [10.3390/app10238662](https://doi.org/10.3390/app10238662). JCR (Engineering, Multidisciplinary - Scie) 2020: Q2.

#### 4.5.2. Conferencias y congresos

Los congresos y conferencias en los que se ha difundido el trabajo realizado son:

- “Evaluation of the effect of non-invasive therapies in Parkinson’s Disease patients through the study of functional connectivity and Deep Learning techniques. (Preliminary results)”. Maitin, A., Álvarez de Toledo, I., Arroyo, A., Serrano, I., del Castillo, M., Romero, J., and García-Tejedor, A. *Movement Disorders*. Estado: Aceptado.
- “Biomarker based on natural language processing methods to detect the stages of Parkinson’s disease through electroencephalography”. Maitin, A., Morales, J., Romero-Azpitarte, S., Romero, J., Arroyo Ferrer, A., González Zamorano, Y., Sánchez Cuesta, F., and García-Tejedor, A. *Movement Disorders* 37 (suppl 2), 107 (2022). DOI: [10.1002/mds.29223](https://doi.org/10.1002/mds.29223). JCR (Clinical Neurology - Scie) 2021: Q1.

- “Time Series Analysis Applied to EEG Shows Increased Global Connectivity During Motor Activation in PD Patients Compared to Controls”. Herráez, D., Maitin, A., Perezzan, R., Castillo, M.D., Serrano, I.J., Arroyo, A., Andreo, J. and Romero, J. P. *Movement Disorders Journal* 35(S1), 675 pp 300 (2020) DOI: [10.1002/mds.28268](https://doi.org/10.1002/mds.28268). JCR (Clinical Neurology - Scie) 2020: Q1.
- “Clasificación automática de los niveles de progresión de la enfermedad de Parkinson por medio de técnicas de Deep Learning. (Resultados Preliminares)”. Maitín López, A. M., García Tejedor, A. J., Arroyo Ferrer, A., Castillo, M. D., Serrano, I., Rocón, E., Romero Muñoz, J. P. LXXII Reunión Anual de la Sociedad Española de Neurología, Nov 23-Dic 3, 2020.
- “Application of Deep Neuronal Networks for the extraction of spectral characteristics in patients with Parkinson’s disease”. Maitin, A., García-Tejedor, Á. J., Castillo, M., Rocón, E., Serrano, I., Arroyo, A., Andreo, J. and Romero, J. P. *Movement Disorders Journal* 35(S1), 681 pp 305 (2020) DOI: [10.1002/mds.28268](https://doi.org/10.1002/mds.28268). JCR (Clinical Neurology - Scie) 2020: Q1.
- “Time Series Analysis Applied to EEG Shows Increased Global Connectivity During Motor Activation in PD Patients Compared to Controls”. Herráez, D., Maitin, A., Perezzan, R., Castillo, M.D., Serrano, I.J., Arroyo, A., Andreo, J. and Romero, J. P. *Movement Disorders Journal* 35(S1), 675 pp 300 (2020) DOI: [10.1002/mds.28268](https://doi.org/10.1002/mds.28268). JCR (Clinical Neurology - Scie) 2020: Q1.
- “Asimetría del efecto de la levodopa en el análisis de EEG para distintos tipos de entropía”. Maitin, A. M., Andreo, J., Herráez, D., Perezzan, R., Arroyo, A., Serrano, I., Castillo, M. D., Rocón, E., Romero, J. P. LXXI Reunión Anual de la Sociedad Española de Neurología, Nov 19-23, 2019.



## Capítulo 5

# Discusión

### 5.1. Estudio mediante análisis de series temporales

Aunque existen estudios electrofisiológicos en la EP, desde cambios en la actividad de fondo del EEG hasta complejas alteraciones en la conectividad funcional a nivel cortico-subcortical y cortico-cortical [246, 252, 309, 310], no hemos encontrado ninguno que utilice el análisis de series temporales para detectar un aumento en la conectividad durante las tareas de activación motora. Esto nos ha motivado a realizar este estudio, verificar así la *hipótesis 1* y proceder al desarrollo del *objetivo 1*, el cual, no obstante, está sometido a ciertas limitaciones:

- En primer lugar, el tamaño de la muestra empleada pudiera considerarse insuficiente para detectar diferencias estadísticamente significativas. No obstante, hemos sido capaces de detectar estas diferencias entre la conectividad de los grupos de estudio, aunque los resultados mostraron una gran dispersión dificultando su aplicación a individuos particulares en el ámbito clínico.
- En segundo lugar, se encuentran las limitaciones asociadas al EEG, tales como la dificultad de detectar la actividad neuronal lejos de la superficie del cuero cabelludo o la relación señal/ruido. Esta última cuestión nos ha hecho eliminar sistemáticamente para todos los participantes varios electrodos situados en la periferia del cerebro para minimizar el ruido en los análisis realizados. Además, debemos tener en cuenta que el proceso de limpieza influye en los resultados de las medidas, pudiendo eliminar componentes relevantes de la EP y, por tanto, minimizando las diferencias entre ambos grupos.
- En tercer lugar, algunas de las diferencias en la conectividad entre los pacientes con EP y los controles fueron significativas, pero la gran dispersión derivada de

la heterogeneidad de los participantes al realizar las pruebas impide distinguir a los pacientes con EP de forma individual. Por lo tanto, se necesitan técnicas más refinadas de procesamiento de señales para poder desarrollar un biomarcador que pueda ser aplicado como prueba de apoyo al diagnóstico en la práctica clínica diaria. Debemos notar que un registro más grande de participantes podría ayudar a determinar diferencias significativas, aunque seguiríamos contando con una gran dispersión en las medidas que impide la detección individual.

### 5.1.1. Estudio comparativo entre pacientes con EP y controles

En el conjunto de pacientes que hemos considerado, se hace evidente la independencia parcial entre hemisferios cerebrales. Esta independencia, medida con la Divergencia de Kullback-Leibler, parece ser mayor en los controles en comparación con los pacientes con EP para la prueba de resting, pero no es estadísticamente significativa ( $p = 0,15$ ). Para las medidas del coeficiente de correlación de la entropía espectral y la coherencia, no se observaron diferencias significativas entre ambos grupos, lo que evidencia que su nivel de conectividad es muy similar en estado de reposo.

Con respecto a las pruebas motoras, se observa un aumento notable de la correlación de la entropía espectral cuando la corteza motora se activa para realizar las tareas de FT en los pacientes con EP, indicando una posible sobreconectividad durante la ejecución motora. Dado que las acciones surgen de actividades cerebrales locales, temporalmente coordinadas, se esperan cambios en la conectividad producidos por la activación motora o cognitiva en comparación con el estado de reposo. Estos resultados fueron corroborados por la coherencia, que presentó cambios opuestos en respuesta a la activación motora sin mostrar influencias por el hemisferio activado en la tarea. También existen diferencias significativas entre los cambios de independencia inter-hemisférica de ambos grupos para ambas pruebas motoras, mostrando un aumento de la independencia en los controles y una disminución en los sujetos con EP, indicando que la conectividad entre ambos hemisferios puede haber aumentado durante las tareas motoras para los últimos. Esto requiere de una participación global de las redes biológicas y podría explicar un menor rendimiento de los pacientes enfermos en las actividades motoras. Este comportamiento fue general en ambas manos, sin mostrar diferencias entre ellas. El aumento de la sincronía inter-hemisférica entre las cortezas sensoriomotoras se ha evidenciado previamente en modelos animales experimentales de EP [253] y también se ha asociado con disfunciones cognitivas en estudios de MEG [252].

Cabe señalar que un aumento en los valores del coeficiente de correlación o la coherencia pueden deberse a las conexiones anatómicas o al acoplamiento funcional entre las regiones del cerebro [311, 312]. Los resultados obtenidos sugieren que puede ser

necesaria una mayor conectividad y una red de activación más extendida para cumplir con las demandas de una tarea motora en los cerebros de los pacientes con EP, reduciendo su eficiencia. Esto se puede traducir en un deterioro del rendimiento cognitivo y motor, en concordancia con lo descrito en [313].

### 5.1.2. Afectación de la levodopa

La levodopa es un precursor de la dopamina, por lo que la ingesta de este medicamento puede alterar algunos de los análisis considerados y por tanto su efecto debe tenerse en cuenta en la búsqueda del biomarcador. Los resultados mostraron que la administración de este fármaco frena el aumento del coeficiente de correlación producido con la activación motora en la mano izquierda. Aunque los valores siguen alterados, estos se acercan más a los observados en los controles. Este efecto asimétrico ha sido evidenciado previamente [314], y merece ser explorado en un estudio futuro.

Cuando se comparan los resultados de la coherencia antes y después de la administración de la levodopa, se observó que el fármaco aumentó estos valores en todo el espectro de frecuencias considerado para la prueba de reposo, tal y como se ha informado en estudios anteriores [315]. La ingesta de levodopa se ha asociado principalmente con aumentos de potencia en los ritmos alpha y beta en las áreas centro-parietales [316], pero otros estudios son bastante contradictorios y muestran que la conectividad del EEG en estado de reposo entre las regiones de la red motora tras la administración de levodopa no cambian significativamente [317]. Durante la activación motora, el fármaco parece evitar el incremento de la potencia en los sujetos con EP, permaneciendo muy similar al estado de reposo. Los efectos de la levodopa sobre la conectividad podrían ser causados por las amplias proyecciones dopaminérgicas que ascienden desde el área tegmental ventral y el sistema nervioso central, que pueden ser importantes para determinar el acoplamiento cortico-cortical durante las tareas ejecutoras [315, 318].

Finalmente, cuando se analiza la “independencia” inter-hemisférica a través de la KLD, se pudo observar que la ingesta de levodopa no cambió la independencia del procesamiento de la información en ambos lados para los pacientes con EP en estado de reposo. Sin embargo, durante la activación motora, disminuyó el aumento de conectividad descrito previamente en los sujetos con EP.

Estos resultados son coherentes con los obtenidos utilizando fMRI (por sus siglas en inglés “functional Magnetic Resonance Imaging”) en estado de reposo, una de las técnicas más utilizadas para estudiar la conectividad. Los sujetos con EP sin medicación, comparados con los controles, muestran una reorganización de la red a gran escala en la fMRI, con un aumento de la conectividad en las áreas sensoriomotoras corticales [309] y la terapia dopaminérgica normalizó este aumento de conectividad [319].

Con este estudio verificamos la *hipótesis 1* aunque no conseguimos desarrollar el biomarcador, tal y como se propuso en el *objetivo 1*. Sin embargo, este estudio nos ha servido de guía para profundizar en la comprensión de las señales del EEG y en las medidas que son sensibles a alteraciones en las señales, proporcionando información acerca de la conectividad.

## 5.2. Análisis bibliográfico

Verificamos la *hipótesis 2* por medio de dos análisis bibliográficos, alcanzando así el *objetivo 2*.

De forma general, los resultados de las revisiones mostraron que el número de artículos relacionados con el tema de investigación de esta tesis ha ido aumentando en los últimos años, en paralelo al desarrollo de las técnicas de ML aplicadas para el análisis del EEG. Es posible que este aumento venga causado por la mayor cantidad de datos disponibles y por el crecimiento de la potencia computacional, que permite el uso de modelos cada vez más complejos y potentes.

Hemos comprobado que el protocolo de limpieza es innecesario cuando se utilizan técnicas de ML, ya que estas no afectaron al rendimiento de los modelos. En la búsqueda de un biomarcador de la enfermedad basado en EEGs, esto se hace imprescindible ya que evitamos pre-procesar las señales aumentando la rapidez de la prueba sin perder eficacia en el diagnóstico ni condicionar los resultados. Sin embargo, tanto las características extraídas como la arquitectura del modelo, son determinantes para obtener un mejor rendimiento en el problema de clasificación. La combinación de estos factores se puede apreciar en los artículos [261, 262], en los cuales se estudiaron los cambios de rendimiento al variar los parámetros de entrada de la red y los parámetros del modelo respectivamente.

Para obtener un valor de referencia para nuestros estudios, filtramos los resultados de los artículos en función de las buenas prácticas que hemos recogido en cada uno de los puntos considerados. Dado que el accuracy fue la métrica común, fue la que consideramos para el problema de diagnóstico de la EP obteniendo un valor de referencia de  $97,35 \pm 3,46\%$ , con un valor medio de sensitivity de  $96,36 \pm 5,00\%$ . Estos resultados fueron obtenidos para registros con diferentes canales. La baja variación en el valor proporcionado por las métricas conduce a pensar que el número de canales no influyó en los resultados de la clasificación.

Aunque los modelos neuronales utilizados para obtener el diagnóstico de la EP son cada vez más complejos, las técnicas utilizadas se limitan al uso de CNN y modelos híbridos basados en CNN+RNN. En estos estudios predominan las ANN y modelos no-



neuronales en igual proporción.

### 5.3. Conectividad funcional mediante teoría de grafos

Algunas de las medidas incluidas en este estudio coinciden con las utilizadas en el análisis mediante series temporales, por lo que ya han mostrado ser sensibles a las alteraciones en los patrones de conectividad. Por ello, estas medidas nos permitirán verificar la *hipótesis 3* de esta tesis, a fin de obtener el biomarcador que buscamos en el *objetivo 3*. Para obtener un biomarcador robusto a las alteraciones dopaminérgicas, desarrollamos 4 modelos con la misma arquitectura para combinar en clasificadores binarios los pacientes con EP en estados OFF y ON y controles. Las combinaciones consideradas se resumen a continuación:

- **Modelo 1: Control-Pre.** Utiliza los EEGs de controles y pacientes con EP en estado OFF.
- **Modelo 2: Control-Post.** Utiliza los EEGs de controles y pacientes con EP en estado ON.
- **Modelo 3: Control-Both.** Utiliza los EEGs de controles y la mitad de pacientes con EP en estado ON y la mitad de pacientes en estado OFF.
- **Modelo 4: Pre-Post.** Utiliza los EEGs de pacientes con EP en estado ON y los EEGs de pacientes en estado OFF.

Para cada uno de los modelos se estudiaron 4 medidas de conectividad, dos de ellas definidas en el dominio temporal de la señal (la correlación de Pearson y la entropía de Shannon), y dos de ellas definidas en el dominio espectral (la coherencia y la entropía espectral). Debemos notar que este estudio también considera medidas que relacionan pares de electrodos (correlación de Pearson y coherencia) y medidas que proporcionan información acerca de un único canal (entropía de Shannon y entropía espectral).

#### 5.3.1. Estudio de conectividad asociado con controles

Comparando los procesos de entrenamiento, validación y prueba para las medidas temporales (Tabla 4.8 para la correlación de Pearson y Tabla 4.10 para la entropía de Shannon), observamos que los resultados asociados al conjunto de test para el *modelo 1*, estaban sesgados por los pacientes con EP incluidos en el conjunto de prueba. Por tanto, podemos deducir que la GCNN entrenada con las medidas de conectividad definidas en el dominio temporal tiene dificultades para abstraer las características fundamentales de

la EP y generalizar los resultados, indicando que con una muestra más grande de participantes podríamos mejorar estos resultados. Esto también puede ser debido a que las medidas consideradas son muy sensibles al ruido o artefactos presentes en las señales de EEG. Los resultados de validación son semejantes en los 3 modelos, por lo que no encontramos un efecto visible producido por la medicación por medio de estas medidas de conectividad. Sin embargo, la entropía de Shannon mostró un mejor rendimiento del modelo para el mismo conjunto ciego aunque ambas medidas reflejaron la existencia de diferencias significativas entre la conectividad de los cerebros sanos frente a cerebros con EP. Esto puede ser debido al hecho de que existe un aumento de conectividad generalizado en pacientes con trastornos neurológicos [320], y la GCNN ha identificado este comportamiento para diferenciar a los pacientes enfermos de los sanos. Estos resultados se correlacionan con los obtenidos anteriormente en el estudio mediante series temporales [248], indicando una mayor independencia en los cerebros de los pacientes sanos. Sin embargo, el *modelo 3* mostró un peor rendimiento en la identificación de los controles (especificidad), por lo que nos lleva a pensar que la LD produce un acercamiento al estado del control y que el *modelo 2* ha sido capaz de salvar las diferencias producidas por la medicación.

Co respecto a las medidas espectrales encontramos que los resultados del conjunto de test no difieren tanto de los de validación para el *modelo 1* (Tabla 4.12), mostrando ser más robustos frente a casos particulares. Debemos señalar la tendencia al overfitting de la entropía espectral. Sin embargo, esta tendencia no ha afectado a la generalización de los resultados para el conjunto de prueba, indicando que nos encontramos en un punto crítico a partir del cual el sobre-ajuste condicionaría los resultados. El *modelo 2* (Tabla 4.14) y el *modelo 3* (Tabla 4.16), tienen un comportamiento similar. Llama la atención que los resultados de validación del *modelo 2* muestran un mayor rendimiento en la tarea de identificar a los pacientes con EP, indicando que las diferencias de conectividad entre controles y pacientes tras la toma de la medicación son más relevantes para distinguir la enfermedad.

Comparando los resultados de test de todas las medidas de conectividad, encontramos que las medidas espectrales para el *modelo 1* mostraron un aumento de la sensibilidad con respecto a las medidas temporales, destacando la banda delta en la coherencia y las bandas beta y gamma en la entropía espectral. Por lo que podemos corroborar que las medidas espectrales han generalizado mejor los resultados del modelo. Llama la atención que las bandas de baja frecuencia (delta, theta y alpha) parecen ser las más representativas para identificar la EP, especialmente en la coherencia. Dado que estas bandas se relacionan con estados de sueño y tareas cognitivas (presentes en las etapas pre-motoras de la EP), estos resultados pueden indicar una posible vía para desarrollar

una prueba de detección precoz en la EP. También se hizo notar la diferencia entre los resultados de la prueba “all” en todos los modelos. Para la coherencia, se observa un aumento generalizado en la sincronización de todas las bandas espectrales, de modo que al introducir todas las bandas en el modelo este ha obtenido un buen rendimiento. Sin embargo, en la entropía espectral no se han observado características comunes entre las bandas y, por tanto, el modelo no ha mostrado un buen rendimiento.

### 5.3.2. Estudio de conectividad para los estados dopaminérgicos

Los resultados de accuracy obtenidos al clasificar los estados de activación dopaminérgica (Pre-Post LD), para los procesos de entrenamiento, validación y prueba, indicaron que el proceso de aprendizaje del modelo no se realizó adecuadamente para ninguna medida de conectividad, con valores de predicción realizados al azar. Sin embargo, debemos notar que todas las ejecuciones del estudio fueron realizadas bajo las mismas condiciones, estos son, la misma cantidad de datos y los mismos parámetros de entrenamiento y arquitectura de la red. Consecuentemente, las diferencias en este entrenamiento que han conducido al proceso fallido de aprendizaje vienen dadas por las características de los datos, esto es, la diferencia de activación dopaminérgica en los pacientes con EP (Pre-Post).

Estos resultados indican que las diferencias de conectividad entre un sujeto sano y uno enfermo son mayores que entre dos sujetos enfermos en diferentes estados de activación dopaminérgica, y por tanto, es necesario una mayor cantidad de datos para identificar las diferencias producidas por la medicación. Además, al no observar diferencias significativas entre los valores de las bandas, concluimos que el efecto de la LD es genérico sobre el cerebro del paciente y no un efecto particular sobre las zonas afectadas, tal y como comprobamos con el análisis de las series temporales.

### 5.3.3. Generalización de los resultados

Dadas las limitaciones del conjunto de prueba en el *modelo 1* y el *modelo 2*, se preparó un conjunto externo compuesto por 13 pacientes con EP y 2 controles para evaluar de una forma más precisa la capacidad de generalización de los resultados. Este conjunto de prueba lo componen pacientes enfermos que han tomado la medicación, pero no se encuentran en el máximo estado de efectividad. Para este conjunto de datos, la correlación de Pearson en el *modelo 1* obtuvo el mejor rendimiento (accuracy:  $93,08 \pm 0,95$ , sensitivity:  $95,29 \pm 1,19$ ). Sin embargo, se redujo el valor de especificidad (specificity:  $78,73 \pm 0,78$ ), esto es, la tasa de aciertos de controles, obteniendo un mejor rendimiento por medio de la entropía de Shannon (specificity:  $89,84 \pm 5,05$ ). Este resultado nos con-

duce a pensar que el modelo ha identificado el aumento de conectividad como una de las características esenciales que definen la EP, y ha marcado un valor umbral a partir del cual se considera que un paciente tiene la enfermedad. Si dicho umbral es bajo, el resultado es un modelo muy sensible a las alteraciones de conectividad, de modo que es posible que algunos controles sean mal clasificados mientras que los pacientes se identifican correctamente. Para ambas medidas de conectividad llama especialmente la atención los resultados del *modelo 3*, para el que encontramos valores de accuracy y sensitivity significativamente inferiores a los obtenidos por los otros dos modelos. Dado que este modelo ha entrenado con pacientes con EP en estados ON y OFF, podemos concluir que la LD produce cambios en la conectividad que influyen en la detección de la enfermedad pero no hay un acercamiento claro al estado del control.

En cuanto a las medidas espectrales encontramos que el *modelo 2* identificó mejor a los pacientes con EP con ambas medidas de conectividad y el *modelo 3* fue el que peor desempeño tuvo en dicha tarea. Por lo que podemos deducir que el efecto de la medicación se produce en todo el espectro y por tanto afecta a ambas medidas. Volvemos a observar que las bandas asociadas a bajas frecuencias son las más representativas para identificar la EP, y las bandas de altas frecuencias (especialmente en la coherencia) son las que obtienen una mayor std. Esto puede ser debido a que el intervalo en el que están definidas las bandas beta y gamma es más amplio que para el resto de bandas, de modo que es posible encontrar más alteraciones que aumenten las fluctuaciones en los resultados de las medidas.

## 5.4. Modelo BERT en el diagnóstico de la EP

Dado que no hemos encontrado estudios previos que apliquen técnicas de PLN a los EEGs, no tenemos una referencia en cuanto a la arquitectura del modelo y los resultados esperados. Para verificar la *hipótesis 4* y proceder al desarrollo del *objetivo 4* se propusieron 4 experimentos basados por un lado, en dos estrategias de entrenamiento (considerar todas las pruebas electroencefalográficas o sólo aquellas relacionadas con la activación motora), y por otro lado, dos casos de estudio en función de los electrodos considerados (con los 64 canales que contiene el EEG o con los 28 canales resultantes de eliminar los electrodos periféricos).

### 5.4.1. Experimentos con 28 canales

En la experimentación que tiene en cuenta todas las pruebas electroencefalográficas, la sensibilidad presenta valores superiores al 90 % en todas las etapas de la experimentación (entrenamiento, validación y prueba), pero con desviaciones muy elevadas, en

torno al 44 % en todos los casos. Por otro lado, la especificidad tiene valores muy bajos, en torno al 20 %. Si a eso añadimos que usando un 5-fold CV hay una alta variabilidad en los TN y FN para cada iteración y que el modelo está entrenado con dos clases desbalanceadas, podemos pensar que puede estar sobre-entrenando la clase mayoritaria, estos son los pacientes con EP. Por tanto, el desequilibrio de clases ha afectado negativamente al entrenamiento. Aunque esto podría ser un problema, BERT ha demostrado resultados prometedores trabajando con conjuntos de datos desbalanceados [308].

En cuanto al experimento de 28 canales considerando únicamente las pruebas relacionadas con la activación motora, encontramos que las fluctuaciones en los resultados de cada iteración para TN y FN (relacionados con las medidas de sensibilidad y especificidad) son menores que en el caso anterior (teniendo en cuenta las pruebas de resting), pero sigue siendo significativamente alta como podemos ver en los resultados de la Tabla 4.28. Debemos tener en cuenta que en esta experimentación disponemos de menos datos que en el caso anterior al prescindir de una de las pruebas de EEG, aunque porcentualmente se mantiene la diferencia entre las clases. Este hecho afecta especialmente a la métrica de especificidad, como podemos observar en los resultados de la Tabla 4.28, que además de tener un valor significativamente bajo, su std exhibe valores altos, alrededor del 30 %. En cuanto a las métricas de exactitud y precisión, no se observan diferencias significativas entre los resultados en ambos procesos.

Al comparar los experimentos de 28 canales entrenados con todas las pruebas de EEGs y con los EEGs de FT que se muestran en la Tabla 4.28, llama la atención que los resultados de las métricas de precisión y exactitud del modelo no difieren mucho en ambos entrenamientos. Donde sí encontramos diferencias es en las métricas de sensibilidad y especificidad, debido a fluctuaciones en las iteraciones del proceso de validación sobre los valores de TN y FN, ambas relacionadas con las predicciones negativas (controles correctamente clasificados y pacientes con EP diagnosticados como controles), que parecen estar influenciadas por el desbalanceo de clases. Dado que el conjunto de controles es menor que el de pacientes con PD, este resultado nos lleva a pensar, de nuevo, que las fluctuaciones en las predicciones pueden deberse al desbalanceo entre clases. Observando los resultados obtenidos en la Tabla 4.28 para cada uno de los procesos, no encontramos grandes diferencias entre ambas estrategias de entrenamiento. Consecuentemente, podemos decir que, en este caso, las pruebas motoras son significativas a la hora de diagnosticar la EP, mientras que las pruebas registradas en estado de reposo, juegan un papel secundario.

### 5.4.2. Experimentos con 64 canales

Para el experimento que considera únicamente los EEGs relacionados con activaciones motoras, podemos observar que tanto la exactitud como la precisión tienen valores altos en las diferentes etapas de la experimentación. Sin embargo, cabe señalar que durante el proceso de validación cruzada, para los FP, hay una iteración que contiene valores muy diferentes al resto de iteraciones y por tanto estos resultados alteran las medidas de las métricas que contienen este valor. Como este suceso sólo aparece en una única iteración, deducimos que es un evento puntual derivado de la elección de los datos y no de un error en el entrenamiento.

En el caso del experimento que incluye todos los EEGs, sólo los valores de FN han mostrado cierta variabilidad en las iteraciones del CV, aunque mucho menos que en los experimentos anteriores de 28 canales. Esto se refleja en los valores obtenidos de las métricas y su desviación estándar, donde aún prevalecen valores bajos de la métrica de especificidad con una alta variabilidad entre los procesos de entrenamiento. Sin embargo, los resultados de este experimento no indican una afectación por el desequilibrio de clases, ya que no existen variaciones significativas en el caso de los TN, mientras que en los FN dicha fluctuación se ha reducido considerablemente. Esta cuestión puede ocurrir porque, al considerar más datos, el modelo puede relacionar mejor la información, minimizando el efecto causado por el desbalanceo de clases. Podemos corroborar este resultado con el aumento de los valores de las métricas de precisión y exactitud relativas al modelo de 28 canales utilizando todos los datos.

La comparativa entre los experimentos con 64 canales que se observa en la Tabla 4.29 arrojó resultados muy diferentes en las métricas consideradas. Los resultados con menos varianza entre las métricas y valores más altos fueron para el experimento *64\_FT*. También se han evaluado los resultados de cada iteración, y salvo un caso concreto cuyos resultados se vieron afectados por la elección de los datos, el resto del entrenamiento durante el método de validación fue homogéneo, por lo que podemos considerar que el entrenamiento del modelo ha sido óptimo. Dado que al comparar los experimentos *28\_all* y *28\_FT* no encontramos diferencias significativas entre sus valores de precisión, podemos deducir que en el caso del experimento *64\_all*, el entrenamiento ha sido insuficiente. Recordemos que las arquitecturas e hiper-parámetros han sido iguales en todos los experimentos para favorecer la comparación entre ellos y favorecer la evaluación tanto de las pruebas de electroencefalografía como de los canales del EEG a los resultados del modelo.

En términos generales, los experimentos con 64 canales han proporcionado mejores resultados que los experimentos con 28 canales. Esto puede ocurrir porque un modelo BERT requiere más datos para realizar un entrenamiento, y quitar parte de los electrodos

no contribuye a mejorar los resultados del problema de clasificación. También podríamos pensar que, en el caso de la EP, la zona afectada se extiende a zonas periféricas, y por tanto estos electrodos también contienen información sobre la enfermedad. Para probar esta hipótesis, se requiere de un nuevo entrenamiento con un número intermedio de canales, además de un dataset con más EEGs tanto de pacientes con EP como de controles.

## 5.5. Discusión global

En esta tesis buscamos un biomarcador de la EP basado en los EEGs de los pacientes. La aplicación de la electroencefalografía como prueba diagnóstica de la EP no está extendida en la práctica clínica diaria debido a la difícil interpretación de las señales del EEG. Sin embargo, debido a las numerosas ventajas que ofrece esta prueba la hemos propuesto como punto de partida de nuestra investigación. Actualmente, el criterio de diagnóstico clínico se ha formalizado mediante los criterios del Banco de Cerebros de la Sociedad de la Enfermedad de Parkinson del Reino Unido, con una precisión diagnóstica de hasta el 90 % [7], en el caso de un paciente con EP establecida, y en torno al 80 % para los pacientes que son clínicamente probables de tener EP [29]. Consideramos estas medidas como referencia para determinar la precisión del marcador.

Como primera aproximación, utilizamos métodos de análisis de series temporales para buscar diferencias entre los EEGs de controles y de pacientes con EP. Para ello utilizamos medidas de conectividad relacionadas con el espectro de frecuencias, entre ellas la coherencia y la correlación de Pearson de la entropía espectral. Aunque la muestra de este estudio fue pequeña (6 pacientes y 6 controles) algunos de los resultados ofrecidos por las diferentes medidas mostraron diferencias estadísticamente significativas en algunas de las pruebas consideradas.

La correlación de la entropía espectral mostró un aumento en los valores de los pacientes con EP con respecto a los controles, especialmente en las pruebas relacionadas con la activación motora de ambos hemisferios. Esto se ha asociado con una mayor independencia del cerebro de los pacientes sanos. Debido a que la EP presenta un daño cerebral en el paciente, el cerebro intenta suplir estas deficiencias con otras regiones aumentando por tanto la conectividad y la dependencia de las distintas zonas cerebrales, pudiendo reducir la eficiencia en la tarea ejecutada y producir un peor rendimiento o lentitud en el proceso. Este resultado se corrobora con los resultados obtenidos para las medidas de la coherencia y la divergencia de KL.

Los resultados obtenidos utilizaron los EEGs de pacientes antes de su toma de medicación, esto es, en estado OFF. Dado que la levodopa puede alterar los patrones de

conectividad y por tanto alterar los análisis realizados, incluimos en los análisis los EEGs de los mismos pacientes tras la toma de la mediación dopaminérgica. En este caso, aunque las diferencias observadas son equivalentes a las del estudio de pre-LD y controles, se puede notar un acercamiento a los valores obtenidos por los controles tras la toma de la medicación, es decir, la levodopa frena el aumento del coeficiente de correlación producido en la activación motora. Este resultado se puede corroborar con las otras medidas de conectividad consideradas.

Debemos notar que no todos los experimentos mostraron diferencias significativas, además de estar influenciados por el tipo de procesamiento de limpieza del EEG, impidiendo realizar predicciones individuales para cada paciente debido a la heterogeneidad presente en la enfermedad. Por tanto, las limitaciones de esta metodología para el desarrollo de un marcador de la enfermedad, nos conducen a buscar otras herramientas que permitan procesar los EEGs en bruto, y que muestren diferencias de modo que nos permita realizar estudios individuales dentro de la heterogeneidad de la sintomatología de cada paciente. En consecuencia, proponemos las técnicas de DL para desarrollar un marcador de la EP.

Para marcar un punto de partida a nuestra investigación realizamos dos revisiones bibliográficas acerca del tratamiento de los EEGs de pacientes con EP mediante técnicas de ML, una con un enfoque clínico en el que se analizaron 9 artículos científicos, y la otra con un enfoque computacional en el que se analizaron 59 estudios. Ambas revisiones coinciden en que el tema propuesto para esta tesis es un área poco explorada que ha cobrado mayor interés en los últimos años (a partir del año 2019). El objetivo de la mayor parte de los estudios resultantes se centró en el problema de diagnóstico, que ha sido abordado desde diferentes perspectivas en cuanto al pre-procesamiento de datos, características introducidas y modelos empleados.

Además de las lecciones prácticas extraídas de estas revisiones, relacionadas con la calidad del conjunto de datos y los tipos de pre-procesamiento empleados, las principales conclusiones que marcaron el rumbo de esta tesis están relacionadas con los modelos de Machine Learning empleados y con los procesos de entrenamiento.

Con respecto al entrenamiento, la metodología más empleada fue el proceso de validación cruzada k-fold, que proporciona valores más fiables al eliminar el bias producido por la elección de los datos. Entre los modelos más utilizados destaca el uso de SVM y las ANN, aunque las arquitecturas que predominan en los últimos años son modelos de DL, particularmente las CNN, solas o en combinación con las RNN por medio de un modelo híbrido.

Las revisiones bibliográficas nos proporcionaron una perspectiva objetiva sobre los aspectos que todavía no habían sido abordadas para el análisis de EEGs en la EP me-



dianter técnicas de DL. Desde el punto de vista clínico, destaca la ausencia de estudios de conectividad funcional, por lo que parte de la investigación de esta tesis se centró en el desarrollo del biomarcador basado en las diferencias de conectividad entre pacientes con EP y controles. Desde el punto de vista computacional, a pesar del gran potencial mostrado por los modelos dentro del campo del PLN, estas técnicas no habían sido propuestas para el análisis de EEGs, por lo que otra rama de la investigación llevada a cabo fue la aplicación de estas técnicas basándonos en la equivalencia entre texto-EEG.

Las alteraciones en la conectividad del cerebro pueden relacionarse con anomalías o trastornos neuropsiquiátricos. Dado los resultados previos relacionados con alteraciones en la conectividad de los pacientes con EP, proponemos un modelo que sea capaz de identificar estas diferencias para clasificar a los pacientes con EP. Para llevar a cabo este estudio, sugerimos la teoría de grafos como la metodología empleada para modelar los EEGs y relacionar las diferentes regiones del cerebro por medio de diferentes medidas de conectividad. Por tanto, en primer lugar, desarrollamos EEGraph, una herramienta en el entorno de Python capaz de realizar este proceso de forma automática permitiendo un sencillo manejo de las variables. Las conexiones del grafo se definen por medio de medidas de conectividad, unas definidas en el dominio temporal y otras en el dominio espectral.

Con esta herramienta modelamos los EEGs de pacientes con EP en diferentes estados de activación dopaminérgica (estados OFF y ON) y los EEGs de controles mediante 4 medidas de conectividad, dos temporales (correlación de Pearson y entropía de Shannon) y dos espectrales (coherencia y entropía espectral). Analizamos la conectividad por medio de una red específica para el análisis de grafos (GCNN) que incluye un módulo de clasificación para poder realizar la predicción de las clases. Los resultados mostraron que las medidas definidas en el dominio de las frecuencias generalizaron mejor los resultados de la red, y entre los modelos que enfrentaban los EEGs de pacientes y controles, el experimento Control frente pacientes post LD para la coherencia de la banda beta ( $90,62 \pm 8,51$ ), ofreció el mejor valor de rendimiento del modelo para el proceso de validación mediante una 5-fold CV. Sobre el conjunto externo de prueba compuesto por pacientes con EP tras la toma de la medicación pero no necesariamente en su estado ON, encontramos que los únicos modelos que ofrecen un valor en la tasa de aciertos superior al 90% son el modelo control-pre LD ( $93,08 \pm 0,95$ ) y el modelo control-post LD ( $92,04 \pm 2,75$ ), ambos para la correlación de Pearson. Estos resultados indican la existencia de diferencias significativas entre los grupos de controles y pacientes con EP independientemente del estado de activación dopaminérgica, corroborando los resultados obtenidos anteriormente. Con respecto a la sensibilidad, esto es la tasa de acierto de pacientes con EP y por tanto proporciona una medida acerca de la capacidad de la

red para identificar la enfermedad, los modelos que superan el 90 % son control-pre LD y control-post para la correlación de Pearson (con  $95,29 \pm 1,19$  y  $93,53 \pm 4,29$  respectivamente) y el modelo control-post para la coherencia de la banda delta ( $92,28 \pm 5,95$ ). Para estos casos los valores de especificidad, asociados con la tasa de aciertos de los controles son respectivamente  $78,73 \pm 0,78$ ,  $82,38 \pm 8,55$  y  $85,40 \pm 8,35$ . Con respecto a las medidas espectrales se debe resaltar la existencia de diferencias significativas en la sincronización de las bandas entre controles y pacientes con EP, corroborando nuevamente los resultados obtenidos mediante el análisis de series temporales.

Debemos notar que las predicciones se realizan a nivel de grafo, y que cada grafo se corresponde con un segmento finito del EEG. De modo que si realizamos la evaluación por cada EEG del paciente, esto es realizando la suma ponderada de las predicciones de los grafos que componen dicho EEG, las tasas de aciertos aumentan considerablemente en cada una de las métricas para todos los modelos implementados.

Para finalizar con el estudio de conectividad se enfrentó los EEGs de los pacientes con EP antes y después de la ingesta de su mediación. Sin embargo, las diferencias entre los estados dopaminérgicos no fue significativa. Dado que los experimentos se llevaron a cabo bajo las mismas condiciones y las diferencias de conectividad están producidas por la medicación, esto nos conduce a que las diferencias entre un paciente con EP y un control son mayores que entre un paciente con EP en distintos estados de activación dopaminérgica, por lo que la LD no produce un claro acercamiento al estado de conectividad del control.

En cuanto a la propuesta de computación, introducimos la técnica de PLN para el análisis de los EEGs. Se realizó un diseño experimental basado en la equivalencia texto-EEG debido al carácter secuencial de los datos implicados y se propuso un modelo BERT para llevar a cabo la clasificación de la enfermedad entre controles y pacientes con EP. Se realizaron varios experimentos teniendo en cuenta los 64 electrodos que contiene el casco, y eliminando los canales periféricos, quedando un total de 28 canales. Aunque los resultados bibliográficos anteriores indican que el número de electrodos no ha sido un factor determinante en los resultados de los modelos, estos resultados no han sido corroborados con modelos dentro del campo del PLN, donde la variación en el número de canales si puede ser un factor relevante. Se consideraron 2 estrategias de entrenamiento, una que utilizaba las 3 pruebas electroencefalográficas registradas, y otra que consideraba únicamente las pruebas relacionadas con la activación motora. A pesar de que el modelo BERT no está condicionado por el desbalanceo de clases, los experimentos con 28 canales mostraban sesgos producidos por las clases no equilibradas (valores de especificidad de  $30,53 \pm 25,06$  para las pruebas motoras y de  $18,08 \pm 20,64$  para todas las pruebas en el proceso de prueba), al presentar grandes fluctuaciones de los

valores de TN y FN durante todo el proceso de validación cruzada. Esto puede ser consecuencia de que la cantidad de señales consideradas en el caso de los 28 canales es menor que en el caso de 64 electrodos y por ello, estos experimentos se ven afectados por la diferencia de clases. Consecuentemente, al aumentar el número de señales con los experimentos de 64 electrodos, encontramos una disminución de estas fluctuaciones (valores de especificidad de  $96,00 \pm 8,94$  para las pruebas motoras y de  $36,00 \pm 8,93$  para todas las pruebas en el proceso de prueba), favoreciendo el aprendizaje del modelo. Comparando los experimentos que utilizan los EEGs que reflejan la activación motora frente a los experimentos que consideran todas las pruebas electroencefalográficas para los 64 electrodos, encontramos que la prueba registrada en estado de reposo no añade información relevante para identificar a los pacientes con EP, obteniendo una tasa de aciertos del  $86,41 \pm 7,08\%$ .

Los resultados presentados muestran que se ha conseguido desarrollar dos marcadores para la EP. Por un lado, mediante el estudio de conectividad funcional, se han obtenido métricas de accuracy, sensibilidad y especificidad con valores de  $93,08 \pm 0,95$ ,  $95,29 \pm 1,19$  y  $78,73 \pm 0,78$  para la correlación de Pearson de controles frente a pacientes en estado ON. Por otro lado, mediante técnicas de PLN, los valores de las métricas de accuracy, sensibilidad y especificidad obtenidas son de  $86,41 \pm 7,08$ ,  $96,00 \pm 8,94$  y  $81,63 \pm 11,58$  para el modelo de 64 canales entrenado con las pruebas motoras. Recordando los valores de referencia del Banco de Cerebros de la Sociedad de la Enfermedad de Parkinson del Reino Unido, tomamos el valor de diagnóstico en  $90\%$  para los casos establecidos (siendo de  $80\%$  para los casos probables). Podemos observar que el marcador basado en grafos ha superado este valor. Debido al estudio previo realizado y a nuestro conocimiento en la materia, creemos que el biomarcador basado en la equivalencia texto-EEG puede superar estas cifras con una mayor cantidad de datos. Entre las limitaciones de estos estudios encontramos que los pacientes enfermos ya han sido diagnosticados con EP. Por ello, es necesario comprobar la efectividad de los métodos desarrollados en la práctica clínica, en el proceso del diagnóstico de la enfermedad.

## 5.6. Trabajo Futuro

Los trabajos llevados a cabo para el desarrollo del biomarcador de la EP abren el camino a nuevas investigaciones que empleen otros modelos de Deep Learning en el ámbito de los dos estudios llevados a cabo con técnicas de IA.

Por parte del estudio de conectividad, proponemos un estudio alternativo basado en la teoría de grafos con el fin de evaluar las diferencias producidas por diferentes terapias en los pacientes con EP. El modelo empleado se basará en GAT (por sus siglas en

inglés “Graph Attention Network”), y realizará una predicción sobre los nodos indicando la probabilidad de estar afectado por la enfermedad tras evaluar las diferencias de conectividad existentes entre pacientes con EP y controles. La arquitectura de esta red incluye módulos de atención (similares a los usados en modelos de PLN) encargados de identificar patrones de conectividad que después serán analizados junto con los restantes grafos que conforman el EEG. De este modo podremos visualizar las zonas más significativas de distintos pacientes, lo que nos permitiría realizar el diagnóstico de la enfermedad o comparar los de un mismo paciente antes y después recibir una terapia, permitiendo evaluar los beneficios sobre la conectividad funcional.

Además, los resultados de la conectividad funcional proporcionados por la coherencia indicaron que la banda espectral más representativa para identificar la EP es la banda delta. Dado que esta banda está asociada con estados del sueño, y las alteraciones del sueño pueden estar presentes en las etapas pre-motoras de la enfermedad, proponemos un estudio basado en EEGs del sueño mediante técnicas de DL para desarrollar un marcador precoz de la EP.

Por parte del tratamiento de los EEGs con técnicas de PLN, proponemos un estudio basado en modelos Transformers para realizar una estadificación de la enfermedad, esto es identificar las diferentes fases de la EP indicando los canales más representativos para cada estado de progresión de la enfermedad. En base a la equivalencia texto-EEG, identificaremos una frase con un canal y un segmento del canal con un palabra. Así, se buscará que el modelo sea capaz de hallar relaciones entre las diferentes partes del EEG, para encontrar las características que identifiquen cada estado de la enfermedad. Para ello utilizaremos EEGs de pacientes con EP, etiquetados con el valor obtenido en la escala HY, junto con EEGs de controles. Esto nos permitiría obtener la matriz de atención para recoger la importancia que el modelo asocia a cada canal a la hora de identificar los diferentes estados de la enfermedad. De este modo podremos generar una imagen indicando las zonas más afectadas en cada estadio de la enfermedad.

## Capítulo 6

# Conclusiones

En este capítulo se resumen, a modo de conclusiones, las aportaciones realizadas por esta tesis.

La principal aportación de esta tesis es la confirmación de que, a falta de una prueba barata y accesible que sea específica para apoyar el diagnóstico de la EP, el análisis de registros electroencefalográficos mediante técnicas de Deep Learning puede constituir un marcador de la enfermedad. Los resultados obtenidos mediante las dos líneas de investigación propuestas, muestran una tasa de aciertos del  $93,08 \pm 0,95\%$  para el biomarcador basado en la conectividad funcional y un  $86,41 \pm 7,08\%$  para el biomarcador basado en técnicas de PLN. Ambos modelos han mostrado su potencial como biomarcadores de la enfermedad, especialmente la Graph Convolutional Neural Network (GCNN), cuyos resultados mejoran el valor propuesto como referencia del criterio de diagnóstico del Banco de Cerebros de la Sociedad de la Enfermedad de Parkinson del Reino Unido, siendo este valor del 90 % para pacientes con EP establecida. Para llevar a cabo la aplicación de estos marcadores a la práctica clínica diaria, es necesario que:

- Sean validados con EEGs de pacientes no diagnosticados todavía, de manera que el rendimiento de los modelos alcanzasen una precisión mayor que el 80 % considerado por el brain bank.
- Sean entrenados con un mayor conjunto de datos, especialmente el modelo de PLN, para mejorar y generalizar los resultados.

De forma más detallada, y siguiendo el orden marcado por los objetivos específicos planteados en esta tesis, especificamos a continuación las restantes conclusiones.

Respecto al *objetivo específico 1*, basado en el estudio de las diferencias de conectividad mediante el análisis de EEGs usando técnicas de series temporales, se concluye que estas técnicas:

- Detectan un aumento generalizado de la conectividad en los pacientes con EP en las pruebas relacionadas con la activación motora de ambos hemisferios, aumento que asociamos con un intento de suplir las alteraciones presentes en el cerebro enfermo que afectan a la eficiencia de la tarea realizada.
- Aunque son sensibles a las alteraciones de conectividad producidas por la enfermedad, no son aplicables a la evaluación de pacientes individuales al presentar una gran dispersión producida por la heterogeneidad de los síntomas de los participantes. Por tanto, esta aproximación no es válida para desarrollar el biomarcador.

Respecto al *objetivo específico 2*, por medio de una revisión sistemática de la literatura concluimos que:

- Las técnicas de DL son capaces de analizar los EEGs de enfermedades neurológicas con mayor nivel de abstracción, permitiendo detectar patrones con un alto grado de precisión.
- El proceso de limpieza no influye de manera significativa en los resultados de rendimiento de los modelos de IA.

Respecto al *objetivo específico 3*, el estudio de conectividad funcional cerebral mediante la aplicación de modelos GCNN a los EEGs modelados como grafos permite concluir que:

- Los resultados de las medidas de conectividad corroboran las conclusiones obtenidas con el análisis de series temporales.
- Las medidas de conectividad definidas en el dominio espectral muestran una mayor capacidad para generalizar los resultados que las definidas en el dominio temporal.
- La banda más representativa para identificar la enfermedad es la banda delta, que corresponde con estados del sueño, lejos de las bandas de alta frecuencia asociadas con estados de actividad en los que se suele detectar la EP.

- Es posible plantear un biomarcador basado en un modelo GCNN que apoye el diagnóstico de la enfermedad, en los términos expresados al comienzo de este capítulo.

Respecto al *objetivo específico 4*, basado en la detección de la EP mediante el estudio de los EEGs utilizando técnicas de PLN, podemos concluir que:

- Es válido suponer la equivalencia EEG-texto y la aplicación de modelos BERT basados en Transformers para encontrar patrones en el EEG que identifiquen la enfermedad.
- Las pruebas de estado de reposo no añaden información relevante para mejorar el rendimiento del modelo, corroborando los resultados obtenidos en el análisis mediante series temporales.
- El uso de modelos de PLN requiere una mayor cantidad de datos para evitar el overfitting y favorecer la generalización de los resultados.
- Es posible plantear un biomarcador basado en modelos de PLN que apoye el diagnóstico de la enfermedad, en los términos expresados al comienzo de este capítulo.

Consecuentemente, hemos alcanzado los objetivos propuestos en esta tesis obteniendo dos biomarcadores de la enfermedad de Parkinson con potencial de ser aplicados en la clínica para apoyar el diagnóstico de la enfermedad.





# Apéndice A

## Tablas auxiliares MDS

The first essential criterion is parkinsonism, which is defined as bradykinesia, in combination with at least 1 of rest tremor or rigidity. Examination of all cardinal manifestations should be carried out as described in the MDS–Unified Parkinson Disease Rating Scale.<sup>30</sup> Once parkinsonism has been diagnosed:

Diagnosis of Clinically Established PD requires:

1. Absence of absolute exclusion criteria
2. At least two supportive criteria, and
3. No red flags

Diagnosis of Clinically Probable PD requires:

1. Absence of absolute exclusion criteria
2. Presence of red flags counterbalanced by supportive criteria  
If 1 red flag is present, there must also be at least 1 supportive criterion  
If 2 red flags, at least 2 supportive criteria are needed  
No more than 2 red flags are allowed for this category

Supportive criteria

(Check box if criteria met)

- 1. Clear and dramatic beneficial response to dopaminergic therapy. During initial treatment, patient returned to normal or near-normal level of function. In the absence of clear documentation of initial response a dramatic response can be classified as:
  - a) Marked improvement with dose increases or marked worsening with dose decreases. Mild changes do not qualify. Document this either objectively (>30% in UPDRS III with change in treatment), or subjectively (clearly-documented history of marked changes from a reliable patient or caregiver).
  - b) Unequivocal and marked on/off fluctuations, which must have at some point included predictable end-of-dose wearing off.
- 2. Presence of levodopa-induced dyskinesia
- 3. Rest tremor of a limb, documented on clinical examination (in past, or on current examination)
- 4. The presence of either olfactory loss or cardiac sympathetic denervation on MIBG scintigraphy

Absolute exclusion criteria: The presence of any of these features rules out PD:

- 1. Unequivocal cerebellar abnormalities, such as cerebellar gait, limb ataxia, or cerebellar oculomotor abnormalities (eg, sustained gaze evoked nystagmus, macro square wave jerks, hypermetric saccades)
- 2. Downward vertical supranuclear gaze palsy, or selective slowing of downward vertical saccades
- 3. Diagnosis of probable behavioral variant frontotemporal dementia or primary progressive aphasia, defined according to consensus criteria<sup>31</sup> within the first 5 y of disease
- 4. Parkinsonian features restricted to the lower limbs for more than 3 y
- 5. Treatment with a dopamine receptor blocker or a dopamine-depleting agent in a dose and time-course consistent with drug-induced parkinsonism
- 6. Absence of observable response to high-dose levodopa despite at least moderate severity of disease
- 7. Unequivocal cortical sensory loss (ie, graphesthesia, stereognosis with intact primary sensory modalities), clear limb ideomotor apraxia, or progressive aphasia
- 8. Normal functional neuroimaging of the presynaptic dopaminergic system
- 9. Documentation of an alternative condition known to produce parkinsonism and plausibly connected to the patient's symptoms, or, the expert evaluating physician, based on the full diagnostic assessment feels that an alternative syndrome is *more likely* than PD

Red flags

- 1. Rapid progression of gait impairment requiring regular use of wheelchair within 5 y of onset
- 2. A complete absence of progression of motor symptoms or signs over 5 or more y unless stability is related to treatment
- 3. Early bulbar dysfunction: severe dysphonia or dysarthria (speech unintelligible most of the time) or severe dysphagia (requiring soft food, NG tube, or gastrostomy feeding) within first 5 y
- 4. Inspiratory respiratory dysfunction: either diurnal or nocturnal inspiratory stridor or frequent inspiratory sighs
- 5. Severe autonomic failure in the first 5 y of disease. This can include:
  - a) Orthostatic hypotension<sup>32</sup>—orthostatic decrease of blood pressure within 3 min of standing by at least 30 mm Hg systolic or 15 mm Hg diastolic, in the absence of dehydration, medication, or other diseases that could plausibly explain autonomic dysfunction, or
  - b) Severe urinary retention or urinary incontinence in the first 5 y of disease (excluding long-standing or small amount stress incontinence in women), that is not simply functional incontinence. In men, urinary retention must not be attributable to prostate disease, and must be associated with erectile dysfunction
- 6. Recurrent (>1/y) falls because of impaired balance within 3 y of onset
- 7. Disproportionate anterocollis (dystonic) or contractures of hand or feet within the first 10 y
- 8. Absence of any of the common nonmotor features of disease despite 5 y disease duration. These include sleep dysfunction (sleep-maintenance insomnia, excessive daytime somnolence, symptoms of REM sleep behavior disorder), autonomic dysfunction (constipation, daytime urinary urgency, symptomatic orthostasis), hyposmia, or psychiatric dysfunction (depression, anxiety, or hallucinations)
- 9. Otherwise-unexplained pyramidal tract signs, defined as pyramidal weakness or clear pathologic hyperreflexia (excluding mild reflex asymmetry and isolated extensor plantar response)
- 10. Bilateral symmetric parkinsonism. The patient or caregiver reports bilateral symptom onset with no side predominance, and no side predominance is observed on objective examination

Criteria Application:

1. Does the patient have parkinsonism, as defined by the MDS criteria?  
If no, *neither* probable PD nor clinically established PD can be diagnosed. *If yes:* Yes  No
2. Are any absolute exclusion criteria present?  
If "yes," *neither* probable PD nor clinically established PD can be diagnosed. *If no:* Yes  No
3. Number of red flags present \_\_\_\_\_
4. Number of supportive criteria present \_\_\_\_\_
5. Are there at least 2 supportive criteria *and* no red flags?  
If yes, patient meets criteria for clinically established PD. *If no:* Yes  No
6. Are there more than 2 red flags?  
If "yes," probable PD *cannot* be diagnosed. *If no:* Yes  No
7. Is the number of red flags equal to, or less than, the number of supportive criteria?  
If yes, patient meets criteria for probable PD Yes  No

Figura A.1: Tabla con criterios de diagnóstico clínicos de la EP. Fuente: [29].



## Apéndice B

# Protocolo de Validación de EEGs

## Protocolo de Validación de EEGs

El presente protocolo establece los pasos para realizar la validación visual de un electroencefalograma (EEG). Los ficheros utilizados tienen el formato “.gdf” y utilizaremos el software Matlab en su versión 15.b compatible con la extensión EEGLAB en su versión eeglab15\_0b.

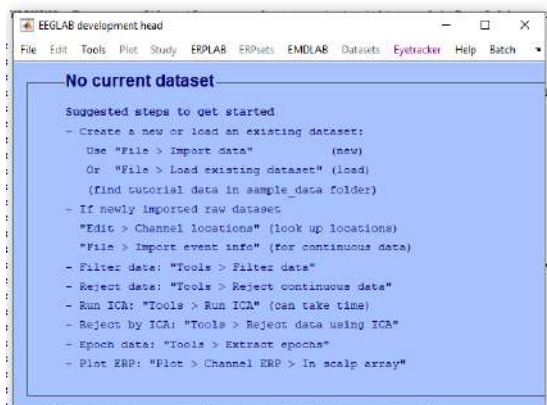
El resultado de la inspección visual será registrado en un fichero “.xlsx”, junto con una pequeña descripción acerca de las características más significativas del EEG.

1.- Abrimos Matlab y seleccionamos en el directorio la carpeta de EEGLAB.

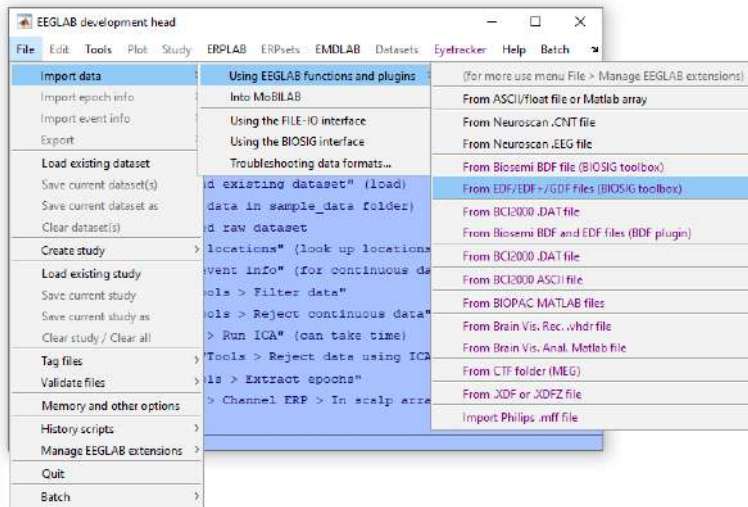
2.- Una vez tenemos abierta la carpeta de EEGLAB en el directorio, introducimos en la ventana de comandos la orden

```
>>eeglab
```

Aparecerá el siguiente cuadro de trabajo.



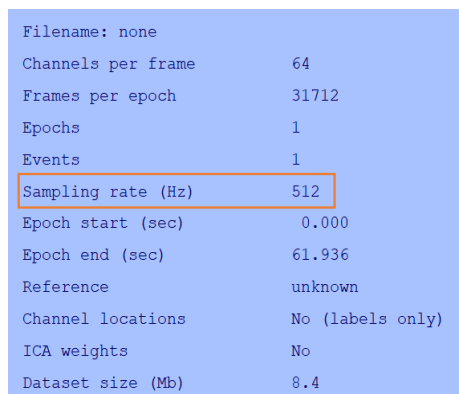
3.- Para abrir un EEG en formato GDF, seguimos la ruta [File -> Import data -> Using EEGLAB functions and plugins -> From EDF/EDF+/GDF files \(BIOSIG toolbox\)](#) y seleccionamos el EEG que queremos abrir.



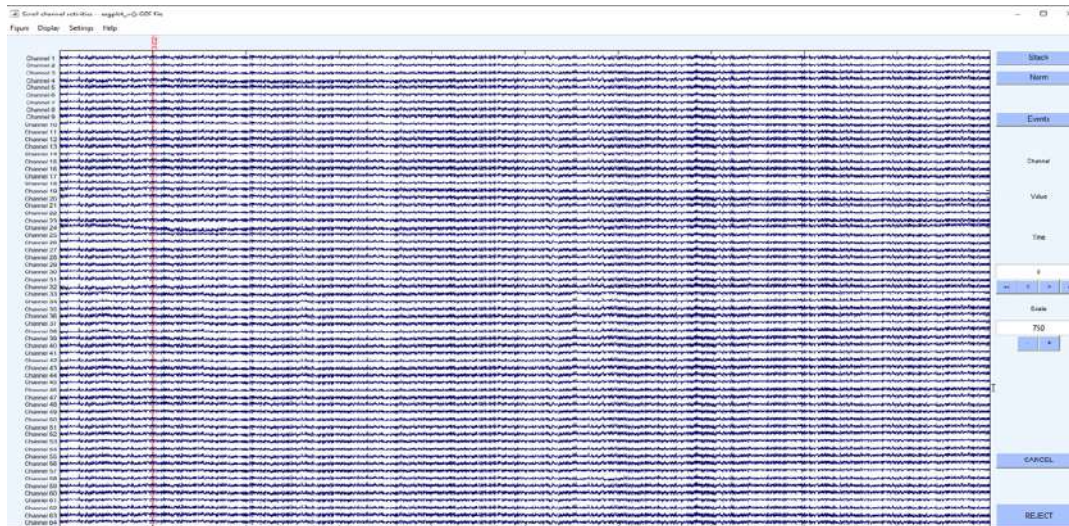
4.- El resultado es una ventana con la información del EEG tal y como se muestra en a continuación.



5.- Confirmamos la frecuencia de muestreo a 512 Hz. De no ser así procedemos a la exclusión del EEG.

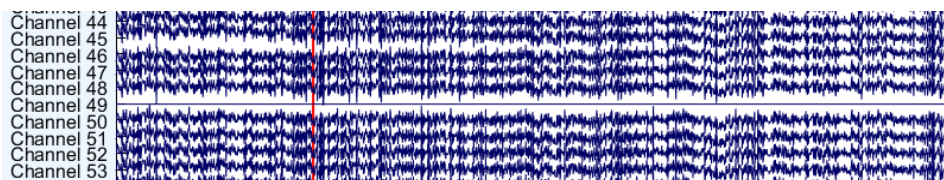


6.- Visualizamos el archivo por medio de la ruta **Plot -> Channel data (scroll)+**. Cuyo resultado se muestra a continuación

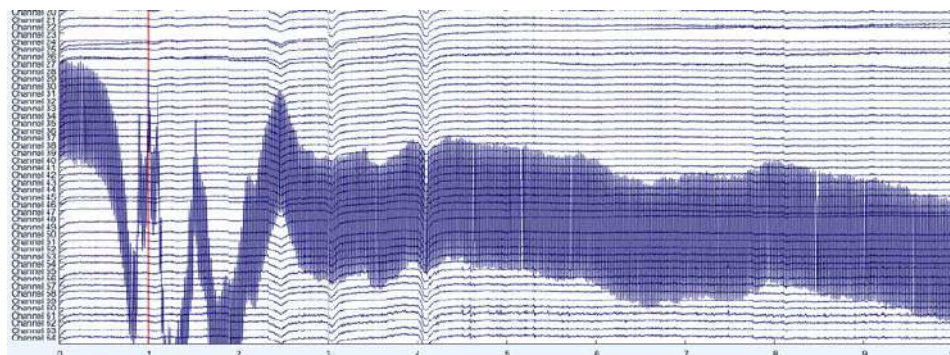


7.- El EEG no debe presentar más de 13 canales anómalos (correspondiendo al  $\approx 20\%$  de pérdida de información), o unos 6 canales contiguos, que representan la pérdida de información de una región completa. A continuación, se especificarán los casos denominados como canales anómalos.

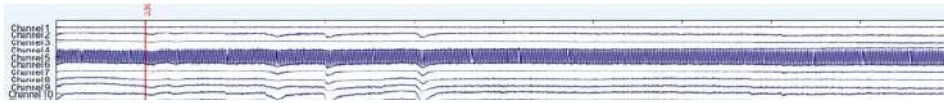
8.- Se identificarán los electrodos planos, y que por tanto no registran señal.



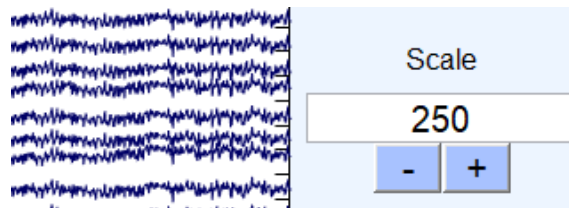
9.- Se identificarán los electrodos defectuosos cuya impedancia no ha sido regulada y por tanto se registran a una escala diferente que el resto de señales del EEG.



10.- Se identificarán los electrodos defectuosos cuya señal contiene una alta tasa de artefactos.



11.- La escala a la que se visualizan los canales del EEG debe estar en el rango (250-750), pudiendo en algunos casos ampliarse a (1000-100). Se considera que se visualizan correctamente si se pueden apreciar los picos característicos de cada canal y se puede diferenciar claramente del resto de canales.



12.- Se observará el registro completo del EEG para comprobar que las características de las señales se mantienen durante toda la grabación y que no aparecen alteraciones en las señales como las que hemos mencionado. En caso de aparecer en un tramo superior a un 15% del total del registro, se procederá a la exclusión del EEG.



# Apéndice C

## Protocolo de Limpieza de EEGs

Protocolo limpieza EEGs

### *Protocolo limpieza EEGs*

En este protocolo describiremos el proceso de limpieza para un Electroencefalograma (EEG). Trabajaremos con el software Matlab y su extensión EEGLAB. Antes de comenzar debemos corroborar que EEGLAB sea compatible con la versión que tengamos de Matlab. Para estos ejemplos se ha utilizado la versión Matlab15.b de Matlab y la versión eeglab15\_0b de EEGLAB.

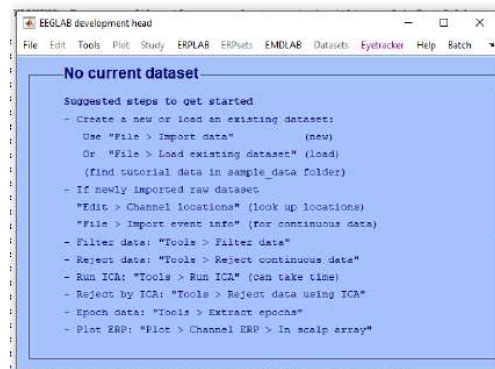
Además de limpiar los EEGs, generaremos una lista donde podamos identificar cada uno de los EEGs junto con las propiedades que nos vayamos encontrando, entre ellas el número de electrodos anómalos y su identificación, si no se ha podido realizar alguno de los procedimientos de limpieza, la existencia e identificación de señales planas.

1.- Abrimos Matlab y seleccionamos en el directorio la carpeta de EEGLAB.

2.- Una vez tenemos abierta la carpeta de EEGLAB en el directorio, introducimos en la ventana de comandos la orden

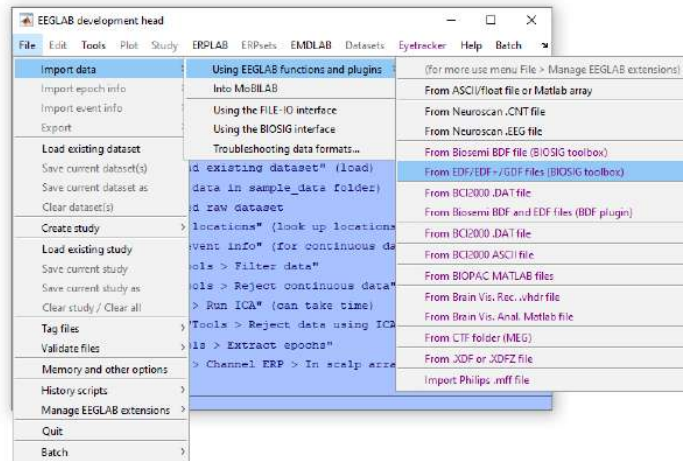
```
>> eeglab
```

Aparecerá el siguiente cuadro de trabajo.

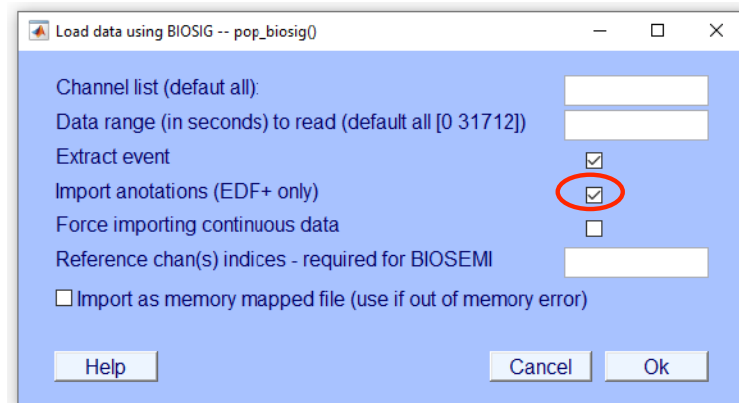


Protocolo limpieza EEGs

3.- Para abrir un EEG en formato GDF, seguimos la ruta **File -> Import data -> Using EEGLAB functions and plugins -> From EDF/EDF+/GDF files (BIOSIG toolbox)** y seleccionamos el EEG que queremos abrir.



4.- Una vez abierto, aparecerá la siguiente ventana, donde desmarcaremos la casilla “Import anotaciones (EDF + only)” y pulsamos “ok”.



5.- En la siguiente pantalla presionamos “ok”.



## Protocolo limpieza EEGs

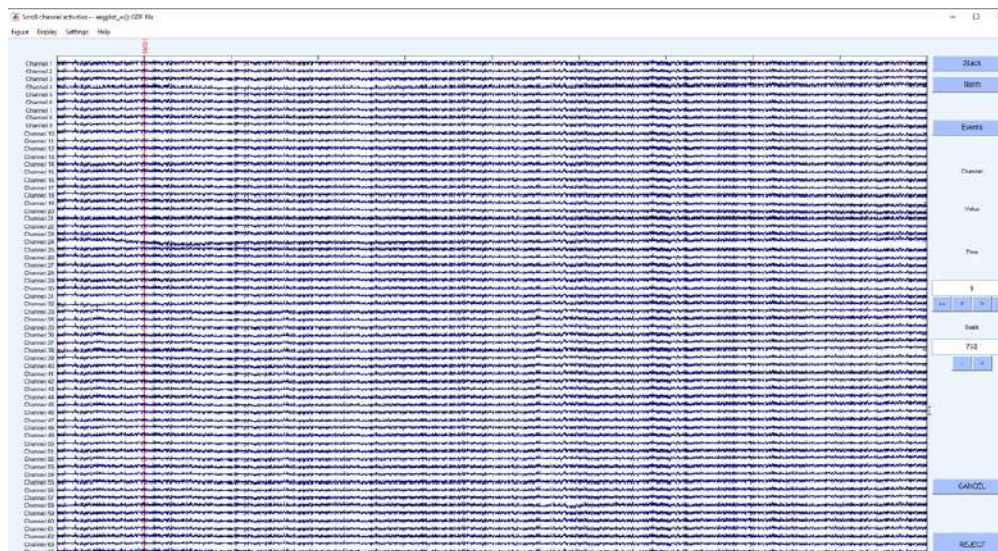
6.- El resultado es una nueva ventana con la información del EEG que se irá modificando según vayamos realizando el proceso de limpieza de la señal. Comprobamos que la frecuencia de muestreo se corresponde con la que utilizemos en el estudio. De no ser así, se procedería a la exclusión del EEG.



7.- Comprobamos que el EEG se ha grabado correctamente siguiendo la ruta [Plot -> Channel data \(scroll\)+](#).

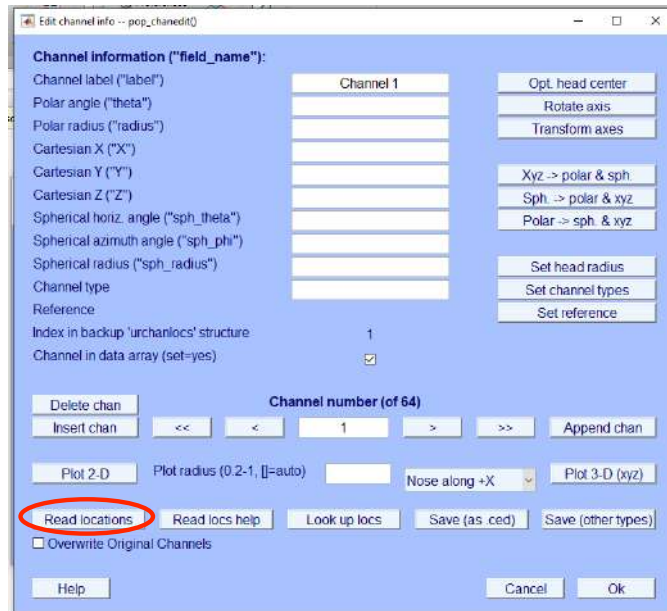
8.- El resultado del paso 7 es una imagen que muestra los canales del EEG.

A continuación, se muestra un ejemplo de un EEG “correcto”. En el caso de encontrar alguna alteración y tener un segundo EEG del mismo paciente con la misma prueba, volvemos al paso 3 para abrir otro EEG y elegir el que muestre menos alteraciones. En el caso de encontrarnos con un EEG con alteraciones significativas en varios electrodos, procedemos a descartarlo.



Protocolo limpieza EEGs

9.- Para localizar los canales del EEG siguiendo la ruta [Edit -> Channel locations](#). En la siguiente pantalla presionamos “ok” y obtenemos la siguiente ventana.

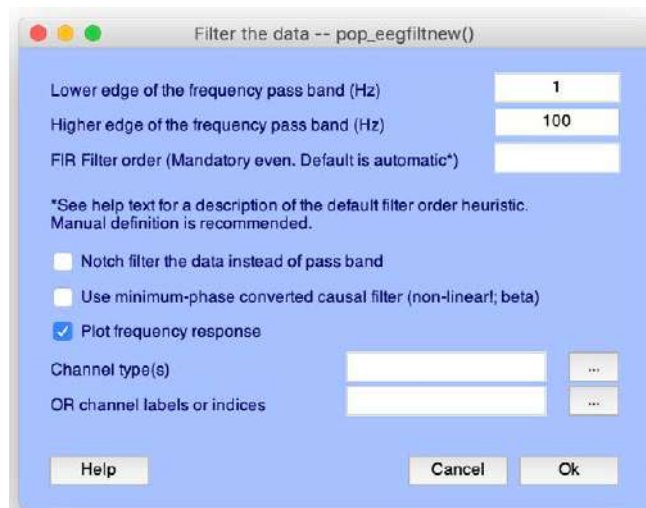


Presionamos “Read locations” para abrir el fichero con la localización de los canales. Una vez abierto dicho fichero, presionamos “ok”. En la pantalla emergente presionamos nuevamente “ok” para implementar dicha localización.

Para comprobar que hemos realizado correctamente la localización de canales, veremos que en la ventana del paso 6, en la línea “Channel locations” aparecerá “Yes”. En esta ventana iremos comprobando cada uno de los pasos realizados.

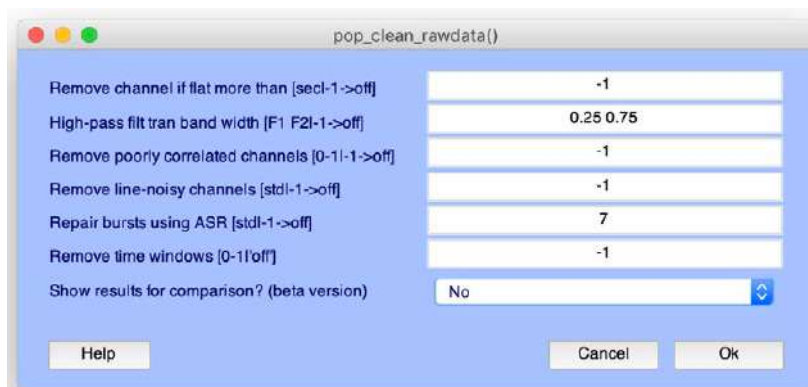
10.- Para proceder al filtrado de la señal, seguimos la ruta [Tools -> Filter the data -> Basic FIR filter \(new, default\)](#). En la ventana emergente introducimos los valores del filtro 1-100 Hz según la imagen.

Protocolo limpieza EEGs



En la pantalla emergente presionamos “ok”.

11.- Continuamos reconstruyendo la señal y eliminando las componentes problemáticas mediante la ruta [Tools](#) -> [Clean continuous data using ASR](#). Aparecerá una ventana que rellenaremos del siguiente modo:

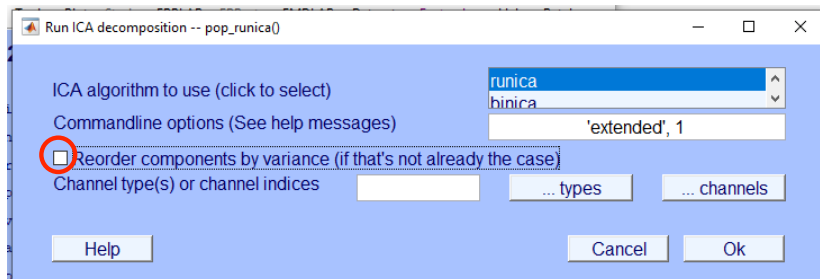


A continuación, presionamos “ok”.

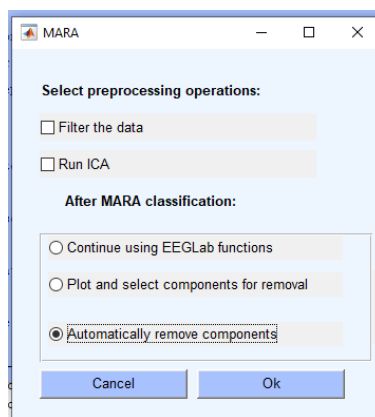
12.- Repetimos el filtrado del paso 10 entre 1-31 Hz ó 1-50 Hz según convenga para el análisis.

Protocolo limpieza EEGs

13.- Con la ruta **Tools -> Run ICA** identificaremos las componentes independientes, haciendo especial hincapié en la banda de frecuencias que introdujimos en el paso anterior. En la ventana emergente, desactivaremos la opción “Reorder components by variance” y presionamos “ok”.



14.- Para eliminar los artefactos de la señal seguimos la ruta **Tools -> IC Artifact Classification (MARA) -> MARA Classification**. En la ventana emergente seleccionamos la opción “Automatically remove components” y presionamos “ok”.



15.- Si al hacer MARA aparece un mensaje de error como el siguiente:



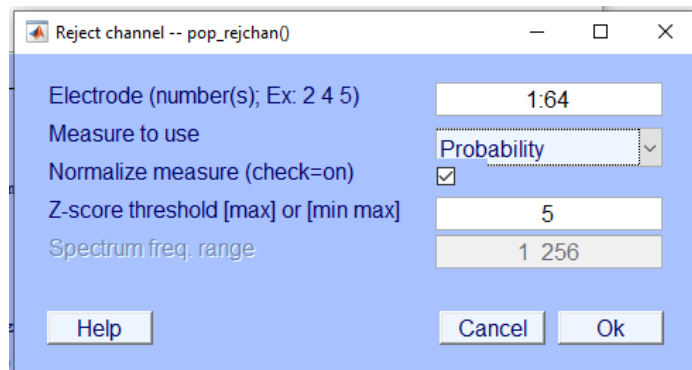
Volveremos al paso 3 y seleccionaremos un nuevo EEG.

16.- Una vez hemos pasado MARA, aceptamos las tres pantallas emergentes presionando en cada una de ellas “ok”.

17.- En este punto, comprobamos el resultados de los pasos dados repitiendo el paso 7, ya que podríamos encontrar alteraciones significativas.

Protocolo limpieza EEGs

18.- Continuamos con la ruta [Tools -> Automatic channel rejection](#), y en la ventana emergente seleccionamos la opción “Probability” y presionamos “ok”



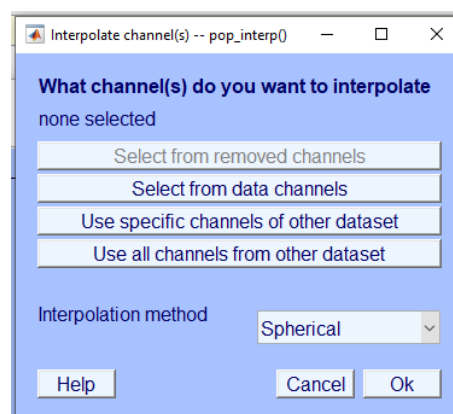
19.- Tras este paso, en la ventana de comandos aparecerá una lista de todos los electrodos y algunos de ellos señalados como “\*Bad\*”. Anotamos aquellos que tienen dicha etiqueta.

```

15 F8 0.73
16 FT9 0.62
17 FT7 0.71
18 FC5 0.66
19 FC3 5.64 *Bad*
20 FC1 4.39
21 FCz 0.10
22 FC2 -0.42
23 FC4 0.45
24 FC6 -3.69

```

20.- Ahora reconstruiremos aquellos electrodos “\*Bad\*” siguiendo la ruta [Tools -> Interpolate electrodes](#) donde encontraremos la siguiente ventana, donde seleccionaremos “Select from data channels”



Protocolo limpieza EEGs

Aparecerá una lista con todos los electrodos y seleccionaremos aquellos que tenían la anotación “Bad”. También podemos eliminar aquellos electrodos anómalos que el programa no haya detectado y nosotros consideremos tras el paso 15. Comprobamos que el método de interpolación sea “Spherical” y presionamos “ok”.

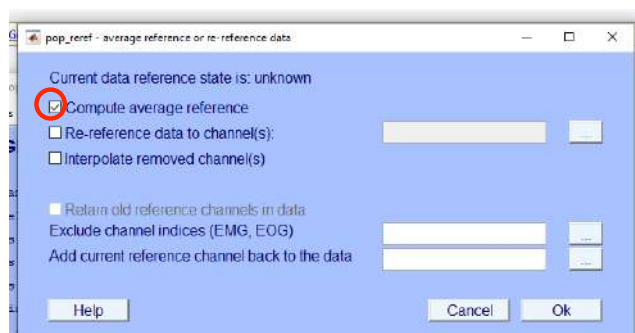
21.- Puede darse el caso de que no encontremos ningún electrodo y tampoco haya electrodos que consideremos anómalos. En dicho caso, omitiremos el paso 18.

22.- Finalmente, guardaremos el archivo de dos maneras.

Para guardar el archivo con una extensión tipo “.dat” seguimos la secuencia [File -> Export -> Data and ICA activity to text file](#). En la ventana emergente, seleccionamos “Browse” para guardar el archivo donde deseemos y presionamos “ok”.

Para guardar el archivo en extensión “.set” seguiremos la secuencia [File -> Save current dataset\(s\)](#) y elegimos la carpeta donde lo deseemos guardar.

23.- Según el estudio se referenciarán los electrodos con respecto a la media, para ello seguiremos la ruta [Tools -> Re-reference](#) y en la ventana emergente seleccionamos la opción “Compute average reference” y presionamos “ok”



24.- Guardamos repitiendo el paso 22.

## **Apéndice D**

# **Tablas de la Revisión Sistemática desde el Punto de Vista Médico**

ITEM	SECTION	TOPIC	CHECKLIST ITEM
1	Title	Nature of study	Identify the report as introducing a predictive model.
2	Abstract	Structured summary	Background. Objectives. Data sources. Performance metrics of the predictive models (in both point estimates and confidence intervals). Conclusion including the practical value of the developed predictive model or models.
3	Introduction	Rationale	Identify the clinical goal. Review the current practice and prediction accuracy of any existing models.
4		Objectives	State the nature of study being predictive modeling, defining the target of prediction. Identify how the prediction problem may benefit the clinical goal.
5	Methods	Describe the setting	Identify the clinical setting for the target predictive model. Identify the modeling context in terms of facility type, size, volume, and duration of available data.
6		Define the prediction problem	Define a measure for the prediction goal. Identify the problem to be prognostic or diagnostic. Determine the form of the prediction model: classification, regression, or survival prediction. Explain practical costs of prediction errors. Define quality metrics for prediction models. Define the success criteria for prediction.
7		Prepare data for model building	Identify relevant data sources and quote the ethics approval number for data access. State the inclusion and exclusion criteria for data. Describe the time span of data and the sample or cohort size. Define the observational units on which the response variable and predictor variables are defined. Define the predictor variables. Describe the data preprocessing performed, including data cleaning and transformation. State any criteria used for outlier removal. State how missing values were handled. Describe the basic statistics of the dataset as the ratio of positive to negative classes for a classification problem. Define the model validation strategies. Define the validation metrics. For classification problems, the metrics should include sensitivity, specificity, positive predictive value, negative predictive value, area under the ROC curve, and calibration plot.
8		Build the predictive model	Identify independent variables that predominantly take a single value. Identify and remove redundant independent variables. Identify the independent variables that may suffer from the perfect separation problem. Assess whether sufficient data are available for a good fit of the model. Determine a set of candidate modeling techniques. If only one type of model was used, justify the decision for using that model. Define the performance metrics to select the best model.



ITEM	SECTION	TOPIC	CHECKLIST ITEM
9	Results	Report the final model and performance	Report the predictive performance of the final model in terms of the validation metrics specified in the methods section. If possible, report the parameter estimates in the model and their confidence intervals. When the direct calculation of confidence intervals is not possible, report nonparametric estimates from bootstrap samples. Comparison with other models in the literature should be based on confidence intervals. Interpretation of the final model. If possible, report what variables were shown to be predictive of the response variable. State which subpopulation has the best prediction and which subpopulation is most difficult to predict.
10	Discussion	Clinical implications	Report the clinical implications derived from the obtained predictive performance.
11		Limitations of the model	Discuss the following potential limitations: Assumed input and output data format. Potential pitfalls in interpreting the model. Potential bias of the data used in modeling. Generalizability of the data.
12		Unexpected results during the experiments	Report unexpected signs of coefficients, indicating collinearity or complex interaction between predictor variables.

Tabla D.1: Ítems a incluir al informar acerca de modelos predictivos en investigación biomédica. Fuente: [235].

Ref.	Objective	Subjects	Sex (F/M)	Age	Hoehn-Yahr (n)	UPDRS	Disease Duration	EEG parameters	EEG Pre-Processing
[255]	Selection of the best QEEG characteristics to identify different levels of cognitive impairment in PD.	118 PD with LD classified into 5 groups according to the severity of the disease. G1: n=28, G2: n=33, G3: n=43, G4: n=5, G5: n=9.	G1 = 8/20 G2 = 9/24 G3 = 16/27 G4 = 3/2 G5 = 0/9	G1 = 60.54 ± 8.75 G2 = 66.09 ± 6.65 G3 = 67.04 ± 7.94 G4 = 73.19 ± 5.29 G5 = 67.56 ± 5.51	G1 = 1.93 ± 0.4 G2 = 2.14 ± 0.55 G3 = 2.21 ± 0.59 G4 = 2.40 ± 0.55 G5 = 2.00 ± 0.97	G1 = 26.00 ± 11.73 G2 = 29.55 ± 12.22 G3 = 28.74 ± 11.44 G4 = 31.00 ± 12.79 G5 = 29.00 ± 18.87	G1 = 7.75 ± 5.29 G2 = 8.36 ± 7.49 G3 = 8.81 ± 5.02 G4 = 6.60 ± 3.58 G5 = 12.00 ± 6.56	122-channel EEG recorded during 10 min in Resting state with 512 Hz sample rate.	Average Reference and 0.1-100 Hz bandwidth filter. Ocular artifacts were corrected and 50 Hz filter was applied. Periods of drowsiness were removed, and semi-automatic rejection of artifacts was performed to eliminate muscle activity. Each channel was divided into 4s epochs. At least 20 segments were used for the analysis.
[256]	Selection of the QEEG parameters that best distinguish between controls and PD.	50 PD and 41 controls.	PD: 17/33 HC: 20/21	PD: 68.8 ± 7 HC: 71.1 ± 7	Not specified.	Not specified.	5.3 ± 5.1	256-channel EEG recorded during 12 min in Resting state with eyes-closed and 500 Hz sample rate.	3 min of EEG were constructed with segments of at least 30 s without artifacts. 0.5-70 Hz filter was applied. An inverse Hanning window was used to join segments. It was referenced with respect to mean and defective channels were interpolated with the spherical spline method. 'Runica' was used with default settings to remove further artifacts.

Ref.	Objective	Subjects	Sex (F/M)	Age	Hoehn-Yahr (n)	UPDRS	Disease Duration	EEG parameters	EEG Pre-Processing
[257]	Classification of patients vs controls for the diagnosis of PD.	20 PD with LD and 20 controls. All right-handed.	PD: 10/10 HC: 11/9	PD: [45, 65] HC: 58.1 ± 2.95	1: (2) 2: (11) 3: (7)	Not specified.	5.75 ± 3.52	14-channel EEG recorded during 5 min in Resting state with 128 Hz sampling rate.	Epochs of 2s were segmented and a threshold technique was applied at $\pm 100 \mu V$ . A sixth order Butterworth band-pass filter was applied with direct reverse filtering technique at 1-49 Hz.
[258]	Classification of patients with RBD and controls. Some of the patients with RBD were eventually diagnosed with PD and dementia.	118 RBD and 74 controls. 14 RBD became PD. No direct patient data.	Not specified.	Not specified.	Not specified.	Not specified.	Not specified.	14-channel EEG at 256 Hz sampling rate in Resting state with open-eye periods followed by closed-eye periods.	The first EEG of each patient was considered baseline. A band-pass filter was passed at 0.3-100 Hz with a notch filter at 60 Hz to minimize noise from the power line. It was also filtered at 4-44 Hz. The signals were referenced to the ears.
[261]	Classification of patients with PD in the early stages of the disease vs. controls.	30 PD and 30 controls.	Not specified.	PD: [50, 70] HC: [50, 70]	1 – 1.5	Not specified.	Not specified.	2-channel EEG recorded during 30 min for the flexion and extension of the wrist.	Not specified.
[262]	Classification of patients with PD in the early stages of the disease vs. controls using various algorithms.	100 PD and 100 controls.	Not specified.	PD: [50, 70] HC: [50, 70]	1 – 1.5	Not specified.	Not specified.	2-channel EEG recorded during 30 min for the flexion and extension of the wrist.	5-50 Hz band-pass filter was applied

Ref.	Objective	Subjects	Sex (F/M)	Age	Hoehn-Yahr (n)	UPDRS	Disease Duration	EEG parameters	EEG Pre-Processing
[259]	Classification of patients with neurological diseases vs. controls to search for spectral equivalence between various neurological and neuropsychiatric disorders with TCD.	31 PD and 264 controls.	PD: 14/17 HC:112/152	PD: 56.62 ± 12.32 HC: 49.51 ± 12.54	1 - 3	43.44 ± 15.53	Not specified.	19-channel EEG recorded during 5 min in Resting state with eyes-closed and 1024 Hz sampling rate.	The data was referenced to the ears, and the impedances were <5 kΩ at all electrodes during the recording. High-pass filter at 0.15 Hz and low-pass filter at 200 Hz were used. Data was resampled at 128 Hz, and band-pass filtered at 2-44 Hz. Artifacts were manually removed.
[321]	Classification of PD and controls to demonstrate the utility of EEG as biomarker for PD.	21 PD (18 for DAT PET) OFF LD during 12 h and 25 controls (24 for DAT PET).	PD: 7/14 HC: 9/16	PD: 62.7 ± 7.32 HC: 54.6 ± 10.5	2.07 ± 0.39	PD: 31.00 ± 10.37 HC: 0.83 ± 1.27	Not specified.	20-channel EEG and 4 additional channels for ECG, EMG or EOG. 2 recordings per patient are recorded during 5 min in Resting state with eyes-closed and 256 Hz sampling.	Not specified.
[254]	Selection the best classifier of PD vs controls using the minimum number of HOS features.	20 PD with LD and 20 Controls. All right-handed.	PD: 10/10 HC: 11/9	PD: 59.05 ± 5.64 HC: 58.10 ± 2.95	1: (2) 2: (11) 3: (7)	Not specified.	5.75 ± 3.52	14-channel EEG recorded during 5 min in Resting state with eyes-closed and 128 Hz sampling rate.	Threshold technique at 80μV. Butterworth sixth-order band-pass filter at 1-49 Hz. Each channel is separated into 2 s epochs with 50 % overlap.

Tabla D.2: Summary of the clinical variables, such as objectives, subjects, EEG recording protocol, EEG cleaning protocol and dataset pre-processing. Fuente: [234].



Ref.	Objective	Features	Models	Model Parameters	Training	Validation Strategy	Metrics	Best Result
[255]	Selection of the best QEEG characteristics to identify different levels of cognitive impairment in PD.	The relative and absolute spectral power was obtained for each epoch using a FFT and a 50% overlap for the Delta, Theta, Alpha and Beta bands. Moreover, a division into 5 ROI was performed. For each case, high and low electrode density were considered. A statistical dependency study with analysis of variance and selection of characteristics with Pearson's correlation method was carried out.	SVM, KNN	SVM: Gaussian kernel KNN: k = 9 and the Euclidean distance as a metric.	The dataset is randomly split into k folds (for this case k=5). k-1 folds were used to train the models and the rest fold was the testing set. The dataset used for the k-fold cross-validation was the set with n=100.	Two validation strategies. First: divide the full dataset into training set with n=100 and validation set with n=18. Second: the training set was used for 5-fold cross-validation.	Accuracy.	SVM: Accuracy = $87 \pm 3.5$ KNN: Accuracy = $88 \pm 2.8$ Both were achieved for the relative power with low-electrode density. Groups with few patients had worse results.
[256]	Selection of the QEEG parameters that best distinguish between controls and PD.	Ten brain regions were considered with 79 different measurements. All of the features were extracted from the frequency spectrum.	RF, SVM, DT, LR and LR with LAS-SO	SVM: Non-linear kernels such as RBF were used.	10-fold cross-validation was considered and optimization was carried out for tuning parameters.	10-fold cross-validation.	Accuracy, AUC.	The most significant models were: RF: Accuracy = 78, AUC = 0.8 LR with LASSO: AUC = 0.76

Ref.	Objective	Features	Models	Model Parameters	Training	Validation Strategy	Metrics	Best Result
[257]	Classification of patients vs controls for the diagnosis of PD.	There was no pre-processing of data.	CNN	13 layers with 4 1D convolution layers, 4 max-pooling layers and 3 fully connected layers. Adam optimizer (learning_rate = $10^{-4}$ ). Activation function Relu and the last one Softmax. Dropout of 0.5.	10-fold cross-validation with 9 parts for training and 1 for testing. 20% of the training data was also used for validation.	Two validation strategies. First: 10-fold cross-validation with all the data. Second: 20% of the training data was also used for validation at the end of each epoch.	Accuracy, Sensitivity, Specificity.	CNN: Accuracy = 88.25, Sensitivity = 84.71, Specificity = 91.77
[258]	Classification of patients with RBD and controls. Some of the patients with RBD were eventually diagnosed with PD and dementia.	Several spectrograms per subject were generated, each of them of 20 s and artifact-free, until about 2.5 min per patient was obtained. The data was centered and normalized to unit variance.	CNN, RNN	CNN: 4 hidden-layer convolutional net with pooling. Dropout was used as regularization, max-pooling layers and using a cross-entropy loss function. RNN: with LSTM and GRU, with 3 cells which 32 units each. Dropout was used.	For training, datasets were balanced by random replication preserving the distribution of the subjects.	Leave-pair out (LPO) cross-validation, where one subject from each class was left out for validation.	Accuracy, AUC.	The results for controls vs. PD: CNN: Accuracy = $79 \pm 1$ AUC = $0.87 \pm 0.1$ RNN: Accuracy = $81 \pm 1$ AUC = $0.87 \pm 0.1$ In RNN, there was no difference between LSTM and GRU.

Ref.	Objective	Features	Models	Model Parameters	Training	Validation Strategy	Metrics	Best Result
[261]	Classification of patients with PD in the early stages of the disease vs. controls.	EEG: Shannon Entropy, Lyapunov and Inverse Lyapunov Exponent EMG: Power, Standard deviation, Root mean square, Variance, Waveform length, Modified median, and mean frequency.	MLP	Back Propagation was used as the learning algorithm and 'trainlm' was used as the training function. Sigmoid transfer function was used for the hidden layer.	The dataset was divided into: training 70 %, validation 15 % and testing 15 %.	Validation with validation and test sets.	Accuracy.	MLP with inputs: EEG: Accuracy = 62 EMG: Accuracy = 73 EEG+EMG: Accuracy = 98.8
[262]	Classification of patients with PD in the early stages of the disease vs. controls using various algorithms.	EEG: Lyapunov and Inverse Lyapunov Exponent, Shannon Entropy EMG: Power, Standard deviation, Root mean square, Variance, Waveform length, Modified median, and mean frequency.	MLP	3 algorithms were tested: Gradient Descent algorithms (trainingd, trainingdm), Conjugate Gradient algorithms (traininscg, traincgp), and Quasi-Newton algorithms (trainbfg, trainlm). Sigmoid function was used in the hidden layer. The number of hidden neurons was checked for 5, 7, 9,10,20,30.	The dataset was divided into: training 70 %, validation 15 % and testing 15 %.	Validation with validation and test sets.	Accuracy RMSE, R value.	ANN with trainlm and 10 neurons: Accuracy = 100 RMSE = $4.03 \cdot 10^{-3}$ R value = 0.9998



Ref.	Objective	Features	Models	Model Parameters	Training	Validation Strategy	Metrics	Best Result
[259]	Classification of patients with neurological diseases vs. controls to search for spectral equivalence between various neurological and neuropsychiatric disorders with TCD.	The power spectrum was calculated for each subject and the five frequency bands (delta, theta, alpha, beta, and gamma) were considered for each ROI.	SVM	The default settings were used as the running parameters.	10-fold cross-validation was performed on the full data set with 90 % of the data for training and 10 % of the data for testing. The distribution of the patients was kept.	10-fold cross-validation. Moreover, controls with obesity were used to validate the model.	Accuracy, TPR, FPR, ROC, MAE, RMSE.	The results for controls vs. PD: Accuracy = $94.34 \pm 1.81$ TPR = $0.93 \pm 0.02$ FPR = $0.11 \pm 0.01$ ROC = $0.95 \pm 0.02$ MAE = $0.07 \pm 0.02$ RMSE = $0.16 \pm 0.02$
[321]	Classification of PD and controls to demonstrate the utility of EEG as biomarker for PD.	Coherence analysis with 2 s windows with 50 % overlap was carried out. Pearson's correlation was calculated to assess the relationships between coherence and disease severity. The relative and absolute PSD was calculated at 1-40 Hz. Only 14 EEG-based features were finally used.	DFA	A linear DFA was used to build the classifier. The classifier input was selected by utilizing the step-wise discriminant analysis procedure in the SPSS software package.	Not specified.	Cross-validation.	Accuracy, Sensitivity, Specificity, PPV (precision), NPV.	Accuracy = 95.24 Sensitivity = 94.74 Specificity = 95.65 PPV = 94.74 NPV = 95.65 An excessive coherence was observed in the Beta and Gamma bands for PD.

Ref.	Objective	Features	Models	Model Parameters	Training	Validation Strategy	Metrics	Best Result
[254]	Selection the best classifier of PD vs controls using the minimum number of HOS features.	For each epoch a total of 13 HOS characteristics were calculated. The Student's t-test is also obtained to determine the importance of the characteristics.	DT, KNN, FKNN, NB, PNN, SVM	FKNN: Euclidean distance, $m = 1.24$ and $k = 3$ . KNN: $k = 2$ and Euclidean distance. PNN: exponential activation function and $\sigma = 0.284$ . SVM: polynomial kernel functions of order 2 and 3, RBF and linear kernels.	The characteristics were added one by one to each classifier until maximum precision is achieved.	10-fold cross-validation.	Accuracy, Sensitivity, Specificity, Precision, F-score.	The best model was SVM with RBF kernel: Accuracy = $99.62 \pm 0.58$ Sensitivity = $100 \pm 0.0$ Specificity = $99.25 \pm 0.53$ Precision = $99.38 \pm 0.47$ F-score = $0.98 \pm 0.05$

Tabla D.3: Summary of the model parameters, such as features extracted, models used, model architecture, training and validation methods, results and metric used for the articles included in this review. The columns with the references and objectives of the selected articles are included again for the sake of readability. Fuente: [234].

## **Apéndice E**

# **Tablas de la Revisión Sistemática desde el Punto de Vista de las Ciencias Computacionales**

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[303]	Classification of PD patients vs. controls.	Subjects: 20 PD and 21 controls Age: PD: 67.6 ± 7.0 HC: 67.5 ± 6.4	HY scale: 1-2 UPDRS: 23.5 ± 9.8 Disease Duration: 7.6 ± 4.3 Medication: ON and OFF	64-channel EEG recorded during 1 min in resting state at 1 kHz. 27 electrodes were considered. Impedances were kept below 15 kΩ. The EEG was divided into non-overlapping 3 s segments.	The segments of EEG were introduced as input.	CNN + RNN	2 1D-conv layers with 64 each, LSTM with 80 cells and a fully connected layer with 50 units. The activation function was sigmoid, the loss function was "binary crossentropy" and the optimizer was Adam with learning rate = 0.001.	Data were split into 80 % for training and 20 % for test sets.	Results without medication: Accuracy = 96.9 Precision = 100 Recall = 93.4
[292]	Classification of PD patients vs. controls.	Subjects: 20 PD and 21 controls Age: PD: 67.6 ± 7.0 HC: 67.5 ± 6.4	HY scale: 1-2 UPDRS: 23.5 ± 9.8 Disease Duration: 7.6 ± 4.3 Medication: ON and OFF	64-channel EEG recorded during 1 min in resting state at 1 kHz. 27 electrodes were considered. Impedances were kept below 15 kΩ. A band-pass filtered was used at 1-55 Hz, and re-referenced to average reference. Artifacts were removed using ICA. The data were standardized and segmented into 1 s or 2s epochs.	For the CNN/CRNN models, 1 s segments without overlapping were introduced. For the others, two datasets were considered. First, with 2 s epochs and 13 features of HOS. Second, 794 time-series features in 1 s epochs using the bands theta, alpha beta and gamma. Top significant features were selected through ANOVA.	CNN + RNN, CNN, KNN, SVM, RF	The final architecture for KNN, SVM and RF were not specified. CNN+RNN: 2 1D-conv (kernel size 3, filters 32 and 64), max-pooling, a GRU cells with 35 units, and 2 fully connected layers (time distributed and dense with 35 units). Dropout were used. Activation function was ReLU and a softmax for the final layer. Optimizer Adam (learning rate = 0.001) and a binary cross-entropy as loss. CNN: 4 consecutive blocks, with 8, 12, 12 and 16 filters in each 1D-conv layer (kernel size = 9). Max-pooling layer. 3 fully connected layers with 30 and 5 units, activation function ReLU and softmax for the last one. Adam optimizer (learning rate = 0.001), and binary cross-entropy as loss function.	Nested cross-validation (inner 5-fold for hyperparameter tuning and outer 10 fold)	The CNN + RNN model obtained: Accuracy = 99.2 Precision = 98.9 Recall = 99.4 F1-score = 99.2 AUC = 99.2

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[257]	Classification of PD patients vs. controls.	Subjects: 20 PD and 20 controls. Age: PD: (45-65) HC: 58.1 ± 2.95	HY scale: 1: n=2; 2: n=11; 3: n=7 UPDRS: Not specified Disease Duration: 5.75 ± 3.52 Medication: ON	14-channel EEG recorded during 5 min in resting state at 128 Hz. Epochs of 2 s were segmented and a threshold technique was applied at ±100 μV. A band-pass filter was used at 1-49 Hz.	The EEG signals were the input.	CNN	13 layers with 4 1D-conv layers, 4 max-pooling layers, and 3 fully connected layers. Adam optimizer (learning rate = 10 <sup>-4</sup> ). Activation function Relu and the last one softmax. Dropout of 0.5.	10-fold cross-validation with stratified data. 20 % of the training data was also used for validation at the end of each epoch.	The results were: Accuracy = 88.25 Sensitivity = 84.71 Specificity = 91.77
[322]	Classification of PD patients with MCI vs. patients without MCI.	Subjects: 27 PD with MCI and 43 PD without MCI Age: MCI: (53-84) Non-MCI: (46-82)	HY scale: MCI: 0-5 Non-MCI: 0-4 UPDRS: MCI: 0-41 Non-MCI: 0-36 Disease Duration: MCI: 0-23 Non-MCI: 0-17 Medication: Not specified	256-channel EEG recorded during 12 min in resting state with eyes-closed at 1 kHz. 214 electrodes were associated with 10 ROI. Signals were filtered at 0.5-70 Hz, with a 50 Hz notch, and an inverse Hanning window was used to stitch together segments to get 3 min of EEG data. Artifacts were removed. The average of all "good" channels was used to reference the signals to an average value.	The spectral power was calculated in 10 ROI and globally for 6 frequency intervals resulting in 66 spectral features. The PLI was calculated between all pairs of ROI resulting in 330 connectivity features.	RF	The standard implementation in R was applied.	Data were split into 70 % for training and 30 % for test sets. 20 runs of 5-fold cross-validation.	The results with the combination of both features. ROI Train AUC = 0.73 ± 0.16 Test AUC = 0.71 Without ROI: Train AUC = 0.7 ± 0.14 Test AUC = 0.875

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[256]	Selection of the QEEG parameters that best distinguish between controls and PD patients.	Subjects: 50 PD and 41 controls. Age: PD: $68.8 \pm 7$ HC: $71 \pm 7$	HY scale: Not specified UPDRS: Not specified Disease Duration: $5.3 \pm 5.1$ Medication: Not specified	256-channel EEG recorded during 12 min in resting state with eyes-closed at 500 Hz. 3 min were constructed with segments of at least 30 s without artifacts, and a 0.5-70 Hz filter was applied. An inverse Hanning window was used to join segments. It was referenced with respect to mean and defective channels that were interpolated with the spherical spline method. Artifacts were removed.	10 brain regions were considered with 79 different measurements. All features were extracted from the frequency spectrum.	RF, SVM, DT, LR, LR with LASSO	SVM: RBF kernel. Optimization was carried out for tuning parameters.	10-fold cross-validation.	The most significant models were: RF: Accuracy = 78 AUC = 0.8 LR with LASSO: AUC = 0.76

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[323]	Cognition classification of patients with PD	Subjects: 20 PD H-COG, 20 PD L-COG and 72 inter-COG Age: H-COG: 59.5 (54.6-66.4) L-COG: 67.8 (60.1-72.1) Inter-COG: 63.5 (57.7-68.0)	HY scale: Not specified UPDRS: H-COG: 18.5 L-COG: 23 Inter-COG: 20.5 Disease Duration: H-COG: 11.2 ± 4.5, L-COG: 10.9 ± 5.1, Inter-COG: 11.8 ± 8.0 Medication: ON	21-channel EEG recorded in resting state with eyes-closed. Data were re-referenced. After visual confirmation of artefact-free signals, 5 consecutive non-overlapping 4096-point (8.192 s) epochs were selected. Recordings with less than five epochs were excluded.	16674 features were extracted per patient. Feature-selection was performed using a Boruta algorithm. Small feature sets were considered as input.	RF	The hyperparameters were optimized with a variant of Bayesian Optimization technique called Mixed Integer Parallel Efficient Global Optimization (MIP-EGO) for mixed-integer categorical search spaces.	10-fold cross-validation. Additional assessment with a combination of cross-validation and split-sample validation.	Using all features from all cross-validation runs in L-COG vs. H-COG: Accuracy: 92 Sensitivity: 90 Specificity: 94

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[279]	Prediction of FOG episodes	Subjects: 16 PD Age: PD: 70.88 ± 6.92	HY scale: 2.75 ± 0.61 UPDRS: 42.50 ± 14.25 Disease Duration: 8.63 ± 6.58 Medication: Not specified	4-channel EEG during 404 FOG episodes in structured series of Timed Up and Go tasks at 500 Hz, with a duration between 1 and 220 s. Segments with artefacts were rejected using visual inspection. 1902 selected samples of data were filtered using band-pass (0.5-60 Hz) and band-stop (50Hz) Butterworth IIR filters with zero phase shift. The data were normalized and ICA was applied.	DTF was applied. The non-parametric Wilcoxon Sum Rank Test was used to select the most significant feature. A p-value <0.05 and r-value >0.25 were chosen for further processing.	BNN	3 layers Back Propagation BNN was used as a classifier with Bayesian regularisation and Levenberg-Marquardt optimization.	11 patients were randomly chosen. Fifty runs of random training/validation (50%) and test (50%). Remaining 5 patients were considered for test.	In train set: Sensitivity: 82.65 Specificity: 86.60 In test set: Sensitivity: 85.86 Specificity: 80.25



Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[324]	Classification of early-stage PD patients vs. HC	Subjects: 19 PD and 30 HC Age: PD: $63.7 \pm 7.8$ HC: $64.4 \pm 6.2$	HY scale: $1.8 \pm 0.6$ UPDRS: $20.1 \pm 8.8$ Disease Duration $1.1 \pm 0.9$ Medication: OFF	64-channel EEG recorded at 250 Hz while performing visual Go/No-Go and AOB during 15 min in cognitive tasks. Signals were referenced to average mastoid electrodes and band-pass filtered in 4 bands with overlap. Then they were cut into epochs based on stimulus onset and response, and averaged across trials of the same condition.	199 features extracted by the BNA analysis from the HC and PD groups were used. In each iteration, the FPR feature selection method was applied.	LR	Not specified	10-fold cross-validation with stratified data.	Cross validation results in discriminating HC vs. early stage PD: AUC: 79 Sensitivity: 74 Specificity: 73
[297]	Classification of PD patients vs. controls using 6 emotional stimuli.	Subjects: 20 PD and 20 controls Age: PD: (40-65) HC: (40-65)	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	14-channel EEG recorded during 6 emotional stimuli. The signals were segmented into 10 s epochs with overlapping of 75%. Then, pass-band elliptic filters were used to obtain the alpha, beta, and gamma bands.	3 spectral features were calculated, that is, Spectral Entropy (SEN), Spectral Energy-Entropy (SEEN), and SpectralTeager Energy-Entropy (STEEN) for each band.	PNN, KNN, SVM	KNN: k between 1 and 10. PNN: A multilayer feed-forward network with 4 layers was considered. It used an exponential activation function with $\sigma$ ranged from 0.55 to 0.65.	Not specified	The best accuracy for each emotion: Happiness: 96.8 with PNN and SEEN Sadness: 90.2 with KNN and SEN Fear: 95.07 with KNN and SEEN Anger: 91.4 with SVM and SEEN Surprise: 94.53 with KNN and SEEN Disgust: 88.18 with SVM and SEEN.

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[300]	Classification of PD patients vs. controls	<p>Subjects: Dataset1: 15 PD and 16 HC Dataset2: 20 PD and 20 HC</p> <p>Age: Dataset1: PD: <math>63.5 \pm 9.6</math> HC: <math>63.2 \pm 8.2</math> Dataset2: PD: <math>58.1 \pm 2.95</math> HC: <math>59.05 \pm 5.94</math></p>	<p>HY scale: Dataset1: 2 and 3 Dataset2: 1-3</p> <p>UPDRS: Not specified</p> <p>Disease</p> <p>Duration: Dataset1: <math>4.5 \pm 3.5</math> Dataset2: <math>5.75 \pm 3.52</math></p> <p>Medication: Dataset1: ON and OFF</p>	<p>Dataset1: 32-channel EEG recorded during 3 min in resting state at 512 Hz. Artifacts were manually removed and a highpass filter at 0.5 Hz was used.</p> <p>Dataset2: 14-channel EEG during 5 min in resting state at 128 Hz. Signals were segmented into 2 s windows. Eye blinking artifacts were removed with a threshold at <math>\pm 100V</math> and a forward and reverse filtering technique using sixth-order Butterworth filter at 1-49 Hz was used.</p>	<p>EEG signals are subjected to SPWVD to obtain TFR. After Kaiser window selection and resizing, the two-dimensional plots are fed to the model. Experiments are carried out by maintaining the same setup for both datasets.</p>	CNN	<p>4 2D-Conv Layers, 2 MaxPooling layers, 2 Fully Connected Layers (50 and 32 neurons), and a softmax layer. The number of filters selected were 96, 32, 16, and 8. A filter size of <math>7 \times 7</math>, <math>5 \times 5</math>, and <math>3 \times 3</math> with a stride of 2 was used. A dropout of 0.5 was considered. Adam optimizer was used with learning rate = <math>10^{-4}</math>.</p>	10-fold cross validation	<p>The best results were for dataset1 HC vs. PD ON medication: Accuracy: 100 Specificity: 100 Sensitivity: 100 Precision: 100 F-1 score: 100</p>
[301]	Classification of controls vs. PD patients with ON and OFF medication	<p>Subjects: 15 PD and 16 HC</p> <p>Age: PD: <math>63.2 \pm 8.2</math> HC: <math>63.5 \pm 9.6</math></p>	<p>HY scale: 2 or 3 UPDRS: ON: <math>33.7 \pm 10.9</math> OFF: <math>45.5 \pm 13.0</math></p> <p>Disease</p> <p>Duration: Not specified</p> <p>Medication ON and OFF</p>	<p>32-channel EEG during 3 min focusing on an image at 512 H.</p>	<p>EEG recordings were split in half and then converted into spectrograms using Gabor transform.</p>	CNN	<p>2D-Conv (16 filters kernel <math>5 \times 5</math>, ReLu), Dropout (0,2), 2D-Conv (32 filters, kernel <math>3 \times 3</math>, ReLu), MaxPooling, Flatten, Dense (unit size of 512, ReLu), Dropout (0,7), Dense (unit size of 3 for Softmax and 1 for sigmoid). Adam optimizer was used with a learning rate = 0.001 and a decay rate of 0.01.</p>	10-fold cross validation	<p>Results for 3 class classification: Accuracy: <math>99.46 \pm 0.73</math> Precision: <math>99.48 \pm 0.01</math> Sensitivity: <math>99.46 \pm 0.01</math> F1 Score: <math>99.46 \pm 0.01</math></p>

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[280]	Determine the optimal montage to detect FOG.	Subjects: 7 PD Age: Not specified	HY scale: Not specified UPDRS: Not specified DiseaseDuration: Not specified Medication: OFF	32-channel EEG in main cortical regions at 512 Hz during a structured series of Timed Up and Go tasks. Average of 2 ear lobes electrodes was taken as reference. Data were segmented into 1 s windows and filtered at 0.5-50 Hz. 343 s of EW and 343 s of FOG samples were collected.	Division in bands was implemented. Z-transformation was applied to normalize EEG data. Power spectral density, centroid frequency and power spectral entropy were extracted.	Feed-forward neural network	6 hidden nodes. Levenberg Marquardt's algorithm with early stopping was used.	50 times runs. Data was divided in training 34 %, validation 33 %, and testing 33 %.	Results with only 2 channels C4-O2: Sensitivity: 72.54 Accuracy: 69.71
[262]	Classification of PD patients vs. controls.	Subjects: 100 PD and 100 controls. Age: PD: (50-70) HC: (50-70)	HY scale: 1-1.5 UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	2-channel EEG recorded for 30 min for the flexion and extension of the wrist. 5-50 Hz band-pass filter was applied.	EEG: Lyapunov and inverse Lyapunov exponent, Shannon Entropy EMG: power, standard deviation, root mean square, variance, waveform length, modified median, and mean frequency.	MLP	3 algorithms were tested. 1. Gradient Descent algorithms (traingd, traingdm), 2. Conjugate Gradient algorithms (trainscg, traingcg), and 3. Quasi-Newton algorithms (trainbfg, trainlm). Sigmoid function was used in the hidden layer. The number of hidden neurons was checked for 5, 7, 9, 10, 20, 30.	The dataset was divided into training 70 %, validation 15 %, and testing 15 %.	ANN with trainlm and 10 neurons: Accuracy = 100 RMSE = $4.03 \cdot 10^{-3}$ R value = 0.9998

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[293]	Classification of PD patients vs. controls.	Subjects: 40 PD and 30 controls Age: PD: 63.53 ± 4.95 HC: 64.72 ± 5.74	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	64-channel EEG recorded during 40 vocalizations of 5-6 s of the vowel /u/ with 5 pitch shifts each. The impedances were kept below 50 kΩ. A band-pass filter at 1-20 Hz was applied. The signals were segmented in epochs of 700 ms that contained the pitch shift. Referenced to the average of the mastoid electrodes. Trials with artifacts were rejected.	The epochs were the input data.	CNN, RNN, 2D-CNN-RNN, 3D-CNN-RNN	CNN: 8 layers. 2 -Conv, max-pooling, Conv, max-pooling, Conv, 2 fully connected layers with 1000 and 500 units. RNN: GRU layer with 6 units and 2 fully connected layers with 1280 and 300 units. 2D-CNN-RNN: Conv, GRU with 6 units and 2 fully connected layers with 2000 and 300 units. 3D-CNN-RNN: 2 Conv, max-pooling, GRU with 6 units and 2 fully connected layers with 2000 and 300 units. All models shared Adam optimizer with learning rate 0.001, ReLU activation, and softmax for the output layer. Filters, strides, and depths were specified.	5-fold cross-validation. The trials from the same patient were involved in either training or test.	The best model was 3D-CNN-RNN: Accuracy = 82.89 ± 9.60
[285]	Classification of PD patients vs. controls.	Subjects: 27 PD and 30 controls Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	256-channel EEG recorded during 2 types of visual stimuli (Stim and No-Stim) of 2.4 s. 185 electrodes were selected removing the channels located on the face and neck. Cleaning was not specified.	For each trial, the FFT of the last 2.24 s was computed. The features consisted of the spectral amplitudes. 6 channels over the occipital area were selected providing a total of 162 patient data and 180 control data.	LR, DT, RF	DT: Gini index measure of entropy was used, maximum depth of 5, minimum of 2 samples for split, and a minimum number of 1 sample per leaf as stopping criterion. RF: 100 trees of a max depth of 30.	For each stimulus 1000 runs of random train-validation splits in 70-30%. The models were tested using the other stimulus.	For both methodologies, the best model was RF. Stim for train: AUC_val = 0.994 AUC_test = 0.71 No-Stim for train: AUC_val = 0.998 AUC_test = 0.66

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[260]	Classification of PD patients vs. controls.	Subjects: 21 PD and 25 controls. Age: PD: 62.7 ± 7.32 HC: 54.6 ± 10.5	HY scale: 2.07 ± 0.39 UPDRS: PD: 31.00 ± 10.37 HC: 0.83 ± 1.27 Disease Duration: Not specified Medication: OFF	20-channel EEG recorded during 5 min in resting state with eyes-closed. 2 recordings were performed per patient. Cleaning was not specified.	Coherence analysis with 2 s windows with 50 % overlap was carried out. Pearson's correlation was calculated to assess the relationships between coherence and disease severity. The relative and absolute PSD were calculated at 1-40 Hz. Only 14 EEG-based features were used.	DFA	A linear DFA was used. The classifier input was selected by utilizing the step-wise discriminant analysis procedure in the SPSS software package.	Cross-validation.	Accuracy = 95.24 Sensitivity = 94.74 Specificity = 95.65 PPV = 94.74 NPV = 95.65 An excessive coherence was observed in the beta and gamma bands for PD.

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[325]	Classification of PD patients with MCI vs. NC	<p>Subjets: 36 PD with MCI and 35 PD with NC</p> <p>Age: PD-MCI: 61.1 ± 8.2 PD-NC: 57.0 ± 11.9</p>	<p>HY scale: PD-MCI: 2.1 ± 0.7 PD-NC: 1.9 ± 0.7</p> <p>UPDRS: PD-MCI: 21.9 ± 8.7 PD-NC: 21.7 ± 9.4</p> <p>Disease Duration: PD-MCI: 2.8 ± 2.3 PD-NC: 3.6 ± 3.6</p> <p>Medication: ON</p>	<p>16-channel EEG recorded during 30 min at 250 Hz. The impedances were set to <math>Z &gt; 100 \text{ M}\Omega</math>.</p> <p>High-pass filter at 0.16 Hz and low-pass filter at 500 Hz were applied before pre-amplification.</p> <p>Wavelet decomposition and reconstruction were made. Within 60 s segments, epochs with eyes-open without any obvious artefacts were selected.</p> <p>Artefacts were further eliminated by ICA.</p>	<p>64 features were calculated from power spectrum for 4 bands. Other features from MR were calculated.</p> <p>The feature importance method of Mean Impact Value was used to categorize the contribution of all features.</p>	SVM	RBF kernel was applied. The regularization parameters were identified using a 'grid search'.	<p>5-fold cross-validation on 80 % of patients. 20 % for test.</p> <p>Additional LOO cross-validation.</p>	<p>Only features from EEG Train:</p> <p>Accuracy: 64 Sensitivity: 68 Specificity: 62 PPV: 54 NPV: 75 AUC: 66 LOO-CV: 66</p> <p>In test:</p> <p>Accuracy: 67 Sensitivity: 67 Specificity: 67 PPV: 75 NPV: 57 AUC: 71</p>
[295]	Classification of HC vs. patients with different types of psychological disorders.	<p>Subjets: 25 PD and 25 controls</p> <p>Age: PD: 69.68 ± 8.73 HC: 69.32 ± 9.58</p>	<p>HY scale: Not specified</p> <p>UPDRS: Not specified</p> <p>Disease Duration: Not specified</p> <p>Medication: Not specified</p>	<p>60-channel EEG in Oddball task at 500 Hz. A band-pass filter at 0.1-60 Hz was applied. Low variance electrodes are dropped, so 30 electrodes were selected.</p>	<p>Pre-processed data were segmented into 4 s epochs in case of non ERP data and original trial length is kept in case of ERP data.</p> <p>26 linear and non linear features including time and frequency domain features were calculated using linear SVM classifier.</p>	SVM, LR, KNN, DT	SVM: RBF kernel. Grid search technique is used to tune the hyper-parameters	5-fold cross validation	<p>On PD patients vs. HC the best results for the selection of electrodes:</p> <p>Accuracy: 82 F1-score: 80 Precision: 85 Recall: 82 Model not specified.</p>

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[255]	Selection of the best QEEG features to identify different levels of cognitive impairment in PD.	Subjects: 118 PD classified into 5 groups according to the severity of the disease. G1: n=28, G2: n=33, G3: n=43, G4: n=5, G5: n=9. Age: G1= 60.54 ± 8.75, G2= 66.09 ± 6.65, G3= 67.04 ± 7.94, G4= 73.19 ± 5.29, G5= 67.56 ± 5.51	HY scale: G1= 1.93 ± 0.4, G2= 2.14 ± 0.55, G3= 2.21 ± 0.59, G4= 2.40 ± 0.55, G5= 2.00 ± 0.97 UPDRS: G1= 26.00 ± 11.73, G2= 29.55 ± 12.22, G3= 28.74 ± 11.44, G4= 31.00 ± 12.79, G5= 29.00 ± 18.87 Disease Duration: G1= 7.75 ± 5.29, G2= 8.36 ± 7.49, G3= 8.81 ± 5.02, G4= 6.60 ± 3.58, G5= 12.00 ± 6.56 Medication: ON	122-channel EEG recorded during 10 min in resting state. Average reference and 0.1-100 Hz bandwidth filter. Ocular artifacts were corrected and a 50 Hz filter was applied. Periods of drowsiness were removed, and the semi-automatic rejection of artifacts was performed to eliminate muscle activity. Each channel was divided into 4 s epochs. At least 20 segments were used for the analysis.	The relative and absolute spectral power were obtained for each epoch using a FFT and a 50 % overlap for the delta, theta, alpha, and beta bands. A division into 5 ROI was performed. For each case, high and low electrode density were considered. A statistical dependency study with an analysis of variance and the selection of characteristics with Pearson's correlation method was carried out.	SVM, KNN	SVM: Gaussian kernel KNN: k=9 and the Euclidean distance as a metric.	Data were split into a training set (n=100) and a test set (n=18). The training set was used for 5-fold cross-validation.	SVM: Accuracy = 87 ± 3.5 KNN: Accuracy = 88 ± 2.8 Both were achieved for the relative power with low-electrode density. Groups with few patients had worse results.

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[326]	Identify patients with early PD.	<p>Subjets: 29 drug-off early PD, 12 drug-on early PD and 22 controls</p> <p>Age: PD-OFF: 62.4 ± 6.3 PD-ON: 65.3 ± 5.4 HC: 63.8 ± 5.5</p>	<p>HY scale: 1 UPDRS: PD-OFF: 15.8 ± 7.5 PD-ON: 14.3 ± 6.2</p> <p>Disease Duration: Not specified Medication: ON and OFF</p>	<p>19-channel EEG recorded in resting state with eyes-closed at 500 Hz with additional channels for ECG, EMG, and EOG. Impedances were kept below 5 kΩ. Fast-ICA was applied to remove artifacts. Epochs with amplitude &gt;80 μV were rejected. More than 5 min signals were kept for each subject. A band-pass FIR filter at 0.5-45 Hz was used. Signals were segmented into 2 s non-overlapping epochs.</p>	<p>EEG signals were decomposed into two bands through the FIR filter. The P-Welch function was used to calculate the PSD of each channel within each epoch at 0.5-45 Hz with step size of 0.5 Hz. Models were fed with channel-frequency PSD and structured PSD. A personalized characteristic index of frequency domain was calculated for statistical anlysis.</p>	<p>CNN, SVM, MLP</p>	<p>SVM: linear kernel CNN: 3 2D-convolutional layers (with 3x3 filters and 8, 16 and 32 neurons respectively), with layer normalization after each of them, and 2 full-connect layers (with 4576 and 40 neurons) before the "Softmax" layer. Nadam optimizer with learning rate of 0.001 was used.</p>	<p>The dataset was obtained shuffling drug-off early PD group and HC group. It was divided into training (80%) and test (20%). The training set was used for 8-fold cross-validation.</p>	<p>The CNN model on the test set of structured PSD yielded: Accuracy: 99.87 ± 0.03 AUC: 99.88 ± 0.05.</p>



Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[302]	Classification of HC vs. PD patients.	Subjects: 15 PD and 18 HC Age: PD: 67.3 ± 6.5 HC: 67.6 ± 8.9	HY scale: 1.3 UPDRS: 23.3 ± 9.1 Disease Duration: 7.4 ± 4.3 Medication OFF	27-channel EEG recorded during 60 s in resting-state with eyes-open at 1000 Hz. The noise was removed.	Signals were filtered into four frequency bands, by a two-way FIR filter. A general orthogonalized directed coherence was used in each band to compute directional connectivity maps, which were normalized and converted into 2D images, resized to fed the VGG-16 model. LASSO regression models were computed for 30 runs separately on latent and non-latent cases. In each run, data were randomly divided in train (25) and test (8) sets. 30 results with least MSE at each run were considered as the LASSO coefficients.	VGG-16	VGG-16 with modifications: functional VGG-16, maxpooling layer (512 units), fully connected layer (512 units), fully connected layer (64 units) with ReLU activation function, and fully connected layer (2 units) with Softmax activation function. Optimizer: SGD. Learning rate: 0,01. Decay: 0,001. Batch size: 8. Loss function: binary cross entropy.	The model was tested 10 times on randomly chosen train/test partitions. The test data set was 25 % of data that had been randomly selected.	After 10 random repetitions with deep transfer learning Accuracy: 99.62 Precision: 100 Recall: 99.17 F1-score: 0.996 AUC: 0.996 Latent features were correlated with five clinical indices.

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[327]	Classification of HC vs. PD patients.	Subjects: 9 PD and 9 HC Age: PD: 55.22 ± 6.25 HC: 52.11 ± 4.98	HY scale: 2.28 ± 0.71 UPDRS: 25.89 ± 7.32 Disease Duration: Not specified Medication: Not specified	32-channels EEG recorded in resting state with eyes-open during 15 min at 1000 Hz. Impedances were kept below 5 kΩ. An online bandpass filter at 0.1-100 Hz and an offline bandpass filter at 0.1- 45 Hz were used. Artifacts were removed with ICA. Segments exceeding 150 μV were removed. ECG, PPG and RA signals were synchronously recorded. 14 epochs of 60 s without artifacts were selected per subject.	Absolute and relative powers, were computed for each electrode for 4 frequency bands, as well as sensory-motor rhythm, and the ratio of alpha to theta spectrums. For feature selection, the elastic network was employed with optimal norm regularization parameters L1 and L2.	SVM	A linear kernel was used.	9-fold cross validation. Experiments were randomly repeated 10 times. In each one, the model parameters and feature selection were determined by inner iterations	For only EEG: Accuracy: 87.54 ± 13.46 Sensitivity: 86.19 ± 15.14 Specificity: 88.89 ± 19.75 The EEG of PD patients had a significant decrease in high-frequency power.
[286]	Classification of 3 diseases between them and vs. controls.	Subjects: 16 PD and mached controls Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: ON	8-channel EEG during 10 s in a CNV trial. Some artifacts were removed and a low-pass filter at 30 Hz was applied. Mean level and baseline corrections were performed. For each trial, 2 epochs of 512 ms were extracted.	16 amplitude measures and a time measure were generated. The time measure included the post-imperative negative variation.	MLP	Input layer with 17 nodes, 1 hidden layer with 40 nodes, and 1 node for the output layer. The parameters were: gain = 1, momentum = 0.6 and learning rate = 0.9. Back-propagation was used.	LOO cross-validation.	For PD vs. controls: Sensitivity = 100 Specificity = 94 False-ve = 0 False+ve = 6 PPV = 94 NPV = 100 FAR = 6 FRR = 0

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[328]	Classification of HC vs. PD patients with and without medication.	Subjects: 15 PD and 16 HC Age: PD: 63.2 ± 8.2 HC: 63.5 ± 9.6	HY scale: 2-3 UPDRS Not specified Disease Duration: 4.5 ± 3.5 Medication: ON and OFF	32-channel EEG during 3 min in resting state with eyes-open at 512 Hz. The mean of the data was removed and were re-referenced to the common average. A highpass filtering at 0.5 Hz was used. The artifacts were manually examined and removed. Data were segmented into 2 s epochs.	An automated tunable Q wavelet transform was used to extract representative subbands. Five features were extracted from the subbands. The clinical significance of features are tested using the Kruskal Wallis test.	SVM, ANN, KNN, RF, LSSVM	LSSVM: polynomial (d = 10), RBF (sigma = 0.05), Morlet (ai = 0.01), Sinc and Mexican Hat (ai = 22 and constant omega = 0.5) were tested. KNN: k=10. RF: total number of learner was 10. SVM: Fine Gaussian kernel with automatic boxconstraint level. ANN: 10 hidden neurons.	10-fold cross-validation.	The results with LSSVM were: HC vs PD OFF Accuracy: 96.13 AUC: 97 HC vs PD ON Accuracy: 97.65 AUC: 98.56
[287]	Identification of the cognitive decline in PD patients.	Subjects: 20 PD patients with good cognition and 20 PD patients with poor cognition Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	21-channel EEG recorded during 15-20 min in resting state with eyes-closed. The signals were band-pass filtered at 0.16-70 Hz. For each patient, 5 visually inspected artifact-free epochs of 8.192 s were extracted.	794-time features were extracted from each channel. In addition, several clinical and spectral features were considered. The boruta algorithm was used for feature selection.	RF	The Bayesian algorithm was chosen for hyperparameter optimization. The depth of each tree and the number of trees were 1 to 100. The minimum number of samples in the leaf node and to split a node were 1 to 10 and 2 to 20, respectively.	5 runs of 10-fold cross-validation	The modeling approach (4) had the best performance. The boruta algorithm selected 1 to 3 features. None of them were clinical. Accuracy = 84.0 ± 4.2 F1 = 85.6 ± 3.2 Precision = 88.3 ± 6.9 Sensitivity = 83.0 ± 2.7 AUC = 86.8 ± 6.0

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[296]	Classification of PD patients vs HC.	Subjects: 25 PD and 25 HC Age: PD: $58.7 \pm 7.7$ HC: $54.6 \pm 7.7$	HY scale: 1-2 UPDRS: Not specified Disease Duration: $5.6 \pm 3.5$ Medication: Not specified	17-channel EEG recorded during 5 min in resting state with eyes-closed at 200 Hz. Impedances were kept below $5 k\Omega$ . Signals were amplified, a notch filter at 50Hz, and a bandpass filter at 0.5-30 Hz were applied. ICA was used. The data were re-referenced to the average reference. Signals were segmented into 5 s windows with 2.5 s overlap.	EEG segments were decomposed, using dynamic mode decomposition, into stable and unstable components at four frequency band. By Pearson correlation, stable brain network, unstable brain network, and inter-connected brain network were constructed and thresholded separately. Traditional brain network was also constructed. Topological attributes were extracted.	SVM, BN, RF, SGD, KNN, Ada-boost, RT, bag-ging, SL, vote methods	Not specified	10-fold cross-validation	The average results for all classifiers: Using the stable brain network attributes Precision: 89.6 Recall: 89.5 AUC: 90.8 Using traditional brain network attributes Precision: 85.4 Recall: 85.0 AUC: 85.4

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[281]	Detection of Turning Freezing in PD patients.	Subjects: 6 PD Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: OFF	15-channel EEG recorded during Timed Up and Go tasks. The EEG was segmented to 1 s epochs associated with normal turning or turning freezing, which resulted in 204 s of each of them. A band-pass filter at 0.5-40 Hz was used. Artifacts were removed. Z-transformation was applied to normalize the signals.	2 parameters were extracted from the theta, alpha, low beta, and high beta spectral bands using S-transform. These were the maximum amplitude for each band and the sum of amplitude of each band. ICA-EBM was considered for source separation.	BNN	3-layer (input, hidden, and output layers) feed-forward structure. Either 4 (for 15 channels) or 7 (for 4 channels) hidden nodes were considered.	Data were randomly split into 50 % for training and 50 % for test sets.	Best results with ICA, 4-channels and 7 hidden nodes. Train: Accuracy = 86.8 Sensitivity = 85.8 Specificity = 88.0 Test: Accuracy = 86.2 Sensitivity = 84.2 Specificity = 88.0 AUC = 0.9296

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[282]	Detecting the occurrence of GIF	Subjects: 4 PD Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: OFF	32-channel EEG during a structured series of Timed Up and Go tasks at 512 Hz. Only data from 9 electrodes positioned in locations of interest were processed. Average of 2 ear lobes electrodes was taken as reference. Data were segmented into 1 s windows and filtered using a nonlinear IIR band-pass filter with a cut-off frequency lower than 1 Hz and higher than 50Hz to remove artifacts. 122 EEG samples were collected, associated to Good Start (61) and Gait Initiation Failure (61).	Welch's method with a 256 points FFT with 55 % overlapping was used to analyze four frequency sub-bands. Power spectra density and centroid frequency of each band were calculated.	BP-NN	Two-layer feed-forward neural network with 10 hidden nodes. Activation function of the hidden layer was tan sig. Levenberg Marquardt's algorithm with early stopping was used.	50 times runs. Data was randomly divided in training 34 %, validation 33 %, and testing 33 %.	The best performance of the classification system was achieved with a combination of nine channels: Sensitivity: 84.27 Specificity: 85.02 Accuracy: 84.80

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[329]	Classification of PD patients with medication, without medication, and controls.	Subjects: 10 PD and 12 controls Age: PD: (40-80) HC: (40-80)	HY scale: 1-2 UPDRS: Not specified Disease Duration: Not specified Medication: ON and OFF	61-channel EEG recorded during 2 min in resting state followed by 140 s of intermittent photic stimulation with eyes-closed. 10 electrodes were selected. Band-pass filter at 0.5-50 Hz and a 60 Hz notch filter were applied. Average reference was used. The artifacts were removed. The impedance remained below 5 k $\Omega$ .	The last 10 s stretch of stimulation was divided into 20 segments of 0.5 s which were used to calculate the partial directed coherence for 6 bands (delta, theta, alpha, beta, gamma1, and gamma2). A total of 60 features were calculated. 19 features were selected using GA with a population of 20 individuals, 20 generations, a crossover probability of 0.6, and a mutation probability of 0.03. Both approaches were considered as inputs.	BN, NB, MLP, SVM, J48, RT, RF, ELM, mELM	MLP: hidden layers 1 and 2, learning rate 0.3, momentum 0.2, iterations 500. SVM: polynomial kernel (exponent = 1 to 5) RBF kernel (gamma = 0.25 and 0.5). RF: Trees 10 and 50. ELM: 100 neurons in the hidden layer and sigmoid kernel. mELM: 100 neurons in the hidden layer, dilatation, and erosion kernels.	K-fold cross-validation.	The best model was RF with 50 trees. All features: Accuracy = 99.22 19 features: Accuracy = 98.09

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[298]	Classification of PD patients vs. controls using 6 emotional stimuli.	Subjects: 20 PD and 20 controls Age: PD: (40-65) HC: (40-65)	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	14-channel EEG recorded during emotional stimuli. The signals were segmented into 10 s epochs with overlapping of 75%. Then, pass-band elliptic filters were used to obtain the alpha, beta, and gamma bands.	For each band 6 features were calculated: Entropy (EN), Energy-Entropy (EEN), Teager Energy-Entropy (TEEN), Spectral Entropy (SEN), Spectral Energy-Entropy (SEEN), and Spectral Teager Energy-Entropy (STEEN).	PNN, KNN, SVM	KNN: k in the range 1 to 10. PNN: used exponential activation function with $\sigma$ ranged from 0.55 to 0.65.	Not specified.	The best accuracy for each emotion: Happiness: 99.59 with SVM and TEEN Sadness: 90.81 with KNN and EN Fear: 95.07 with KNN and SEEN Anger: 91.42 with SVM and SEEN Surprise: 94.53 with KNN and SEEN Disgust: 88.18 with SVM and SEEN.
[299]	Classification of PD patients vs. controls using 6 emotional stimuli.	Subjects: 20 PD and 20 controls Age: PD: (40-65) HC: (40-65)	HY scale: Not specified UPDRS: Not specified Disease duration: Not specified Medication: Not specified	14-channel EEG recorded during emotional stimuli. The signals were segmented into 10 s epochs with overlapping of 75%. Then, pass-band elliptic filters were used to obtain the alpha, beta, and gamma bands.	For each band 4 features were calculated: Entropy (EN), Energy-Entropy (EEN), Spectral Entropy (SEN), and Spectral Energy-Entropy (SEEN).	PNN, KNN	KNN: k in the range 1 to 10. PNN: exponential activation function with $\sigma$ ranged from 0.55 to 0.65.	Not specified.	The best accuracy for each emotion: Happiness: 96.8 with PNN and SEEN Sadness: 90.81 with KNN and EN Fear: 95.07 with KNN and SEEN Anger: 88.65 with KNN and EN Surprise: 94.53 with KNN and SEEN Disgust: 87.43 with KNN and EN.



Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[288]	Classification of 3 diseases and controls through 4 methodologies.	Subjects: 15 PD and 16 controls Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	12-channel EEG recorded during approximately 250 ms. Cleaning and task were not specified.	Features selected through 4 methodologies: 1. The maximum amplitude and frequency of the FFT in alpha, beta, theta, and gamma bands. 2. Amplitudes of the first 8 peaks of the FFT and their frequency values up to 30 Hz. 3. The FFT of the EEG signal below 30 Hz. 4. The maximum amplitude and frequency of the FFT in each band were applied to the combination of the 12 electrodes.	ANN	Feed-forward architecture with 16 nodes in the hidden layer. Activation functions were logsig and tansig. The training algorithm was Levenberg-Marquardt.	75 % training, 12.5 % validation and 12.5 % test. The sets were balanced in terms of pathologies that contained each one.	Bad results were expressed in terms of the correlation coefficient between target and predicted values.

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[258]	Classification of RBD patients vs. controls. Some patients were eventually diagnosed with PD and dementia.	Subjects: 118 RBD and 74 controls. 14 RBD became PD. No direct patient data. Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease: Not specified Duration: Not specified Medication: Not specified	14-channel EEG recorded in resting state with open-eyes periods followed by closed-eyes periods. Only eyes-closed sequences were considered. The EEG was recorded when the patients were RBD. A band-pass filter at 0.3-100 Hz and a notch filter at 60 Hz were applied. Artifacts were removed.	Using a sliding window of 1 s, 148 spectrograms per subject were generated (of 20 seconds of duration each) using only the FFT amplitude bins in the band 4-44 Hz. The spectrograms were centered and normalized to unit variance for each frequency and channel.	CNN, RNN	CNN: 4 hidden-layer convolutional. Dropout, max-pooling layers, and cross-entropy loss function were used. RNN: with LSTM and GRU, with 3 cells which 32 units each. Dropout was used.	LOO cross-validation. For training, the dataset was balanced by random replication preserving the distribution of the subjects.	The results for controls vs. PD: CNN: Accuracy = $79 \pm 1$ AUC = $0.87 \pm 0.1$ RNN: Accuracy = $81 \pm 1$ AUC = $0.87 \pm 0.1$ In RNN, there was no difference between LSTM and GRU.
[289]	Classification of PD patients with RBD vs. controls.	Subjects: 14 RBD with PD and 14 controls. Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease: Not specified Duration: Not specified Medication: Not specified	14-channel EEG recorded in resting state with open-eyes periods followed by closed-eyes periods. Only eyes-closed sequences were considered. The EEG was recorded when the patients were RBD. A band-pass filter at 0.3-100 Hz and a notch filter at 60 Hz were applied. Artifacts were removed.	Several spectrograms were computed to extract temporal series of power for each electrode and band (10 bands in total). The use of 4 s and 1 s spectrogram windowing was explored.	RNN	ESN layer with 3000 nodes, least-squares regularization. Spectral radius ranged from 0.5 to 2.	For each parameter set 50 runs were carried out with random and balanced training (90%) and test (10%) sets.	The best performance was obtained with 1 s. Test set: Average_accuracy = 85

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[261]	Classification of PD patients vs. controls.	Subjects: 30 PD and 30 controls. Age: PD: (50-70) HC: (50-70)	HY scale: 1-1.5 UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	2-channel EEG recorded for 30 min for the flexion and extension of the wrist. Cleaning was not specified.	EEG: Shannon entropy, Lyapunov, and inverse Lyapunov exponent were calculated. EMG: power, standard deviation, root mean square, variance, waveform length, modified median, and mean frequency.	MLP	Back Propagation was used as the learning algorithm and "trainlm" was used as the training function. Sigmoid transfer function was used for the hidden layer.	The dataset was divided into training 70 %, validation 15 %, and testing 15%.	MLP with inputs: EEG: accuracy = 62 EMG: accuracy = 73 EEG+EMG: accuracy = 98.8
[294]	Classification PD and HC	Subjects: 16 PD and 15 controls Age: PD: 62.6 ± 8.3 HC: 63.5 ± 9.6	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: OFF	40-channel EEG during 2 min in resting state at 512 Hz. Signals were segmented into patches of 512 time samples.	EEG patches of 512 samples were used as input. Channels were considered independently.	ANN	12 layers with 512, 512, 128, 128, 64, 64, 32, 32, 16, 16, 8, and 2 (output) units. The conjugate gradient backpropagation with Polak-Ribière updates was used for updating the weights, and biases of the network. Grid search was used for hyper-parameters selection.	20 % for testing and the rest was used for training and validation	The Oz/P8/FC2 channels were selected and used with majority voting. Results on test set were: Accuracy: 98 Sensitivity: 97 Specificity: 100

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[330]	Classification of PD patients with ON medication vs. OFF medication.	Subjects: 28 PD Age: $69.75 \pm 8.43$	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: ON and OFF	64-channel EEG recorded during 12 min in the oddball task. 60 electrodes were selected during the period without stimuli. A band-pass filter at 13-30 Hz was applied. Non-brain activities were subtracted.	The average spectral power over all the frequency ranges was calculated and added together. The differences between OFF and ON medication data were compared. 12 channels were selected. A preprocess with delay ( $\tau=12$ ) and embedding dimension ( $m=3$ ) were applied. 3 sets of 12 time-series (non-delayed, $\tau$ -delayed and $2\tau$ -delayed) were generated.	EEGNet, HNet, DGH-Net	ADAM optimizer with learning rate = $10^{-4}$ for 500 epochs was used. HNet: 3 layers (Convolutional, LSTM, LSTM) with 28 filters, 14 and 2 neurons, with sigmoid, sigmoid, and softmax activation. DGHNet: 2Dconvolutional as input layer with 4 filters and exponential linear activation. Values of $m=4$ and $\tau=3, 6, 9, 12,$ and $17$ were trained. 1Dconvolutional with exponential linear activation. LSTM with 2 outputs and softmax activation.	Intra-patient: the data was randomly shuffled and split into 90% training and 10% test. The training dataset was divided into 85% training and 15% validation. Inter-patient: 2 patients with ON and OFF medication were randomly left out for the test. The remaining patients were partitioned at 85% training and 15% validation.	The results of HNet and DGHNet were similar (accuracy: 99.74 vs. 99.22). DGHNet was considered better for containing fewer parameters. Test: Accuracy = 99.22 Sensitivity = 98.98 Specificity = 99.46 MCC = 98.44 F1 = 99.24

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[290]	Classification of PD patients vs. HC	Subjects: 9 PD and 9 HC Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	20-channel EEG recorded in 3 epochs of 1 min during externally paced bilateral, cyclical foot in the sitting position at 1200 Hz.	Kolmogorov Complexity, Sample Entropy and Approximate Entropy were extracted for each channel. Different subsets of electrodes were considered, and their associated features were used as input.	MLP	Five hidden layers chosen by trial and error method.	10-fold cross validation	The best results were obtained for the channels Fz, F1 and F2: Accuracy: 97.5 Precision: 100 Sensitivity: 96.7 Specificity: 100 AUC: 0.978
[254]	Selection of the best classifier of PD vs. controls using the minimum number of HOS features.	Subjects: 20 PD and 20 controls. All right-handed. Age: PD: 59.05 ± 5.64 HC: 58.10 ± 2.95	HY scale: 1: n=2; 2: n=11; 3: n=7 UPDRS: Not specified Disease Duration: 5.75 ± 3.52 Medication: ON	14-channel EEG recorded during 5 min in resting state with eyes-closed. Threshold technique at 80μV. A band-pass filter at 1-49 Hz was applied. 2 s epochs with 50 % overlap were considered.	For each epoch, a total of 13 HOS characteristics were calculated. The Student's t-test was also obtained to determine the importance of the characteristics.	DT, KNN, FKNN, NB, PNN, SVM	FKNN: Euclidean distance, m=1.24 and k=3. KNN: k=2 and Euclidean distance. PNN: exponential activation and σ=0.284. SVM: polynomial (orders 2 and 3), RBF, and linear kernels.	10-fold cross-validation. The characteristics were added one by one to each classifier until maximum precision was achieved.	The best model was SVM with RBF kernel: Accuracy = 99.62 ± 0.58 Sensitivity = 100 ± 0.0 Specificity = 99.25 ± 0.53 Precision = 99.38 ± 0.47 F-score = 0.98 ± 0.05

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[283]	Identification of the onset of freezing of PD patients during walking	Subjects: 26 PD Age: 69.8 ± 8.41	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	4-channel EEG recorded in periods of 1-2 h per patient at 500 Hz during a structured series of video-recorded Timed Up and Go tasks. Only data from 10 patients and differential channels O1-T4 and P4-T3 were used. Epochs of 1 s from individual freezing events were taken. Band-pass at 0.5-60 Hz, and band-stop at 50 Hz butterworth IIR filters were applied. An additional threshold filter was applied.	Data were divided into normal, onset, and freezing with 40 samples per subject and group. Discrete wavelet transform was used to calculate wavelet entropy for 5 frequency bands, and total wavelet entropy. The model was fed with the wavelet entropy of 3 bands an total for channel O1-T4, and all bands and total for channel P4-T3, separately and in combination.	BP-NN	3 layers with 4-7 hidden nodes depending on the number of inputs dimension and the number of training pairs. Levenberg Marquardt algorithm was used. Activation function was Tangent Sigmoid.	20 runs for each feature. Data was divided in training 56 %, validation 25 %, and test 19%.	The results for Normal vs. Onset for P4-T3 were: Accuracy: 76.6 ± 3.4 Sensitivity: 74.2 ± 6.8 Specificity: 78.9 ± 7.3 Normal vs. Freezing for O1-T4 and P4-T3: Accuracy: 73.9 ± 2.8 Sensitivity: 71.2 ± 6.1 Specificity: 77.2 ± 4.7

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[284]	Predicting transition to FOG from normal walking	Subjects: 26 PD Age: 69.8 ± 8.41	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	4-channel EEG recorded in periods of 1-2 h per patient at 500 Hz during a structured series of video-recorded Timed Up and Go tasks. Data from 10 patients without significant artifacts were selected and differential channels O1-T4 and P4-T3 were used. Epochs of 1 s were taken. Band-pass at 0.5-60 Hz, and band-stop at 50 Hz butterworth IIR filters were applied. Ocular and muscular artifacts were removed.	Data were divided into normal, transition, and freezing with 40 samples per subject and group. Discrete wavelet transform was used to calculate total, global and centroid frequency wavelet cross spectrum for 5 bands, wavelet cross frequency energy ratios por specific pairs of bands. Other statistical features were computed. 131 features were selected to feed the model separately and in combination.	MLP, KNN	MLP: 3 layers with 8 to 12 hidden layer neurons. The Levenberg Marquardt algorithm was used. Error goal of 0.01. KNN: 15 to 40 nearest neighbors based on the Euclidian distance.	20 runs for each feature and their combination. Data was divided in training 56 %, validation 25 %, and test 19 %.	Using MLP with statistical features: Sensitivity: 75.47 Specificity: 71.47 Accuracy: 73.47 For kNN with all features were: Sensitivity: 87.25 Specificity: 70.00 Accuracy: 52.75

Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[291]	Classification of HC vs. PD patients through different emotional states.	Subjects: 5 PD and 5 controls Age: PD: 45-65 HC: 20-25 Age for each subject were specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	1-channel EEG recorded while watching happy, sad, and neutral videos, and during meditation in resting state with eyes-closed. The EEG collecting device produced power values for eight frequency bands, and attention values.	Benjamini-Hochberg corrected F-tests and t-tests were applied on attention level, average power of frequency bands and on absolute value of first and second order derivatives. An initial model was fed with first and second order derivatives for each pair emotion-band. Meditation was not used. The three best pairs were combined in the final model. Input features were preprocessed to have zero mean and unit variance.	MLP	Initial model: 1 hidden layer with 4 neurons. Final model: 2 hidden layers with 4 and 2 neurons, respectively. Adam optimizer, Log-loss loss function, and activation function tanh were used in both.	5-fold cross validation	The results of the final model were: Accuracy: 0.965 F1 score: 0.976 Recall: 0.970 Precision: 0.955



Objective		Evaluation of the quality of the dataset			Data pre-processing	Evaluation of the models used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[259]	Classification of patients with neurological diseases vs. controls.	Subjects: 31 PD and 264 controls. Age: PD: 56.62 ± 12.32 HC: 49.51 ± 12.54	HY scale: 1-3 UPDRS: 43.44 ± 15.53 Disease Duration: Not specified Medication: Not specified	19-channel EEG recorded during 5 min in resting state with eyes-closed. The impedances were kept below 5 kΩ. A high-pass filter at 0.15 Hz and a low-pass filter at 200 Hz were used. A band-pass filter at 2-44 Hz was applied. Artifacts were manually removed. Channels were divided into ROI.	The power spectrum was calculated for each subject and 5 frequency bands (delta, theta, alpha, beta, and gamma) were considered for each ROI.	SVM	The default settings were used as the running parameters.	10-fold cross-validation. The distribution of the patients was kept. Controls with obesity were used to validate the model.	The results for controls vs. PD: Accuracy = 94.34 ± 1.81 Sensitivity = 0.93 ± 0.02 FPR = 0.11 ± 0.01 ROC = 0.95 ± 0.02 MAE = 0.07 ± 0.02 RMSE = 0.16 ± 0.02

Tabla E.1: Summary of the objectives, participants, state of PD, EEG pre-processing, features, models used, model parameters, training and validation methods and best results for each article included in this review. Here n stands for the number of patients. Fuente: [236].



## Apéndice F

# Detalles del software de EEGraph

La librería EEGraph ha sido desarrollada con el objetivo principal de crear una herramienta, tanto para investigadores como para especialistas clínicos en el campo de la neurociencia, que proporcione información directa sobre la conectividad del cerebro a partir de señales de electroencefalografía. A continuación, se detalla el funcionamiento interno de esta herramienta.

### F.1. Paquetes requeridos para su instalación

Esta librería ha sido implementada en el lenguaje de programación Python y puede ser utilizada en varios sistemas operativos. La instalación se realiza mediante el comando “pip install EEGRAPH”. A continuación se detallan las librerías utilizadas junto con una breve especificación acerca de la función realizada:

- **Numpy.** Se utiliza para realizar la división en bandas de frecuencia a través de la transformada rápida de Fourier, para implementar el valor umbral, y para definir algunas operaciones dentro del cálculo de las medidas de conectividad (por ejemplo, PS).
- **Pandas.** Se utiliza tanto para estructurar los datos en dataframes como para leer el fichero con las especificaciones del montaje de los electrodos.
- **Mne.** Se utiliza para leer el archivo EEG en los diferentes formatos soportados.
- **NetworkX.** Se utiliza para la creación y manipulación de grafos.
- **Plotly.** Se utiliza en la implementación de la representación visual de los grafos.
- **Scipy.** Se utiliza para estimar la correlación de Pearson, la correlación cruzada, la coherencia cuadrática, la densidad espectral de potencia, la transformada de

Hilbert para el cálculo de la fase instantánea, la densidad espectral de potencia cruzada y la entropía de Shannon, que son cruciales en la definición de algunas de las medidas de conectividad consideradas (C, CC y CCC, SC, IC, PLV y PLI, WPLI y H, respectivamente).

- **SCoT**. Se utiliza para el cálculo del DTF.
- **AntroPy**. Se utiliza para la estimación del SE.

Las versiones necesarias de cada una de estas librerías se especifican en el archivo *requirements.txt* de GitHub.

## F.2. Datos de entrada

EEGraph permite una gran variedad de formatos de EEGs como entrada para cubrir tanto las señales de EEG en bruto como las procesadas. Los formatos admitidos con sus respectivas extensiones son: Brainvision (.vhdr), Neuroscan CNT (.cnt), formato de datos europeo (.edf), formato de datos Biosemi (.bdf), formato de datos general (.gdf), EGI binario simple (.egi), formato EGI MFF (.mff) y eXimia (.nxe).

Para cargar el archivo asociado a un EEG, en adelante, archivo de EEG, hay que crear un objeto tipo *Graph* y luego llamar al método “*load\_data*” especificando el parámetro “*path*”, que debe coincidir con la ubicación de dicho archivo. Si el archivo de EEG no contiene las etiquetas de los canales, EEGraph ofrece la opción de introducir un archivo adicional a través del parámetro “*electrode\_montage\_path*” que permite a EEGraph etiquetar e identificar cada señal con su correspondiente electrodo. Este archivo debe contener al menos dos columnas, una con la numeración de los electrodos en el archivo EEG y otra con las etiquetas de los canales correspondientes, separadas por los delimitadores “\s+”, “;”, o “:”. Esta opción está desactivada por defecto. Por último, el parámetro “*exclude*” permite al usuario excluir una lista de canales del archivo EEG para el análisis posterior.

## F.3. Construcción de nodos

Una vez que la librería lee el archivo que contiene el EEG, EEGraph identifica automáticamente el número de electrodos que contiene, etiqueta cada una de las señales con el nodo que le corresponde y coloca cada nodo en la posición adecuada con respecto a la posición del electrodo para representar el grafo. La construcción del grafo tiene en cuenta que los sistemas de montaje más pequeños están contenidos en estructuras

de montaje más grandes [331]. Por ello, se ha construido un diccionario con 333 posiciones de electrodos y se ha generado una malla donde se han colocado estos electrodos con sus correspondientes etiquetas para posteriormente identificar las posiciones de los nodos del grafo. Se ha considerado que hay electrodos que pueden ser identificados con más de una etiqueta, por lo que también se han incluido las diferentes opciones de etiquetado, estas son:  $T3 = T7$ ,  $T4 = T8$ ,  $T5 = P7$ ,  $T6 = P6$ . El resultado es un grafo cuyos nodos representan las posiciones de los electrodos del casco de EEG etiquetados con el nombre asociado a cada canal.

## F.4. Construcción de aristas

Para la selección de las medidas implementadas en la librería, se realizó un estudio bibliográfico para determinar cuáles eran las más utilizadas [226, 227, 228, 229, 230]. Como resultado, EEGraph implementa un total de 12 medidas de conectividad, que se detallan en el Capítulo 3. Del conjunto de estas medidas, *Power Spectrum*, *Spectral Entropy* y *Shannon Entropy*, extraen información de un sólo canal de EEG, mientras que las restantes medidas proporcionan información sobre la relación entre pares de electrodos. Las medidas de conectividad que relacionan pares de electrodos proporcionan el valor de la arista que une esos nodos en el grafo. De este modo, se genera un grafo totalmente conectado en el que cada arista tiene un peso asociado que se corresponde con la medida de conectividad elegida. Para facilitar la extracción de patrones, hemos introducido un parámetro modificable por el usuario “threshold” para representar sólo aquellos valores que superen este límite. El valor por defecto se ha definido a partir de la bibliografía anterior. En el caso de las medidas que proporcionan información sobre electrodos individuales, se hace necesario definir un criterio que relacione los nodos del grafo para establecer las conexiones. El criterio elegido consiste en conectar totalmente un porcentaje fijo de los nodos con los valores más altos. El valor por defecto de dicho porcentaje es de 25 %, aunque el usuario puede modificarlo. En este caso, las aristas del grafo no tienen ninguna medida asociada, de modo que el valor que muestran las aristas es 1,00, indicando la existencia de conexiones. El valor de la medida se muestra en los nodos.

## F.5. Medidas de conectividad

Para elegir la medida de conectividad con la que se quiere realizar el estudio, utilizamos el método “*modelate*” con la posibilidad de especificar el valor de los siguientes parámetros:

- **“*window\_size*”**. Este parámetro admite un número entero, coincidiendo con la longitud temporal (en segundos) de los segmentos en los que se divide el EEG, o una lista de números enteros, que permiten al usuario especificar los límites (en segundos) de dichos intervalos.
- **“*connectivity*”**. Este parámetro admite como entrada un *string* (cadena de caracteres en forma de texto) que especifica el nombre de la medida de conectividad que el usuario desea utilizar. Los nombres específicos de estas medidas pueden verse en detalle en la sección de “*Modelate Data*” dentro del apartado wiki de GitHub.
- **“*bands*”**. Este parámetro admite, como entrada, una lista de *strings*, en la que el usuario puede especificar la banda o bandas para las que quiere obtener la información. Este parámetro debe ser especificado para aquellas medidas que trabajen sobre el espectro de frecuencias del EEG.
- **“*threshold*”**. Este parámetro opcional admite como entrada, un *float* (esto es un número en representación de coma o punto flotante), y su valor por defecto se puede encontrar en la sección “*Modelate Data*” de GitHub. En el caso de las medidas que relacionan pares de electrodos, este parámetro especifica un límite a partir del cual se representan las conexiones. En el caso de las medidas que proporcionan información sobre electrodos individuales, este parámetro representa el porcentaje de nodos con los valores más altos que están totalmente conectados. Este porcentaje debe introducirse dividido por 100, es decir, como partes por unidad.

Dado que los estudios de electroencefalografía pueden dividirse principalmente en dos grupos, el análisis del dominio temporal y el análisis del dominio de la frecuencia, las medidas implementadas cubren estos dos aspectos. Estas son:

- **Medidas definidas en el dominio temporal:** *Pearson Correlation*, *Cross-Correlation*, *Corrected Cross-Correlation*, *Phase Lag Index* y *Shannon Entropy*.
- **Medidas definidas en el dominio espectral:** *Coherence*, *Imaginary Coherence*, *Phase Locking Value*, *Phase Lag Index*, *Weighted Phase Lag Index*, *Directed Transfer Function*, *Power Spectrum* y *Spectral Entropy*.

## F.6. Datos de salida

La salida que proporciona EEGraph es doble. Por un lado, proporciona la representación visual del grafo, es decir, la imagen del grafo con los nodos y las conexiones entre ellos que depende de la medida de conectividad elegida. Dentro de esta representación,

existen dos posibilidades. En caso de que la medida de conectividad proporcione una relación entre dos nodos, el valor de la medida puede visualizarse aproximando el cursor al punto medio de cada arista. Además, la anchura de cada línea depende del valor asociado. En el caso de que la medida de conectividad proporcione información sobre nodos individuales, los valores obtenidos se representan en los nodos, y las aristas representadas muestran un valor de 1,00 cuando se aproxima el cursor a su punto medio, indicando que existe una conexión. Esta representación está pensada para facilitar la obtención de información de una forma directa y visual. Por otro lado, EEGGraph proporciona una representación matricial del grafo, que contiene todos los valores de la medida de conectividad elegida por el usuario. Si el EEG se ha dividido en varias ventanas, EEGGraph proporciona tanto la representación visual como matricial del grafo asociado a cada uno de los segmentos, lo que permite al usuario evaluar la evolución temporal de las conexiones cerebrales.

Para obtener la representación visual del grafo el usuario puede utilizar una de las salidas proporcionadas por el método *“modelate”*. La salida *“graphs”* es un diccionario que contiene todos los grafos generados a partir del archivo EEG, como objetos *NetworkX graph*. Cada uno de estos objetos contiene información sobre las conexiones del grafo y los pesos de las conexiones que superan el valor umbral del parámetro *“threshold”*, que son las que finalmente se representarán. Para obtener la imagen del grafo, es necesario llamar al método *“visualize”*. Este método requiere de dos parámetros: *“graph”*, que debe ser un objeto tipo *NetworkX graph* como los que componen el diccionario *“graphs”* anterior, y *“name”*, un nombre para el grafo que el usuario quiera visualizar. Por tanto, el usuario debe seleccionar el grafo contenido en el diccionario *“graphs”* que quiere representar e introducirlo como valor del parámetro *“graph”*. El resultado de este método es un grafo como el que se muestra en la Fig. F.1, que se proporciona al usuario a través de un fichero HTML que se guarda automáticamente en la carpeta del usuario con el nombre especificado anteriormente, y se abre en el navegador web.

Como se puede ver en la Fig. F.1, hay dos opciones de visualización: *“Show edge markers”*, que muestra los marcadores donde el usuario debe situar el cursor para ver el valor del peso asociado a esa arista (o nodo en el caso de las medidas que proporcionan información sobre un solo canal), y *“Hide edge markers”*, que hace desaparecer los marcadores, aunque el usuario puede visualizar los valores de la medida de la misma manera. Hay que tener en cuenta que el grosor de cada arista depende del valor obtenido.

Finalmente, para obtener la representación matricial del grafo, el usuario puede utilizar la otra salida del método *“modelate”*. Precisamente, la salida *“connectivity\_matrix”* es un *numpy.ndarray* con la forma  $G \times N \times N$ , donde  $G$  es el número de grafos y  $N$  es el nú-

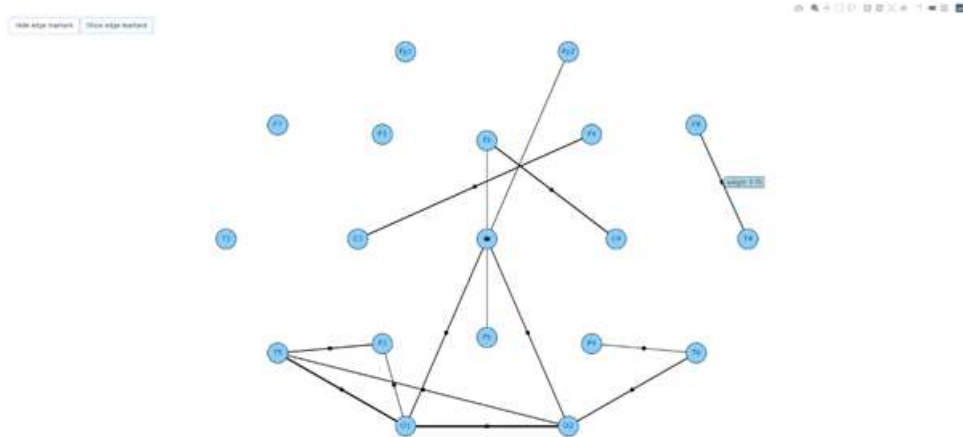


Figura F.1: Ejemplo de salida de un grafo por medio del método “visualize” con la opción de visualización. Fuente: elaboración propia.

mero de electrodos del EEG o nodos del grafo. Cabe destacar que “*conectividad\_matriz*” contiene todos los valores proporcionados por la medida de conectividad independientemente del valor del parámetro “*threshold*”. Una vez obtenida la representación matricial, a través de la variable “*connectivity\_matrix*” el usuario puede utilizarla dentro de un script de Python, o guardarla en un archivo “.CSV” utilizando la función “*savetxt*” de la librería *Numpy*.

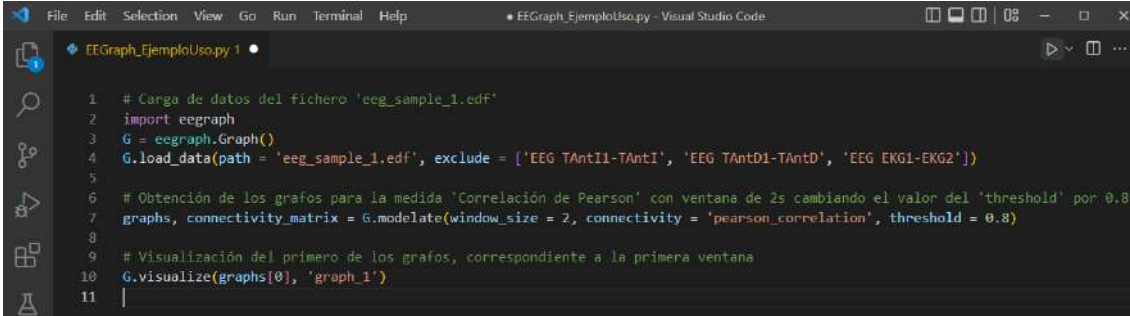
## F.7. Guía de usuario de EEGraph

Tras corroborar los paquetes junto con las versiones necesarias para ejecutar la librería, que podemos encontrar en el fichero *requirements.txt* de GitHub, procedemos a la instalación de EEGraph mediante el comando “`pip install eegraph`”.

Tras la instalación procedemos a la lectura del EEG. EEGraph admite varios formatos de entrada de EEGs, que incluye los formatos de salida de diferentes dispositivos de electroencefalografía y herramientas de preprocesamiento. Para cargar el EEG, utilizamos el método “*load\_data*” especificando el parámetro “*path*”. Si el archivo de EEG no contiene las etiquetas de los canales (como es el caso de nuestros EEGs), introducimos un archivo adicional a través del parámetro “*electrode\_montage\_path*” que contenga la identificación de cada canal con su etiqueta. Así mismo, se puede indicar si se desean excluir canales del análisis mediante el parámetro “*exclude*”. En la Fig. F.2 se recoge un ejemplo de código para realizar cada proceso con EEGraph.

Una vez que la librería lee el archivo que contiene el EEG, EEGraph identifica automáticamente el número de electrodos que contiene, etiqueta cada una de las seña-





```
1 # Carga de datos del fichero 'eeg_sample_1.edf'
2 import eeagraph
3 G = eeagraph.Graph()
4 G.load_data(path = 'eeg_sample_1.edf', exclude = ['EEG TAntI1-TAntI', 'EEG TAntD1-TAntD', 'EEG EKG1-EKG2'])
5
6 # Obtención de los grafos para la medida 'Correlación de Pearson' con ventana de 2s cambiando el valor del 'threshold' por 0.8
7 graphs, connectivity_matrix = G.modelate(window_size = 2, connectivity = 'pearson_correlation', threshold = 0.8)
8
9 # Visualización del primero de los grafos, correspondiente a la primera ventana
10 G.visualize(graphs[0], 'graph_1')
11 |
```

Figura F.2: Muestra el código empleado para cargar, modelar y visualizar un EEG mediante la librería EEGraph para la correlación de Pearson, con un threshold de 0,8 y con un tamaño de ventana de 2 segundos. Fuente: elaboración propia.

les con el nodo que le corresponde y coloca cada nodo en la posición adecuada con respecto al casco de electroencefalografía para representar el grafo. A continuación, seleccionamos una medida de conectividad de entre las 12 incluidas en la librería. Para especificar la medida de conectividad utilizamos el método “*modelate*” con la posibilidad de especificar los parámetros “*window\_size*”, “*connectivity*”, “*bands*” y “*threshold*”. En el caso del parámetro “*bands*”, la librería incluye las 5 bandas que han sido descritas y el usuario debe especificarlas para aquellas medidas que trabajen sobre el espectro de frecuencias del EEG, es decir, para las medidas: “*squared\_coherence*”, “*imag\_coherence*”, “*power\_spectrum*”, “*spectral\_entropy*”, “*wpli*”, “*plv*”, “*pli\_bands*”, y “*dtf*”.

La salida proporcionada por EEGraph se realiza mediante forma de imagen y de estructura de datos por medio de una matriz con el fin de facilitar su interpretación y el manejo de los datos respectivamente.



# Financiación

Esta tesis doctoral se ha realizado con datos obtenidos financiados por el Ministerio de Ciencia e Innovación (PID2020-113222RB-C21/AEI/10.13039/501100011033) y (PID2020-113222RB-C22/AEI/10.13039/501100011033).



## Bibliografía

- [1] GBD 2016 Parkinson's Disease Collaborators, "Global, regional, and national burden of Parkinson's disease, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016," *The Lancet Neurology*, vol. 17, pp. 939–953, Nov. 2018.
- [2] P. Mahlknecht, F. Krismer, W. Poewe, and K. Seppi, "Meta-analysis of dorsolateral nigral hyperintensity on magnetic resonance imaging as a marker for Parkinson's disease," *Movement Disorders*, vol. 32, pp. 619–623, Apr. 2017.
- [3] J. Opara, A. Małecki, E. Małecka, and T. Socha, "Motor assessment in Parkinson's disease," *Annals of Agricultural and Environmental Medicine*, vol. 24, pp. 411–415, Sept. 2017.
- [4] E. R. Dorsey, T. Sherer, M. S. Okun, and B. R. Bloem, "The Emerging Evidence of the Parkinson Pandemic," *Journal of Parkinson's Disease*, vol. 8, no. s1, pp. S3–S8, 2018.
- [5] D. W. Dickson, "Neuropathology of Parkinson disease," *Parkinsonism & Related Disorders*, vol. 46, pp. S30–S33, Jan. 2018.
- [6] R. Bahena-Trujillo, G. Flores, and J. A. Arias-Montaño, "Dopamina: síntesis, liberación y receptores en el Sistema Nervioso Central," *REVISTA BIOMÉDICA*, vol. 11, no. 1, pp. 39–60, 2000. Number: 1.
- [7] P. Rizek, N. Kumar, and M. S. Jog, "An update on the diagnosis and treatment of Parkinson disease," *CMAJ*, vol. 188, pp. 1157–1165, Nov. 2016.
- [8] M. C. de Rijk, C. Tzourio, M. M. Breteler, J. F. Dartigues, L. Amaducci, S. Lopez-Pousa, J. M. Manubens-Bertran, A. Alperovitch, and W. A. Rocca, "Prevalence of parkinsonism and Parkinson's disease in Europe: the EUROPARKINSON Collaborative Study. European Community Concerted Action on the Epidemiology of Parkinson's disease," *Journal of Neurology, Neurosurgery, and Psychiatry*, vol. 62, pp. 10–15, Jan. 1997.

- [9] D. Twelves, K. S. M. Perkins, and C. Counsell, "Systematic review of incidence studies of Parkinson's disease," *Movement Disorders: Official Journal of the Movement Disorder Society*, vol. 18, pp. 19–31, Jan. 2003.
- [10] G. Petzinger, D. Holschneider, B. Fisher, S. McEwen, N. Kintz, M. Halliday, W. Toy, J. Walsh, J. Beeler, and M. Jakowec, "The Effects of Exercise on Dopamine Neurotransmission in Parkinson's Disease: Targeting Neuroplasticity to Modulate Basal Ganglia Circuitry," *Brain Plasticity*, vol. 1, no. 1, pp. 29–39, 2015.
- [11] J. Parkinson, "An essay on the shaking palsy. 1817," *The Journal of Neuropsychiatry and Clinical Neurosciences*, vol. 14, no. 2, pp. 223–236; discussion 222, 2002.
- [12] C. G. Goetz, "The history of Parkinson's disease: early clinical descriptions and neurological therapies," *Cold Spring Harbor Perspectives in Medicine*, vol. 1, p. a008862, Sept. 2011.
- [13] L. V. Kalia and A. E. Lang, "Parkinson's disease," *The Lancet*, vol. 386, pp. 896–912, Aug. 2015.
- [14] A. Berardelli, J. C. Rothwell, P. D. Thompson, and M. Hallett, "Pathophysiology of bradykinesia in Parkinson's disease," *Brain: A Journal of Neurology*, vol. 124, pp. 2131–2146, Nov. 2001.
- [15] M. Bologna, G. Fabbrini, L. Marsili, G. Defazio, P. D. Thompson, and A. Berardelli, "Facial bradykinesia," *Journal of Neurology, Neurosurgery, and Psychiatry*, vol. 84, pp. 681–685, June 2013.
- [16] Y.-T. Chang, W.-N. Chang, N.-W. Tsai, K.-Y. Cheng, C.-C. Huang, C.-T. Kung, Y.-J. Su, W.-C. Lin, B.-C. Cheng, C.-M. Su, Y.-F. Chiang, and C.-H. Lu, "Clinical Features Associated with Frozen Shoulder Syndrome in Parkinson's Disease," *Parkinson's Disease*, vol. 2015, p. 232958, 2015.
- [17] A. D. Ha and J. Jankovic, "Pain in Parkinson's disease," *Movement Disorders: Official Journal of the Movement Disorder Society*, vol. 27, pp. 485–491, Apr. 2012.
- [18] J. Jankovic, "Progression of Parkinson Disease: Are We Making Progress in Charting the Course?," *Archives of Neurology*, vol. 62, pp. 351–352, Mar. 2005.
- [19] J. M. Beitz, "Parkinson's disease: a review," *Frontiers in Bioscience-Scholar*, vol. 6, pp. 65–74, Jan. 2014.

- [20] H. Braak, K. D. Tredici, U. Rüb, R. A. I. de Vos, E. N. H. Jansen Steur, and E. Braak, "Staging of brain pathology related to sporadic Parkinson's disease," *Neurobiology of Aging*, vol. 24, pp. 197–211, Mar. 2003.
- [21] R. L. Doty, "Olfactory dysfunction in Parkinson disease," *Nature Reviews Neurology*, vol. 8, pp. 329–339, June 2012.
- [22] R. B. Postuma, D. Aarsland, P. Barone, D. J. Burn, C. H. Hawkes, W. Oertel, and T. Ziemssen, "Identifying prodromal Parkinson's disease: Pre-Motor disorders in Parkinson's disease," *Movement Disorders*, vol. 27, no. 5, pp. 617–626, 2012.
- [23] T. K. Khoo, A. J. Yarnall, G. W. Duncan, S. Coleman, J. T. O'Brien, D. J. Brooks, R. A. Barker, and D. J. Burn, "The spectrum of nonmotor symptoms in early Parkinson disease," *Neurology*, vol. 80, pp. 276–281, Jan. 2013.
- [24] Y. Niimi, S. Ito, K. Murate, S. Hirota, C. Hikichi, T. Ishikawa, T. Maeda, R. Nagao, S. Shima, Y. Mizutani, A. Ueda, and T. Mutoh, "Usefulness of combining 123I-FP-CIT-SPECT striatal asymmetry index and cardiac 123I-metaiodobenzylguanidine scintigraphy examinations for diagnosis of parkinsonisms," *Journal of the Neurological Sciences*, vol. 377, pp. 174–178, June 2017.
- [25] C. Scherfler, K. Seppi, K. J. Mair, E. Donnemiller, I. Virgolini, G. K. Wenning, and W. Poewe, "Left hemispheric predominance of nigrostriatal dysfunction in Parkinson's disease," *Brain*, vol. 135, pp. 3348–3354, Nov. 2012.
- [26] K. C. Stewart, H. H. Fernandez, M. S. Okun, R. L. Rodriguez, C. E. Jacobson, and C. J. Hass, "Side onset influences motor impairments in Parkinson disease," *Parkinsonism & Related Disorders*, vol. 15, pp. 781–783, Dec. 2009.
- [27] A. Tarakad and J. Jankovic, "Diagnosis and Management of Parkinson's Disease," *Seminars in Neurology*, vol. 37, pp. 118–126, Apr. 2017.
- [28] M. A. Thenganatt and J. Jankovic, "The relationship between essential tremor and Parkinson's disease," *Parkinsonism & Related Disorders*, vol. 22 Suppl 1, pp. S162–165, Jan. 2016.
- [29] R. B. Postuma, D. Berg, M. Stern, W. Poewe, C. W. Olanow, W. Oertel, J. Obeso, K. Marek, I. Litvan, A. E. Lang, G. Halliday, C. G. Goetz, T. Gasser, B. Dubois, P. Chan, B. R. Bloem, C. H. Adler, and G. Deuschl, "MDS clinical diagnostic criteria for Parkinson's disease," *Movement Disorders*, vol. 30, no. 12, pp. 1591–1601, 2015.

- [30] C. H. Adler, T. G. Beach, J. G. Hentz, H. A. Shill, J. N. Caviness, E. Driver-Dunckley, M. N. Sabbagh, L. I. Sue, S. A. Jacobson, C. M. Belden, and B. N. Dugger, "Low clinical diagnostic accuracy of early vs advanced Parkinson disease," *Neurology*, vol. 83, pp. 406–412, July 2014.
- [31] T. G. Beach and C. H. Adler, "Importance of Low Diagnostic Accuracy for Early Parkinson's Disease," *Movement disorders : official journal of the Movement Disorder Society*, vol. 33, pp. 1551–1554, Oct. 2018.
- [32] A. J. Hughes, S. E. Daniel, and A. J. Lees, "Improved accuracy of clinical diagnosis of Lewy body Parkinson's disease," *Neurology*, vol. 57, pp. 1497–1499, Oct. 2001.
- [33] J. Jankovic, "Parkinson's disease: clinical features and diagnosis," *Journal of Neurology, Neurosurgery, and Psychiatry*, vol. 79, pp. 368–376, Apr. 2008.
- [34] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *Journal of Neural Engineering*, vol. 16, p. 051001, Aug. 2019.
- [35] J. Sheng, B. Wang, Q. Zhang, Q. Liu, Y. Ma, W. Liu, M. Shao, and B. Chen, "A novel joint HCPMMP method for automatically classifying Alzheimer's and different stage MCI patients," *Behavioural Brain Research*, vol. 365, pp. 210–221, June 2019.
- [36] R. Prashanth and S. Dutta Roy, "Novel and improved stage estimation in Parkinson's disease using clinical scales and machine learning," *Neurocomputing*, vol. 305, pp. 78–103, Aug. 2018.
- [37] O. Rascol, D. J. Brooks, A. D. Korczyn, P. P. De Deyn, C. E. Clarke, and A. E. Lang, "A five-year study of the incidence of dyskinesia in patients with early Parkinson's disease who were treated with ropinirole or levodopa," *The New England Journal of Medicine*, vol. 342, pp. 1484–1491, May 2000.
- [38] C. G. Goetz, B. C. Tilley, S. R. Shaftman, G. T. Stebbins, S. Fahn, P. Martinez-Martin, W. Poewe, C. Sampaio, M. B. Stern, R. Dodel, B. Dubois, R. Holloway, J. Jankovic, J. Kulisevsky, A. E. Lang, A. Lees, S. Leurgans, P. A. LeWitt, D. Nyenhuis, C. W. Olanow, O. Rascol, A. Schrag, J. A. Teresi, J. J. van Hilten, N. LaPelle, and Movement Disorder Society UPDRS Revision Task Force, "Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results," *Movement Disorders: Official Journal of the Movement Disorder Society*, vol. 23, pp. 2129–2170, Nov. 2008.



- [39] Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease, "The Unified Parkinson's Disease Rating Scale (UPDRS): status and recommendations," *Movement Disorders: Official Journal of the Movement Disorder Society*, vol. 18, pp. 738–750, July 2003.
- [40] A. Schrag and N. Quinn, "Dyskinesias and motor fluctuations in Parkinson's disease. A community-based study," *Brain: A Journal of Neurology*, vol. 123 ( Pt 11), pp. 2297–2305, Nov. 2000.
- [41] C. G. Goetz, W. Poewe, O. Rascol, C. Sampaio, G. T. Stebbins, C. Counsell, N. Giladi, R. G. Holloway, C. G. Moore, G. K. Wenning, M. D. Yahr, L. Seidl, and Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease, "Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations," *Movement Disorders: Official Journal of the Movement Disorder Society*, vol. 19, pp. 1020–1028, Sept. 2004.
- [42] J. Booij, G. Tissingh, G. J. Boer, J. D. Speelman, J. C. Stoof, A. G. Janssen, E. C. Wolters, and E. A. van Royen, "[123I]FP-CIT SPECT shows a pronounced decline of striatal dopamine transporter labelling in early and advanced Parkinson's disease," *Journal of Neurology, Neurosurgery, and Psychiatry*, vol. 62, pp. 133–140, Feb. 1997.
- [43] L. J. Evers, J. H. Krijthe, M. J. Meinders, B. R. Bloem, and T. M. Heskes, "Measuring Parkinson's disease over time: The real-world within-subject reliability of the MDS-UPDRS," *Movement Disorders*, vol. 34, no. 10, pp. 1480–1487, 2019.
- [44] R. M. Kostrzewa, P. Nowak, J. P. Kostrzewa, R. A. Kostrzewa, and R. Brus, "Peculiarities of L-DOPA treatment of Parkinson's disease," *Amino Acids*, vol. 28, pp. 157–164, Mar. 2005.
- [45] J. A. Obeso, M. C. Rodriguez-Oroz, C. G. Goetz, C. Marin, J. H. Kordower, M. Rodriguez, E. C. Hirsch, M. Farrer, A. H. V. Schapira, and G. Halliday, "Missing pieces in the Parkinson's disease puzzle," *Nature Medicine*, vol. 16, pp. 653–661, June 2010.
- [46] E. Tolosa, M. J. Martí, F. Valldeoriola, and J. L. Molinuevo, "History of levodopa and dopamine agonists in Parkinson's disease treatment," *Neurology*, vol. 50, pp. S2–10; discussion S44–48, June 1998.
- [47] U. Raghavendra, U. R. Acharya, and H. Adeli, "Artificial Intelligence Techniques for Automated Diagnosis of Neurological Disorders," *European Neurology*, vol. 82, no. 1-3, pp. 41–64, 2019.

- [48] J. E. Ahlskog and M. D. Muentner, "Frequency of levodopa-related dyskinesias and motor fluctuations as estimated from the cumulative literature," *Movement Disorders: Official Journal of the Movement Disorder Society*, vol. 16, pp. 448–458, May 2001.
- [49] M. S. Okun, "Deep-brain stimulation for Parkinson's disease," *The New England Journal of Medicine*, vol. 367, pp. 1529–1538, Oct. 2012.
- [50] M. S. Okun and K. D. Foote, "Parkinson's disease DBS: what, when, who and why? The time has come to tailor DBS targets," *Expert Review of Neurotherapeutics*, vol. 10, pp. 1847–1857, Dec. 2010.
- [51] H. S. Dafsari, P. Reddy, C. Herchenbach, S. Wawro, J. N. Petry-Schmelzer, V. Visser-Vandewalle, A. Rizos, M. Silverdale, K. Ashkan, M. Samuel, J. Evans, C. A. Huber, G. R. Fink, A. Antonini, K. R. Chaudhuri, P. Martinez-Martin, L. Timmermann, and IPMDS Non-Motor Symptoms Study Group, "Beneficial Effects of Bilateral Subthalamic Stimulation on Non-Motor Symptoms in Parkinson's Disease," *Brain Stimulation*, vol. 9, pp. 78–85, Feb. 2016.
- [52] C. Loane and M. Politis, "Positron emission tomography neuroimaging in Parkinson's disease," *American Journal of Translational Research*, vol. 3, pp. 323–341, Aug. 2011.
- [53] O. Rascol and L. Schelosky, "<sup>123</sup>I-metaiodobenzylguanidine scintigraphy in Parkinson's disease and related disorders," *Movement Disorders*, vol. 24, no. S2, pp. S732–S741, 2009.
- [54] S. Orimo, A. Takahashi, T. Uchihara, F. Mori, A. Kakita, K. Wakabayashi, and H. Takahashi, "Degeneration of Cardiac Sympathetic Nerve Begins in the Early Disease Process of Parkinson's Disease," *Brain Pathology*, vol. 17, no. 1, pp. 24–30, 2007.
- [55] S. Braune, M. Reinhardt, R. Schnitzer, A. Riedel, and C. H. Lücking, "Cardiac uptake of [<sup>123</sup>I]MIBG separates Parkinson's disease from multiple system atrophy," *Neurology*, vol. 53, pp. 1020–1020, Sept. 1999.
- [56] G. Treglia, A. Stefanelli, E. Cason, F. Cocciolillo, D. Di Giuda, and A. Giordano, "Diagnostic performance of iodine-123-metaiodobenzylguanidine scintigraphy in differential diagnosis between Parkinson's disease and multiple-system atrophy: A systematic review and a meta-analysis," *Clinical Neurology and Neurosurgery*, vol. 113, pp. 823–829, Dec. 2011.

- [57] A. E. King, J. Mintz, and D. R. Royall, "Meta-analysis of 123I-MIBG cardiac scintigraphy for the diagnosis of Lewy body-related disorders," *Movement Disorders*, vol. 26, no. 7, pp. 1218–1224, 2011.
- [58] A. Berardelli, G. K. Wenning, A. Antonini, D. Berg, B. R. Bloem, V. Bonifati, D. Brooks, D. J. Burn, C. Colosimo, A. Fanciulli, J. Ferreira, T. Gasser, F. Grandas, P. Kanovsky, V. Kostic, J. Kulisevsky, W. Oertel, W. Poewe, J.-P. Reese, M. Relja, E. Ruzicka, A. Schrag, K. Seppi, P. Taba, and M. Vidailhet, "EFNS/MDS-ES recommendations for the diagnosis of Parkinson's disease," *European Journal of Neurology*, vol. 20, no. 1, pp. 16–34, 2013.
- [59] R. Hilker, N. Razai, M. Ghaemi, S. Weisenbach, J. Rudolf, B. Szelies, and W.-D. Heiss, "[18F]fluorodopa uptake in the upper brainstem measured with positron emission tomography correlates with decreased REM sleep duration in early Parkinson's disease," *Clinical Neurology and Neurosurgery*, vol. 105, pp. 262–269, Sept. 2003.
- [60] P. Piccini, D. J. Burn, R. Ceravolo, D. Maraganore, and D. J. Brooks, "The role of inheritance in sporadic Parkinson's disease: evidence from a longitudinal study of dopaminergic function in twins," *Annals of Neurology*, vol. 45, pp. 577–582, May 1999.
- [61] P. Remy, M. Doder, A. Lees, N. Turjanski, and D. Brooks, "Depression in Parkinson's disease: loss of dopamine and noradrenaline innervation in the limbic system," *Brain: A Journal of Neurology*, vol. 128, pp. 1314–1322, June 2005.
- [62] L. Wang, Q. Zhang, H. Li, and H. Zhang, "SPECT Molecular Imaging in Parkinson's Disease," *Journal of Biomedicine and Biotechnology*, vol. 2012, p. 412486, 2012.
- [63] K. L. Chou, H. I. Hurtig, M. B. Stern, A. Colcher, B. Ravina, A. Newberg, P. D. Mozley, and A. Siderowf, "Diagnostic accuracy of [99mTc]TRODAT-1 SPECT imaging in early Parkinson's disease," *Parkinsonism & Related Disorders*, vol. 10, pp. 375–379, Aug. 2004.
- [64] A. C. Felicio, C. Godeiro-Junior, M. C. Shih, V. Borges, S. M. A. Silva, P. d. C. Aguiar, M. Q. Hoexter, O. G. P. Barsottini, L. A. F. Andrade, R. A. Bressan, and H. B. Ferraz, "Evaluation of patients with Clinically Unclear Parkinsonian Syndromes submitted to brain SPECT imaging using the technetium-99m labeled tracer TRODAT-1," *Journal of the Neurological Sciences*, vol. 291, pp. 64–68, Apr. 2010.

- [65] V. Marshall and D. Grosset, "Role of dopamine transporter imaging in routine clinical practice," *Movement Disorders: Official Journal of the Movement Disorder Society*, vol. 18, pp. 1415–1423, Dec. 2003.
- [66] W.-C. Shyu, S.-Z. Lin, M.-F. Chiang, C.-Y. Pang, S.-Y. Chen, Y.-L. Hsin, P. Thajeb, Y.-J. Lee, and H. Li, "Early-onset Parkinson's disease in a Chinese population: 99mTc-TRODAT-1 SPECT, Parkin gene analysis and clinical study," *Parkinsonism & Related Disorders*, vol. 11, pp. 173–180, May 2005.
- [67] W. S. Huang, S. Z. Lin, J. C. Lin, S. P. Wey, G. Ting, and R. S. Liu, "Evaluation of early-stage Parkinson's disease with 99mTc-TRODAT-1 imaging," *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine*, vol. 42, pp. 1303–1308, Sept. 2001.
- [68] C. Skowronek, L. Zange, and A. Lipp, "Cardiac 123I-MIBG Scintigraphy in Neurodegenerative Parkinson Syndromes: Performance and Pitfalls in Clinical Practice," *Frontiers in Neurology*, vol. 10, 2019.
- [69] K. Tsujikawa, Y. Hasegawa, S. Yokoi, K. Yasui, I. Nanbu, T. Yanagi, and A. Takahashi, "Chronological changes of 123I-MIBG myocardial scintigraphy and clinical features of Parkinson's disease," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 86, pp. 945–951, Sept. 2015.
- [70] J.-S. Kim, H.-E. Park, I.-S. Park, Y.-S. Oh, D.-W. Ryu, I.-U. Song, Y.-A. Jung, I. R. Yoo, H.-S. Choi, P. H. Lee, and K.-S. Lee, "Normal 'heart' in Parkinson's disease: is this a distinct clinical phenotype?," *European Journal of Neurology*, vol. 24, no. 2, pp. 349–356, 2017.
- [71] H. Matsui, K. Nishinaka, M. Oda, K. Komatsu, T. Kubori, and F. Udaka, "Does cardiac metaiodobenzylguanidine (MIBG) uptake in Parkinson's disease correlate with major autonomic symptoms?," *Parkinsonism & Related Disorders*, vol. 12, pp. 284–288, June 2006.
- [72] A. Umemura, T. Oeda, R. Hayashi, S. Tomita, M. Kohsaka, K. Yamamoto, and H. Sawada, "Diagnostic accuracy of apparent diffusion coefficient and 123I-metaiodobenzylguanidine for differentiation of multiple system atrophy and parkinson's disease," *PLOS ONE*, vol. 8, pp. 1–7, April 2013.
- [73] H. Sawada, T. Oeda, K. Yamamoto, N. Kitagawa, E. Mizuta, R. Hosokawa, M. Ohba, R. Nishio, K. Yamakawa, H. Takeuchi, S. Shimohama, R. Takahashi, and T. Kawamura, "Diagnostic accuracy of cardiac metaiodobenzylguanidine scintigraphy in

- Parkinson disease,” *European Journal of Neurology*, vol. 16, pp. 174–182, Feb. 2009.
- [74] J. Puñal Riobóo, A. Serena-Puig, L. Varela-Lema, A. M. Alvarez-Páez, and A. Ruano-Ravina, “Clinical utility of (18)F-DOPA-PET in movement disorders. A systematic review,” *Revista Española De Medicina Nuclear*, vol. 28, pp. 106–113, June 2009.
- [75] P. Sasannezhad, A. G. Juibary, K. Sadri, R. Sadeghi, M. Sabour, V. R. D. Kakhki, and H. Alizadeh, “99mTc-TRODAT-1 SPECT Imaging in Early and Late Onset Parkinson’s Disease,” *Asia Oceania Journal of Nuclear Medicine and Biology*, vol. 5, no. 2, pp. 114–119, 2017.
- [76] J. Eerola, P. J. Tienari, S. Kaakkola, P. Nikkinen, and J. Launes, “How useful is [123I]beta-CIT SPECT in clinical practice?,” *Journal of Neurology, Neurosurgery, and Psychiatry*, vol. 76, pp. 1211–1216, Sept. 2005.
- [77] V. L. Marshall, C. B. Reiningger, M. Marquardt, J. Patterson, D. M. Hadley, W. H. Oertel, H. T. S. Benamer, P. Kemp, D. Burn, E. Tolosa, J. Kulisevsky, L. Cunha, D. Costa, J. Booij, K. Tatsch, K. R. Chaudhuri, G. Ulm, O. Pogarell, H. Höffken, A. Gerstner, and D. G. Grosset, “Parkinson’s disease is overdiagnosed clinically at baseline in diagnostically uncertain cases: a 3-year European multicenter study with repeat [123I]FP-CIT SPECT,” *Movement Disorders: Official Journal of the Movement Disorder Society*, vol. 24, pp. 500–508, Mar. 2009.
- [78] H. Nagasawa, H. Tanji, Y. Itoyama, H. Saito, I. Kimura, T. Fujiwara, R. Iwata, M. Itoh, and T. Ido, “Brain 6-[18F]fluorodopa metabolism in early and late onset of Parkinson’s disease studied by positron emission tomography,” *Journal of the Neurological Sciences*, vol. 144, pp. 70–76, Dec. 1996.
- [79] R. De La Fuente-Fernández, A. S. Lim, V. Sossi, M. J. Adam, T. J. Ruth, D. B. Calne, A. J. Stoessl, and C. S. Lee, “Age and severity of nigrostriatal damage at onset of Parkinson’s disease,” *Synapse (New York, N.Y.)*, vol. 47, pp. 152–158, Feb. 2003.
- [80] G. R. Uhl, “Neurotransmitter transporters (plus): a promising new gene family,” *Trends in Neurosciences*, vol. 15, pp. 265–268, July 1992.
- [81] M. J. Kaufman and B. K. Madras, “Severe depletion of cocaine recognition sites associated with the dopamine transporter in Parkinson’s-diseased striatum,” *Synapse (New York, N.Y.)*, vol. 9, pp. 43–49, Sept. 1991.

- [82] G. Pagano, F. Niccolini, and M. Politis, "Imaging in Parkinson's disease," *Clinical Medicine*, vol. 16, pp. 371–375, Aug. 2016.
- [83] R. C. Helmich, D. E. Vaillancourt, and D. J. Brooks, "The Future of Brain Imaging in Parkinson's Disease," *Journal of Parkinson's Disease*, vol. 8, no. Suppl 1, pp. S47–S51, 2018.
- [84] R. Balestrino and A. Schapira, "Parkinson disease," *European Journal of Neurology*, vol. 27, no. 1, pp. 27–42, 2020.
- [85] A. de Micheli, "En torno a los primeros estudios de electrofisiología," *Archivos de Cardiología de México*, vol. 81, pp. 337–342, Oct. 2011.
- [86] A. v. Haller, *Elementa physiologiae corporis humani*. Lausannae: Sigismundi D'Arnay, 1757-1759.
- [87] J. Walsh and S. Seignette, *Of the Electric Property of the Torpedo. In a Letter from John Walsh, Esq; F. R. S. to Benjamin Franklin, Esq; LL.D., F. R. S., Ac. R. Par. Soc. Ext., &c.* Royal Society of London, Jan. 1773.
- [88] E. Niedermeyer and F. H. Lopes da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*. Philadelphia: Lippincott Williams & Wilkins, 2005.
- [89] L. Galvani, A. Volta, J. Zambelli, and d. D. Burndy Library, *Aloysii Galvani De viribus electricitatis in motu musculari commentarius*. Bononiae : Ex Typographia Institutii Scientiarium, 1791.
- [90] G. J. Tortora and B. Derrickson, *Principios de anatomía y fisiología*. Buenos Aires: Editorial Médica Panamericana, Ed. 3, 2013.
- [91] J. E. Hall, *Guyton y Hall. Tratado de fisiología médica*. Elsevier Health Sciences, Aug. 2011.
- [92] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of Physiology*, vol. 117, no. 4, pp. 500–544, 1952.
- [93] A. Coenen, E. Fine, and O. Zayachkivska, "Adolf Beck: A Forgotten Pioneer in Electroencephalography," *Journal of the History of the Neurosciences*, vol. 23, pp. 276–286, July 2014.
- [94] v. W. W. Neminsky, "Ein Versuch der Registrierung der elektrischen Gehirnerscheinungen," *Zentralblatt für Physiologie*, vol. 27, pp. 951–960, 1913.

- [95] L. F. Haas, "Hans Berger (1873-1941), Richard Caton (1842-1926), and electroencephalography," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 74, no. 1, pp. 9–9, 2003.
- [96] D. Millet, "The origins of EEG," in *7th Annual Meeting of the International Society for the History of the Neurosciences*, (Los Angeles, California, USA), ISHN, July 2002.
- [97] H. Berger, "Über das Elektrenkephalogramm des Menschen," *Archiv für Psychiatrie und Nervenkrankheiten*, vol. 87, pp. 527–570, Dec. 1929.
- [98] "EEG Machine." <https://www.emotiv.com/glossary/eeg-machine/>. Accedido: 13-12-2022.
- [99] F. A. Gibbs, H. Davis, and W. G. Lennox, "The electro-encephalogram in epilepsy and in conditions of impaired consciousness," *Archives of Neurology & Psychiatry*, vol. 34, pp. 1133–1148, Dec. 1935.
- [100] S. Noachtar and J. Rémi, "The role of EEG in epilepsy: a critical review," *Epilepsy & Behavior: E&B*, vol. 15, pp. 22–33, May 2009.
- [101] W. O. Tatum, *Handbook of EEG interpretation*. Demos Medical, 2014.
- [102] C. C. Chernecky and B. J. Berger, *Laboratory Tests and Diagnostic Procedures*. Elsevier Health Sciences, Oct. 2012.
- [103] "EEG: MedlinePlus Medical Encyclopedia." <https://medlineplus.gov/ency/article/003931.htm>. Accedido: 18-08-2022.
- [104] E. A. Wolpert, "A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects.," *Archives of General Psychiatry*, vol. 20, pp. 246–247, Feb. 1969.
- [105] C. Kwasnica, "Evoked Potentials," in *Encyclopedia of Clinical Neuropsychology* (J. S. Kreutzer, J. DeLuca, and B. Caplan, eds.), pp. 986–986, New York, NY: Springer, 2011.
- [106] S. J. Luck, *An Introduction to the Event-Related Potential Technique*. MIT Press. Ed. 2., 2014.
- [107] P. Fries, "A mechanism for cognitive dynamics: neuronal communication through neuronal coherence," *Trends in Cognitive Sciences*, vol. 9, pp. 474–480, Oct. 2005.

- [108] J. Fell and N. Axmacher, "The role of phase synchronization in memory processes," *Nature Reviews Neuroscience*, vol. 12, pp. 105–118, Feb. 2011.
- [109] A. Schnitzler and J. Gross, "Normal and pathological oscillatory communication in the brain," *Nature Reviews Neuroscience*, vol. 6, pp. 285–296, Apr. 2005.
- [110] D. Baldauf and R. Desimone, "Neural Mechanisms of Object-Based Attention," *Science*, vol. 344, pp. 424–427, Apr. 2014.
- [111] M. Nuwer, "Assessment of digital EEG, quantitative EEG, and EEG brain mapping: report of the American Academy of Neurology and the American Clinical Neurophysiology Society," *Neurology*, vol. 49, pp. 277–292, July 1997.
- [112] L. Livint Popa, H. Dragos, C. Pantelemon, O. Verisezan Rosu, and S. Strilciuc, "The Role of Quantitative EEG in the Diagnosis of Neuropsychiatric Disorders," *Journal of Medicine and Life*, vol. 13, no. 1, pp. 8–15, 2020.
- [113] W. G. Walter, *The Living Brain*. Penguin Books, 1961.
- [114] P. M. Rossini, R. Di Iorio, M. Bentivoglio, G. Bertini, F. Ferreri, C. Gerloff, R. J. Ilmoniemi, F. Miraglia, M. A. Nitsche, F. Pestilli, M. Rosanova, Y. Shirota, C. Tesoriero, Y. Ugawa, F. Vecchio, U. Ziemann, and M. Hallett, "Methods for analysis of brain connectivity: An IFCN-sponsored review," *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, vol. 130, pp. 1833–1858, Oct. 2019.
- [115] B. C. Lega, J. Jacobs, and M. Kahana, "Human hippocampal theta oscillations and the formation of episodic memories," *Hippocampus*, vol. 22, no. 4, pp. 748–761, 2012.
- [116] G. McLoughlin, M. Gyurkovics, J. Palmer, and S. Makeig, "Midfrontal Theta Activity in Psychiatric Illness: An Index of Cognitive Vulnerabilities Across Disorders," *Biological Psychiatry*, vol. 91, pp. 173–182, Jan. 2022.
- [117] R. İnce, S. S. Adanır, and F. Sevmez, "The inventor of electroencephalography (EEG): Hans Berger (1873-1941)," *Child's Nervous System*, vol. 37, pp. 2723–2724, Sept. 2021.
- [118] S. N. Baker, "Oscillatory interactions between sensorimotor cortex and the periphery," *Current Opinion in Neurobiology*, vol. 17, pp. 649–655, Dec. 2007.



- [119] J. A. van Deursen, E. F. P. M. Vuurman, F. R. J. Verhey, V. H. J. M. van Kranen-Mastenbroek, and W. J. Riedel, "Increased EEG gamma band activity in Alzheimer's disease and mild cognitive impairment," *Journal of Neural Transmission*, vol. 115, pp. 1301–1311, Sept. 2008.
- [120] J. R. Hughes, "Gamma, fast, and ultrafast waves of the brain: Their relationships with epilepsy and behavior," *Epilepsy & Behavior*, vol. 13, pp. 25–31, July 2008.
- [121] X. Jia and A. Kohn, "Gamma Rhythms in the Brain," *PLOS Biology*, vol. 9, p. e1001045, Apr. 2011.
- [122] Y. Höller, C. Helmstaedter, and K. Lehnertz, "Quantitative Pharmacoelectroencephalography in Antiepileptic Drug Research," *CNS drugs*, vol. 32, pp. 839–848, Sept. 2018.
- [123] G. M. Tedrus, L. M. Negreiros, R. S. Ballarim, T. A. Marques, and L. C. Fonseca, "Correlations Between Cognitive Aspects and Quantitative EEG in Adults With Epilepsy," *Clinical EEG and neuroscience*, vol. 50, pp. 348–353, Sept. 2019.
- [124] V. Köpruner, G. Pfurtscheller, and L. M. Auer, "Quantitative EEG in normals and in patients with cerebral ischemia," *Progress in Brain Research*, vol. 62, pp. 29–50, 1984.
- [125] S. Kanna and J. Heng, "Quantitative EEG parameters for monitoring and biofeedback during rehabilitation after stroke," in *2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pp. 1689–1694, July 2009. ISSN: 2159-6255.
- [126] M. R. Nuwer, G. Comi, R. Emerson, A. Fuglsang-Frederiksen, J. M. Guérit, H. Hinrichs, A. Ikeda, F. J. Luccas, and P. Rappelsburger, "IFCN standards for digital recording of clinical EEG," *Electroencephalography and Clinical Neurophysiology*, vol. 106, pp. 259–261, Mar. 1998.
- [127] K. Maurer, *Imaging of the Brain in Psychiatry and Related Fields*. Springer Science & Business Media, Dec. 2012.
- [128] J. A. Coan and J. J. B. Allen, "Frontal EEG asymmetry as a moderator and mediator of emotion," *Biological Psychology*, vol. 67, pp. 7–49, Oct. 2004.
- [129] P. A. d. M. Kanda, R. Anghinah, M. T. Smidth, and J. M. Silva, "The clinical use of quantitative EEG in cognitive disorders," *Dementia & Neuropsychologia*, vol. 3, pp. 195–203, Sept. 2009.

- [130] K. L. Coburn, E. C. Lauterbach, N. N. Boutros, K. J. Black, D. B. Arciniegas, and C. E. Coffey, "The value of quantitative electroencephalography in clinical psychiatry: a report by the Committee on Research of the American Neuropsychiatric Association," *The Journal of Neuropsychiatry and Clinical Neurosciences*, vol. 18, no. 4, pp. 460–500, 2006.
- [131] A. S. Widge, M. T. Bilge, R. Montana, W. Chang, C. I. Rodriguez, T. Deckersbach, L. L. Carpenter, N. H. Kalin, and C. B. Nemeroff, "Electroencephalographic Biomarkers for Treatment Response Prediction in Major Depressive Illness: A Meta-Analysis," *The American Journal of Psychiatry*, vol. 176, pp. 44–56, Jan. 2019.
- [132] J. R. Hughes and E. R. John, "Conventional and quantitative electroencephalography in psychiatry," *The Journal of Neuropsychiatry and Clinical Neurosciences*, vol. 11, no. 2, pp. 190–208, 1999.
- [133] S. M. Bresnahan, J. W. Anderson, and R. J. Barry, "Age-related changes in quantitative EEG in attention-deficit/hyperactivity disorder," *Biological Psychiatry*, vol. 46, pp. 1690–1697, Dec. 1999.
- [134] D.-H. Jeong, Y.-D. Kim, I.-U. Song, Y.-A. Chung, and J. Jeong, "Wavelet Energy and Wavelet Coherence as EEG Biomarkers for the Diagnosis of Parkinson's Disease-Related Dementia and Alzheimer's Disease," *Entropy*, vol. 18, p. 8, Jan. 2016.
- [135] S. Q. A. Rizvi, G. Wang, and X. Xing, "Early Detection of Parkinson Disease Using Wavelet Transform Along with Fourier Transform," in *Smart City and Informatization* (G. Wang, A. El Saddik, X. Lai, G. Martinez Perez, and K.-K. R. Choo, eds.), Communications in Computer and Information Science, (Singapore), pp. 323–333, Springer, 2019.
- [136] R. Yuvaraj, M. Murugappan, N. Mohamed Ibrahim, M. Iqbal, K. Sundaraj, K. Mohamad, R. Palaniappan, E. Mesquita, and M. Satiyan, "On the analysis of EEG power, frequency and asymmetry in Parkinson's disease during emotion processing," *Behavioral and brain functions: BBF*, vol. 10, p. 12, Apr. 2014.
- [137] R. Soikkeli, J. Partanen, H. Soininen, A. Pääkkönen, and P. Riekkinen, "Slowing of EEG in Parkinson's disease," *Electroencephalography and Clinical Neurophysiology*, vol. 79, pp. 159–165, Sept. 1991.
- [138] C.-X. Han, J. Wang, G.-S. Yi, and Y.-Q. Che, "Investigation of EEG abnormalities in the early stage of Parkinson's disease," *Cognitive Neurodynamics*, vol. 7, pp. 351–359, Aug. 2013.

- [139] N. Bigdely-Shamlo, T. Mullen, C. Kothe, K.-M. Su, and K. A. Robbins, "The PREP pipeline: standardized preprocessing for large-scale EEG analysis," *Frontiers in Neuroinformatics*, vol. 9, p. 16, 2015.
- [140] M. Jas, D. A. Engemann, Y. Bekhti, F. Raimondo, and A. Gramfort, "Autoreject: Automated artifact rejection for MEG and EEG data," *NeuroImage*, vol. 159, pp. 417–429, Oct. 2017.
- [141] F. A. Rahman, M. F. Othman, and N. A. Shaharuddin, "Analysis Methods of EEG Signals: A Review in EEG Application for Brain Disorder," *Jurnal Teknologi*, vol. 72, Jan. 2015. Number: 2.
- [142] A. Gramfort, D. Strohmeier, J. Haueisen, M. S. Hämäläinen, and M. Kowalski, "Time-frequency mixed-norm estimates: sparse M/EEG imaging with non-stationary source activations," *NeuroImage*, vol. 70, pp. 410–422, Apr. 2013.
- [143] S. Cole and B. Voytek, "Cycle-by-cycle analysis of neural oscillations," *Journal of Neurophysiology*, vol. 122, pp. 849–861, Aug. 2019.
- [144] D. Acharjya, A. Mitra, and N. Zaman, *Deep Learning in Data Analytics. Recent Techniques, Practices and Applications*. Springer Nature, 2022.
- [145] "Definición de inteligencia por la real academia española.." <https://dle.rae.es/inteligencia>. Accedido: 22-10-2022.
- [146] H. Gardner, *Frames of Mind*. New York: Basic Books, 1983.
- [147] J. Moor, "The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years," *AI Magazine*, vol. 27, pp. 87–87, Dec. 2006. Number: 4.
- [148] A. M. Turing, "Computing Machinery and Intelligence," *Mind*, vol. LIX, pp. 433–460, Oct. 1950.
- [149] J. Searle, "Minds, Brains, and Programs," *Behavioral and Brain Sciences*, vol. 3, no. 3, pp. 417–57, 1980.
- [150] D. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Reading, Massachusetts: Addison-Wesley, 1989.
- [151] L. A. Zadeh, "Fuzzy Logic, Neural Networks and Soft Computing," in *Safety Evaluation Based on Identification Approaches Related to Time-Variant and Nonlinear Structures*, pp. 320–321, Wiesbaden: Vieweg+Teubner Verlag, 1993.

- [152] A. L. Samuel, "Some Studies in Machine Learning Using the Game of Checkers. II-Recent Progress," in *Computer Games I* (D. N. L. Levy, ed.), pp. 366–400, New York, NY: Springer, 1988.
- [153] F. Gómez Quesada, M. Fernández Graciani, M. López Bonal, and M. Alonso Díaz-Mata, "Aprendizaje con redes neuronales artificiales," *Ensayos: Revista de la Facultad de Educación de Albacete*, vol. 9, pp. 169–180, 1994.
- [154] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning*. Cambridge, Massachusetts: MIT Press, 2 ed., 2018.
- [155] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*, vol. 43, pp. 59–69, Jan. 1982.
- [156] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, pp. 115–133, Dec. 1943.
- [157] S. Russell and P. Norvig, *Artificial Intelligence. A Modern Approach*. Essex, England: Pearson, 2016.
- [158] F. Rosenblat, *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Spartan Books, 1962.
- [159] M. Minsky and S. Papert, *Perceptrons: an introduction to computational geometry*. MIT Press, 1969.
- [160] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Singapore: Prentice Hall, 1999.
- [161] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, Oct. 1986.
- [162] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, Massachusetts: MIT Press, 2016 ed., 2016.
- [163] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014. arXiv:1412.6980. Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015.
- [164] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

- [165] M. V. Valueva, N. N. Nagornov, P. A. Lyakhov, G. V. Valuev, and N. I. Chervyakov, "Application of the residue number system to reduce hardware costs of the convolutional neural network implementation," *Mathematics and Computers in Simulation*, vol. 177, pp. 232–243, Nov. 2020.
- [166] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.
- [167] Y. LeCun, L. Jackel, B. Boser, J. Denker, H. Graf, I. Guyon, D. Henderson, R. Howard, and W. Hubbard, "Handwritten digit recognition: applications of neural network chips and automatic learning," *IEEE Communications Magazine*, vol. 27, no. 11, pp. 41–46, 1989.
- [168] S. E. Schaeffer, "Graph clustering," *Computer Science Review*, vol. 1, pp. 27–64, Aug. 2007.
- [169] S. Zhang, H. Tong, J. Xu, and R. Maciejewski, "Graph convolutional networks: a comprehensive review," *Computational Social Networks*, vol. 6, p. 11, Nov. 2019.
- [170] N. Kalchbrenner and P. Blunsom, "Recurrent continuous translation models," in *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, (Seattle, Washington, USA), pp. 1700–1709, Association for Computational Linguistics, 2013.
- [171] P. Koehn, *Neural Machine Translation*. Cambridge: Cambridge University Press, 2020.
- [172] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, pp. 1735–1780, Nov. 1997.
- [173] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention Is All You Need," Dec. 2017. arXiv:1706.03762 [cs].
- [174] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," May 2019. arXiv:1810.04805 [cs].
- [175] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving Language Understanding by Generative Pre-Training," *Online*, pp. 1–12, 2018.
- [176] F. Jiang, Y. Jiang, H. Zhi, Y. Dong, H. Li, S. Ma, Y. Wang, Q. Dong, H. Shen, and Y. Wang, "Artificial intelligence in healthcare: past, present and future," *Stroke and Vascular Neurology*, vol. 2, Dec. 2017.

- [177] A. Nogales, A. J. García-Tejedor, D. Monge, J. S. Vara, and C. Antón, “A survey of deep learning models in medical therapeutic areas,” *Artificial Intelligence in Medicine*, vol. 112, p. 102020, Feb. 2021.
- [178] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017.
- [179] S. Dhivya and A. Nithya, “A Review on Machine Learning Algorithm for EEG Signal Analysis,” in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 54–57, Mar. 2018.
- [180] V. Jahmunah, S. Lih Oh, V. Rajinikanth, E. J. Ciaccio, K. Hao Cheong, N. Arunkumar, and U. R. Acharya, “Automated detection of schizophrenia using nonlinear signal processing methods,” *Artificial Intelligence in Medicine*, vol. 100, p. 101698, Sept. 2019.
- [181] Y. Chen, C. Gong, H. Hao, Y. Guo, S. Xu, Y. Zhang, G. Yin, X. Cao, A. Yang, F. Meng, J. Ye, H. Liu, J. Zhang, Y. Sui, and L. Li, “Automatic Sleep Stage Classification Based on Subthalamic Local Field Potentials,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, pp. 118–128, Feb. 2019. Conference Name: IEEE Transactions on Neural Systems and Rehabilitation Engineering.
- [182] M. Zhou, C. Tian, R. Cao, B. Wang, Y. Niu, T. Hu, H. Guo, and J. Xiang, “Epileptic Seizure Detection Based on EEG Signals and CNN,” *Frontiers in Neuroinformatics*, vol. 12, 2018.
- [183] A. Vourvopoulos, C. Jorge, R. Abreu, P. Figueiredo, J.-C. Fernandes, and S. Bermúdez i Badia, “Efficacy and Brain Imaging Correlates of an Immersive Motor Imagery BCI-Driven VR System for Upper Limb Motor Rehabilitation: A Clinical Case Report,” *Frontiers in Human Neuroscience*, vol. 13, 2019.
- [184] L. E. Ismail and W. Karwowski, “A Graph Theory-Based Modeling of Functional Brain Connectivity Based on EEG: A Systematic Review in the Context of Neuroergonomics,” *IEEE Access*, vol. 8, pp. 155103–155135, 2020. Conference Name: IEEE Access.
- [185] A. Siderowf and A. E. Lang, “Premotor Parkinson’s disease: Concepts and definitions,” *Movement Disorders*, vol. 27, no. 5, pp. 608–616, 2012.

- [186] M. Suchmacher and M. Geller, "Chapter 1 - Study Type Determination," in *Practical Biostatistics* (M. Suchmacher and M. Geller, eds.), pp. 3–15, San Diego: Academic Press, Jan. 2012.
- [187] D. K. Arnett and S. A. Claas, "Chapter 4 - Introduction to Epidemiology," in *Clinical and Translational Science (Second Edition)* (D. Robertson and G. H. Williams, eds.), pp. 53–69, Academic Press, Jan. 2017.
- [188] A. Sánchez-Ferro, M. Matarazzo, P. Martínez-Martín, J. C. Martínez-Ávila, A. Gómez de la Cámara, L. Giancardo, T. Arroyo Gallego, P. Montero, V. Puertas-Martín, I. Obeso, I. Butterworth, C. S. Mendoza, M. J. Catalán, J. A. Molina, F. Bermejo-Pareja, J. C. Martínez-Castrillo, L. López-Manzanares, A. Alonso-Cánovas, J. Herreros Rodríguez, and M. Gray, "Minimal Clinically Important Difference for UPDRS-III in Daily Practice," *Movement Disorders Clinical Practice*, vol. 5, no. 4, pp. 448–450, 2018.
- [189] L. Giancardo, A. Sánchez-Ferro, T. Arroyo-Gallego, I. Butterworth, C. S. Mendoza, P. Montero, M. Matarazzo, J. A. Obeso, M. L. Gray, and R. S. J. Estépar, "Computer keyboard interaction as an indicator of early Parkinson's disease," *Scientific Reports*, vol. 6, p. 34468, Oct. 2016.
- [190] Y. González-Zamorano, J. Fernández-Carnero, F. J. Sánchez-Cuesta, A. Arroyo-Ferrer, A. Vourvopoulos, P. Figueiredo, J. I. Serrano, and J. P. Romero, "New Approaches Based on Non-Invasive Brain Stimulation and Mental Representation Techniques Targeting Pain in Parkinson's Disease Patients: Two Study Protocols for Two Randomized Controlled Trials," *Brain Sciences*, vol. 11, p. 65, Jan. 2021.
- [191] A. Arroyo-Ferrer, F. J. Sánchez-Cuesta, Y. González-Zamorano, M. D. Del Castillo, C. Sastre-Barrios, M. Ríos-Lago, and J. P. Romero, "Validation of Cognitive Rehabilitation as a Balance Rehabilitation Strategy in Patients with Parkinson's Disease: Study Protocol for a Randomized Controlled Trial," *Medicina (Kaunas, Lithuania)*, vol. 57, p. 314, Mar. 2021.
- [192] "actiCHamp series | Brain Products GmbH > Solutions." <https://www.brainproducts.com/solutions/actichamp/>. Accedido: 26-09-2022.
- [193] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, pp. 9–21, Mar. 2004.

- [194] S. Hu, Y. Lai, P. A. Valdes-Sosa, M. L. Bringas-Vega, and D. Yao, "How do reference montage and electrodes setup affect the measured scalp EEG potentials?," *Journal of Neural Engineering*, vol. 15, p. 026013, Jan. 2018.
- [195] T. Oliphant, *Guide to NumPy*. CreateSpace Independent Publishing Platform, 2015.
- [196] W. McKinney, "pandas: a foundational python library for data analysis and statistics," *Python for High Performance and Scientific Computing*, vol. 14, no. 9, pp. 1–9, 2011.
- [197] J. Hunter, "Matplotlib: A 2D graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.
- [198] A. Gramfort, M. Luessi, E. Larson, D. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. Hämäläinen, "MEG and eeg data analysis with mne-python," *Frontiers in Neuroscience*, vol. 7, no. 267, pp. 1–13, 2013.
- [199] A. Hagberg, D. Schult, and P. Swart, "Exploring Network Structure, Dynamics, and Function using NetworkX," in *Proceedings of the 7th Python in Science conference (SciPy 2008)* (G. Varoquaux, T. Vaught, and J. Millman, eds.), pp. 11–15, 2008.
- [200] P. Virtanen, R. Gommers, T. Oliphant, and et al., "SciPy 1.0: fundamental algorithms for scientific computing in python," *Nature Methods*, vol. 17, pp. 261–272, 2020.
- [201] M. Billinger, C. Brunner, and G. Müller-Putz, "SCoT: a Python toolbox for EEG source connectivity," *Frontiers in neuroinformatics*, vol. 8, no. 22, pp. 1–11, 2014.
- [202] "Keras: the Python deep learning API." <https://keras.io/>. Accedido: 26-09-2022.
- [203] "TensorFlow." <https://www.tensorflow.org/>. Accedido: 26-09-2022.
- [204] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: A system for large-scale machine learning," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pp. 265–283, 2016.
- [205] "scikit-learn: machine learning in Python - documentation." <https://scikit-learn.org/>. Accedido: 26-09-2022.



- [206] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011.
- [207] D. Hongyong, “PyMeta. Built on PythonMeta, Python module of Meta-analysis.” <https://www.pymeta.com/>. Accedido: 26-09-2022.
- [208] J. L. Lanciego, N. Luquin, and J. A. Obeso, “Functional Neuroanatomy of the Basal Ganglia,” *Cold Spring Harbor Perspectives in Medicine*, vol. 2, p. a009621, Dec. 2012.
- [209] M. P. van den Heuvel and H. E. Hulshoff Pol, “Exploring the brain network: A review on resting-state fMRI functional connectivity,” *European Neuropsychopharmacology*, vol. 20, pp. 519–534, Aug. 2010.
- [210] A. Khanna, A. Pascual-Leone, C. M. Michel, and F. Farzan, “Microstates in resting-state EEG: Current status and future directions,” *Neuroscience & Biobehavioral Reviews*, vol. 49, pp. 105–113, Feb. 2015.
- [211] J. I. Serrano, M. D. del Castillo, V. Cortés, N. Mendes, A. Arroyo, J. Andreo, E. Roncon, M. del Valle, J. Herreros, and J. P. Romero, “EEG Microstates Change in Response to Increase in Dopaminergic Stimulation in Typical Parkinson’s Disease Patients,” *Frontiers in Neuroscience*, vol. 12, 2018.
- [212] I. Winkler, S. Haufe, and M. Tangermann, “Automatic Classification of Artifactual ICA-Components for Artifact Removal in EEG Signals,” *Behavioral and Brain Functions*, vol. 7, p. 30, Aug. 2011.
- [213] F. Perrin, J. Pernier, O. Bertrand, and J. F. Echallier, “Spherical splines for scalp potential and current density mapping,” *Electroencephalography and Clinical Neurophysiology*, vol. 72, pp. 184–187, Feb. 1989.
- [214] Y. Liu, W. G. Coon, A. d. Pestere, P. Brunner, and G. Schalk, “The effects of spatial filtering and artifacts on electrocorticographic signals,” *Journal of Neural Engineering*, vol. 12, p. 056008, Aug. 2015.
- [215] P. Rossini, R. Di Iorio, M. Bentivoglio, G. Bertini, F. Ferreri, C. Gerloff, R. Ilmoniemi, F. Miraglia, M. Nitsche, F. Pestilli, M. Rosanova, Y. Shirota, C. Tesoriero, Y. Ugawa, F. Vecchio, U. Ziemann, and M. Hallett, “Methods for analysis of brain connectivity:

- An IFCN-sponsored review," *Clinical Neurophysiology*, vol. 130, no. 10, pp. 1833–1858, 2019.
- [216] R. Srinivasan, W. R. Winter, J. Ding, and P. L. Nunez, "EEG and MEG coherence: Measures of functional connectivity at distinct spatial scales of neocortical dynamics," *Journal of Neuroscience Methods*, vol. 166, pp. 41–52, Oct. 2007.
- [217] C. Stam, B. Jones, G. Nolte, M. Breakspear, and P. Scheltens, "Small-world networks and functional connectivity in Alzheimer's disease," *Cereb. Cortex*, vol. 17, pp. 92–99, 2007.
- [218] D. Bassett, E. Bullmore, B. Verchinski, V. Mattay, D. Weinberger, and A. Meyer-Lindenberg, "Hierarchical organization of human cortical networks in health and schizophrenia," *J. Neurosci.*, vol. 28, pp. 9239–9248, 2008.
- [219] S. Leistedt, N. Coumans, M. Dumont, J. Lanquart, C. Stam, and P. Linkowski, "Altered sleep brain functional connectivity in acutely depressed patients," *Hum. Brain Mapp.*, vol. 30, pp. 2207–2219, 2009.
- [220] C. Stam, W. de Haan, A. Daffertshofer, B. Jones, I. Manshanden, A. van Cappellen van Walsum, T. Montez, J. Verbunt, J. de Munck, B. van Dijk, H. Berendse, and P. Scheltens, "Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease," *Brain*, vol. 132, pp. 213–224, 2009.
- [221] J. S. George, J. Strunk, R. Mak-McCully, M. Houser, H. Poizner, and A. R. Aron, "Dopaminergic therapy in Parkinson's disease decreases cortical beta band coherence in the resting state and increases cortical beta band power during executive control," *NeuroImage: Clinical*, vol. 3, pp. 261–270, Jan. 2013.
- [222] N. C. Swann, C. de Hemptinne, A. R. Aron, J. L. Ostrem, R. T. Knight, and P. A. Starr, "Elevated synchrony in Parkinson disease detected with electroencephalography," *Annals of Neurology*, vol. 78, no. 5, pp. 742–750, 2015.
- [223] M. Bočková and I. Rektor, "Impairment of brain functions in Parkinson's disease reflected by alterations in neural connectivity in EEG studies: A viewpoint," *Clinical Neurophysiology*, vol. 130, pp. 239–247, Feb. 2019.
- [224] O. Faust and M. G. Bairy, "Nonlinear analysis of physiological signals: a review | Journal of Mechanics in Medicine and Biology," *Journal of Mechanics in Medicine and Biology*, vol. 12, no. 04, p. 1240015, 2012.

- [225] H. Helakari, J. Kananen, N. Huotari, L. Raitamaa, T. Tuovinen, V. Borchardt, A. Rasilta, V. Raatikainen, T. Starck, T. Hautaniemi, T. Myllylä, O. Tervonen, S. Rytty, T. Keinänen, V. Korhonen, V. Kiviniemi, and H. Ansakorpi, "Spectral entropy indicates electrophysiological and hemodynamic changes in drug-resistant epilepsy - A multimodal MREG study," *NeuroImage: Clinical*, vol. 22, p. 101763, Jan. 2019.
- [226] F. Fallani, L. Astolfi, F. Cincotti, D. Mattia, A. Tocci, M. Marciani, A. Colosimo, S. Salinari, S. Gao, A. Cichocki, and F. Babiloni, "Extracting information from cortical connectivity patterns estimated from high resolution EEG recordings: a theoretical graph approach," *Brain Topography*, vol. 19, no. 3, pp. 125–136, 2007.
- [227] M. Christodoulakis, A. Hadjipapas, E. Papathanasiou, M. Anastasiadou, S. Papacostas, and G. Mitsis, "Graph-theoretic analysis of scalp EEG brain networks in epilepsy-the influence of montage and volume conduction," in *13th IEEE International Conference on BioInformatics and BioEngineering*, pp. 1–4, IEEE, 2013.
- [228] E. Olejarczyk and W. Jernajczyk, "Graph-based analysis of brain connectivity in schizophrenia," *PLoS One*, vol. 12, no. 11, p. e0188629, 2017.
- [229] J. Gomez-Pilar, R. de Luis-García, A. Lubeiro, N. de Uribe, J. Poza, P. Núñez, M. Ayuso, R. Hornero, and V. Molina, "Deficits of entropy modulation in schizophrenia are predicted by functional connectivity strength in the theta band and structural clustering," *NeuroImage: Clinical*, vol. 18, pp. 382–389, 2018.
- [230] M. Anastasiadou, M. Christodoulakis, E. Papathanasiou, S. Papacostas, A. Hadjipapas, and G. Mitsis, "Graph theoretical characteristics of EEG-based functional brain networks in patients with epilepsy: The effect of reference choice and volume conduction," *Frontiers in Neuroscience*, vol. 13, no. 221, pp. 1–18, 2019.
- [231] Izsto, G. Ru, M. Crescio, F. Ingravalle, C. Maurella, Ubesp, D. Gregori, C. Lanera, D. Azzolina, G. Lorenzoni, N. Soriani, S. Zec, Dscb, P. Berchialla, S. Mercadante, Zeta, F. Zobec, M. Ghidina, S. Baldas, B. Bonifacio, A. Kinkopf, D. Kozina, L. Nicolandi, and L. Rosat, "Machine Learning Techniques applied in risk assessment related to food safety," *EFSA Supporting Publications*, vol. 14, no. 7, p. 1254E, 2017.
- [232] D. Moher, L. Shamseer, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, P. Shekelle, L. A. Stewart, and PRISMA-P Group, "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement," *Systematic Reviews*, vol. 4, p. 1, Jan. 2015.

- [233] L. Shamseer, D. Moher, M. Clarke, D. Gherzi, A. Liberati, M. Petticrew, P. Shekelle, and L. A. Stewart, "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation," *BMJ*, vol. 349, p. g7647, Jan. 2015.
- [234] A. M. Maitín, A. J. García-Tejedor, and J. P. R. Muñoz, "Machine Learning Approaches for Detecting Parkinson's Disease from EEG Analysis: A Systematic Review," *Applied Sciences*, vol. 10, p. 8662, Jan. 2020.
- [235] W. Luo, D. Phung, T. Tran, S. Gupta, S. Rana, C. Karmakar, A. Shilton, J. Yearwood, N. Dimitrova, T. B. Ho, S. Venkatesh, and M. Berk, "Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View," *Journal of Medical Internet Research*, vol. 18, p. e5870, Dec. 2016.
- [236] A. M. Maitin, J. P. Romero Muñoz, and A. J. García-Tejedor, "Survey of Machine Learning Techniques in the Analysis of EEG Signals for Parkinson's Disease: A Systematic Review," *Applied Sciences*, vol. 12, p. 6967, Jan. 2022.
- [237] A. Craik, Y. He<sup>1</sup>, and J. Contreras-Vidal, "Deep learning for electroencephalogram (eeg) classification tasks: a review," *Journal of Neural Engineering*, vol. 16, p. 031001, 2019.
- [238] C. Godsil and G. Royle, *Algebraic Graph Theory*. Springer, 2001.
- [239] L. Cabañero Gomez, R. Hervás, I. Gonzalez, and L. Rodríguez-Benitez, "eeglib: A Python module for eeg feature extraction," *SoftwareX*, vol. 15, p. 100745, 2021.
- [240] D. Grattarola and C. Alippi, "Graph Neural Networks in TensorFlow and Keras with Spektral," 2020.
- [241] D. Bahdanau, K. H. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in *3rd International Conference on Learning Representations, ICLR*, Jan. 2015.
- [242] I. I. Goncharova, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "EMG contamination of EEG: spectral and topographical characteristics," *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, vol. 114, pp. 1580–1593, Sept. 2003.
- [243] Y. Lee, J.-M. Kwon, Y. Lee, H. Park, H. Cho, and J. Park, "Deep Learning in the Medical Domain: Predicting Cardiac Arrest Using Deep Learning," *Acute and Critical Care*, vol. 33, pp. 117–120, Aug. 2018.

- [244] R. Kohavi and D. H. Wolpert, "Bias Plus Variance Decomposition for Zero-One Loss Functions," in *Machine Learning: Proceedings of the Thirteenth International*, pp. 275–283, Morgan Kaufmann Publishers, 1996.
- [245] P. Silberstein, A. Pogosyan, A. A. Kühn, G. Hotton, S. Tisch, A. Kupsch, P. Dowsey-Limousin, M. I. Hariz, and P. Brown, "Cortico-cortical coupling in Parkinson's disease and its modulation by therapy," *Brain*, vol. 128, pp. 1277–1291, June 2005.
- [246] D. Stoffers, J. L. W. Bosboom, J. B. Deijen, E. C. Wolters, H. W. Berendse, and C. J. Stam, "Slowing of oscillatory brain activity is a stable characteristic of Parkinson's disease without dementia," *Brain*, vol. 130, pp. 1847–1860, July 2007.
- [247] R. Noldus and P. Van Mieghem, "Assortativity in complex networks," *Journal of Complex Networks*, vol. 3, pp. 507–542, Dec. 2015.
- [248] A. M. Maitín, R. Perezzan, D. Herráez-Aguilar, J. I. Serrano, M. D. Del Castillo, A. Arroyo, J. Andreo, and J. P. Romero, "Time Series Analysis Applied to EEG Shows Increased Global Connectivity during Motor Activation Detected in PD Patients Compared to Controls," *Applied Sciences*, vol. 11, p. 15, Jan. 2021.
- [249] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, July 1948. Conference Name: The Bell System Technical Journal.
- [250] H. Viertiö-Oja, V. Maja, M. Särkelä, P. Talja, N. Tenkanen, H. Tolvanen-Laakso, M. Paloheimo, A. Vakkuri, A. Yli-Hankala, and P. Meriläinen, "Description of the Entropy algorithm as applied in the Datex-Ohmeda S/5 Entropy Module," *Acta Anaesthesiologica Scandinavica*, vol. 48, no. 2, pp. 154–161, 2004.
- [251] S. Kullback and R. A. Leibler, "On Information and Sufficiency," *The Annals of Mathematical Statistics*, vol. 22, pp. 79–86, Mar. 1951.
- [252] D. Stoffers, J. L. W. Bosboom, J. B. Deijen, E. C. Wolters, C. J. Stam, and H. W. Berendse, "Increased cortico-cortical functional connectivity in early-stage Parkinson's disease: An MEG study," *NeuroImage*, vol. 41, pp. 212–222, June 2008.
- [253] C. Monnot, X. Zhang, S. Nikkhou-Aski, P. Damberg, and P. Svenningsson, "Asymmetric dopaminergic degeneration and levodopa alter functional corticostriatal connectivity bilaterally in experimental parkinsonism," *Experimental Neurology*, vol. 292, pp. 11–20, June 2017.

- [254] R. Yuvaraj, U. Rajendra Acharya, and Y. Hagiwara, "A novel Parkinson's Disease Diagnosis Index using higher-order spectra features in EEG signals," *Neural Computing and Applications*, vol. 30, pp. 1225–1235, Aug. 2018.
- [255] N. Betrouni, A. Delval, L. Chaton, L. Defebvre, A. Duits, A. Moonen, A. F. Leentjens, and K. Dujardin, "Electroencephalography-based machine learning for cognitive profiling in Parkinson's disease: Preliminary results," *Movement Disorders*, vol. 34, no. 2, pp. 210–217, 2019.
- [256] M. Chaturvedi, F. Hatz, U. Gschwandtner, J. G. Bogaarts, A. Meyer, P. Fuhr, and V. Roth, "Quantitative EEG (QEEG) Measures Differentiate Parkinson's Disease (PD) Patients from Healthy Controls (HC)," *Frontiers in Aging Neuroscience*, vol. 9, Jan. 2017.
- [257] S. L. Oh, Y. Hagiwara, U. Raghavendra, R. Yuvaraj, N. Arunkumar, M. Murugappan, and U. R. Acharya, "A deep learning approach for Parkinson's disease diagnosis from EEG signals," *Neural Computing & Applications*, vol. 32, pp. 10927–10933, Aug. 2020.
- [258] G. Ruffini, D. Ibanez, M. Castellano, L. Dubreuil-Vall, A. Soria-Frisch, R. Postuma, J.-F. Gagnon, and J. Montplaisir, "Deep Learning With EEG Spectrograms in Rapid Eye Movement Behavior Disorder," *Frontiers in Neurology*, vol. 10, July 2019.
- [259] S. Vanneste, J.-J. Song, and D. De Ridder, "Thalamocortical dysrhythmia detected by machine learning.," *Nature communications*, vol. 9, p. 1103, Mar. 2018.
- [260] S. Waninger, C. Berka, M. S. Karic, S. Korszen, P. D. Mozley, C. Henchcliffe, Y. Kang, J. Hesterman, T. Mangoubi, and A. Verma, "Neurophysiological Biomarkers of Parkinson's Disease," *Journal of Parkinsons Disease*, vol. 10, no. 2, pp. 471–480, 2020.
- [261] A. Saikia, M. Hussain, A. Barua, and S. Paul, "EEG-EMG correlation for parkinson's disease," *International Journal of Engineering and Advanced Technology*, vol. 8, no. 6, pp. 1179–1185, 2019.
- [262] A. Saikia, M. Hussain, A. Barua, and S. Paul, "Performance analysis of various neural network functions for parkinson's disease classification using EEG and EMG," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 1, pp. 3402–3406, 2019.
- [263] R. Yuvaraj, M. Murugappan, N. M. Ibrahim, K. Sundaraj, M. I. Omar, K. Mohamad, and R. Palaniappan, "Optimal set of EEG features for emotional state classification

- and trajectory visualization in Parkinson's disease," *International Journal of Psychophysiology*, vol. 94, pp. 482–495, Dec. 2014.
- [264] R. Yuvaraj, M. Murugappan, N. M. Ibrahim, M. I. Omar, K. Sundaraj, K. Mohamad, R. Palaniappan, and M. Satiyan, "Emotion classification in Parkinson's disease by higher-order spectra and power spectrum features using EEG signals: A comparative study," *Journal of Integrative Neuroscience*, vol. 13, pp. 89–120, Mar. 2014.
- [265] R. Yuvaraj, M. Murugappan, N. M. Ibrahim, K. Sundaraj, M. I. Omar, K. Mohamad, and R. Palaniappan, "Detection of emotions in Parkinson's disease using higher order spectral features from brain's electrical activity," *Biomedical Signal Processing and Control*, vol. 14, pp. 108–116, Nov. 2014.
- [266] M. Murugappan, W. Alshuaib, A. Bourisly, S. Sruthi, and R. Ranjana, "Recurrence Quantification Analysis based Emotion Detection in Parkinson's disease using EEG Signals," in *4th International Conference on Computer, Communication and Signal Processing, ICCSP, 2020*.
- [267] M. Murugappan, W. Alshuaib, A. K. Bourisly, S. K. Khare, S. Sruthi, and V. Bajaj, "Tunable Q wavelet transform based emotion classification in Parkinson's disease using Electroencephalography," *PLOS ONE*, vol. 15, Nov. 2020.
- [268] M. Cesari, J. A. E. Christensen, M.-L. Muntean, B. Mollenhauer, F. Sixel-Doring, H. B. D. Sorensen, C. Trenkwalder, and P. Jennum, "A data-driven system to identify REM sleep behavior disorder and to predict its progression from the prodromal stage in Parkinson's disease," *Sleep Medicine*, vol. 77, pp. 238–248, Jan. 2021.
- [269] G. L. Sorensen, P. Jennum, J. Kempfner, M. Zoetmulder, and H. B. D. Sorensen, "A Computerized Algorithm for Arousal Detection in Healthy Adults and Patients With Parkinson Disease," *Journal of Clinical Neurophysiology*, vol. 29, pp. 58–64, Feb. 2012.
- [270] A. Patanaik, J. L. Ong, J. J. Gooley, S. Ancoli-Israel, and M. W. L. Chee, "An end-to-end framework for real-time automatic sleep stage classification," *Sleep*, vol. 41, May 2018.
- [271] G. L. Sorensen, J. Kempfner, P. Jennum, and H. B. D. Sorensen, "Detection of arousals in Parkinson's disease patients.," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, vol. 2011, pp. 2764–2767, 2011. Place: United States.

- [272] S. Castano-Candamil, T. Piroth, P. Reinacher, B. Sajonz, V. A. Coenen, and M. Tangermann, "Identifying controllable cortical neural markers with machine learning for adaptive deep brain stimulation in Parkinson's disease," *NeuroImage-Clinical*, vol. 28, 2020.
- [273] V. J. Geraedts, M. Koch, R. Kuiper, M. Kefalas, T. H. W. Back, J. J. van Hilten, H. Wang, H. A. M. Middelkoop, N. A. van der Gaag, M. F. Contarino, and M. R. Tannemaat, "Preoperative Electroencephalography-Based Machine Learning Predicts Cognitive Deterioration after Subthalamic Deep Brain Stimulation," *Movement Disorders*, vol. 36, pp. 2324–2334, Oct. 2021.
- [274] M. Stuart, C. S. Wickramasinghe, D. L. Marino, D. Kumbhare, K. Holloway, and M. Manic, "Machine Learning for Deep Brain Stimulation Efficacy using Dense Array EEG," in *12th International Conference on Human System Interaction (HSI)*, pp. 143–150, 2019.
- [275] D. Sand, D. Arkadir, M. Abu Snineh, O. Marmor, Z. Israel, H. Bergman, S. Hassin-Baer, S. Israeli-Korn, Z. Peremen, A. B. Geva, and R. Eitan, "Deep Brain Stimulation Can Differentiate Subregions of the Human Subthalamic Nucleus Area by EEG Biomarkers," *Frontiers in Systems Neuroscience*, vol. 15, Oct. 2021.
- [276] A. Maurer, S. Hanrahan, J. Nedrud, A. O. Hebb, and A. Papandreou-Suppappola, "Suppression of Neurostimulation Artifacts and Adaptive Clustering of Parkinson's Patients Behavioral Tasks using EEG," in *50th Asilomar Conference on Signals, Systems and Computers* (M. Matthews, ed.), pp. 851–855, 2016.
- [277] A. Saikia, V. Majhi, M. Hussain, A. Barua, S. Paul, and J. Verma, "Machine Learning based Diagnostic System for Early Detection of Parkinson's Disease," in *2020 International Conference on Computational Performance Evaluation, ComPE 2020*, pp. 275–279, 2020.
- [278] O. Geman, I. Chiuchisan, M. Covasa, K. Eftaxias, S. Sanei, J. G. Ferreira Madeira, and R. A. Mancebo Boloy, "Joint EEG - EMG Signal Processing for Identification of the Mental Tasks in Patients with Neurological Diseases," in *24th European Signal Processing Conference (EUSIPCO)*, pp. 1598–1602, 2016.
- [279] A. M. A. Handojoseno, G. R. Naik, M. Gilat, J. M. Shine, T. N. Nguyen, Q. T. Ly, S. J. G. Lewis, and H. T. Nguyen, "Prediction of Freezing of Gait in Patients with Parkinson's Disease Using EEG Signals.," *Studies in health technology and informatics*, vol. 246, pp. 124–131, 2018.



- [280] Q. Ly, A. Handojoseno, M. Gilat, N. Nguyen, R. Chai, Y. Tran, S. Lewis, and H. Nguyen, "Identifying montages that best detect the electroencephalogram power spectrum alteration during freezing of gait in Parkinson's disease patients," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2016-October, pp. 6094–6097, 2016.
- [281] Q. Ly, M. Gilat, R. Chai, K. Martens, M. Georgiades, G. Naik, Y. Tran, S. Lewis, and H. Nguyen, "Detection of turning freeze in Parkinson's disease based on S-transform decomposition of EEG signals," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 3044–3047, 2017.
- [282] Q. Ly, A. Handojoseno, M. Gilat, N. Nguyen, R. Chai, Y. Tran, S. Lewis, and H. Nguyen, "Detection of Gait Initiation Failure in Parkinson's disease patients using EEG signals," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2016-October, pp. 1599–1602, 2016.
- [283] A. M. A. Handojoseno, J. M. Shine, T. N. Nguyen, Y. Tran, S. J. G. Lewis, and H. T. Nguyen, "The detection of Freezing of Gait in Parkinson's disease patients using EEG signals based on Wavelet decomposition.," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, vol. 2012, pp. 69–72, 2012. Place: United States.
- [284] A. M. A. Handojoseno, J. M. Shine, T. N. Nguyen, Y. Tran, S. J. G. Lewis, and H. T. Nguyen, "Using EEG spatial correlation, cross frequency energy, and wavelet coefficients for the prediction of Freezing of Gait in Parkinson's Disease patients.," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, vol. 2013, pp. 4263–4266, 2013. Place: United States.
- [285] M. I. Vanegas, M. F. Ghilardi, S. P. Kelly, and A. Blangero, "Machine learning for eeg-based biomarkers in parkinson's disease," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 2661–2665, 2018.
- [286] B. Jervis, M. Saatchi, A. Lacey, T. Roberts, E. Allen, N. Hudson, S. Oke, and M. Grimsley, "Artificial neural network and spectrum analysis methods for detecting brain diseases from the cnv response in the electroencephalogram," *IEE Proceedings - Science, Measurement and Technology*, vol. 141, pp. 432–440, Nov. 1994.

- [287] M. Koch, V. Geraedts, H. Wang, M. Tannemaat, and T. Back, “Automated Machine Learning for EEG-Based Classification of Parkinson’s Disease Patients,” in *2019 IEEE International Conference on Big Data (Big Data)* (C. Baru, J. Huan, L. Khan, X. Hu, R. Ak, Y. Tian, R. Barga, C. Zaniolo, K. Lee, and Y. Ye, eds.), pp. 4845–4852, 2019.
- [288] Rodrigues, P. Miguel, Teixeira, and J. a. Paulo, “Classification of electroencephalogram signals using artificial neural networks,” in *2010 3rd International Conference on Biomedical Engineering and Informatics*, vol. 2, pp. 808–812, 2010.
- [289] G. Ruffini, D. Ibañez, M. Castellano, S. Dunne, and A. Soria-Frisch, “EEG-driven RNN Classification for Prognosis of Neurodegeneration in At-Risk Patients,” in *Artificial Neural Networks and Machine Learning – ICANN 2016* (A. E. Villa, P. Masulli, and A. J. Pons Rivero, eds.), pp. 306–313, Springer International Publishing, 2016.
- [290] K. Shreya Prabhu and R. Joy Martis, “Diagnosis of Parkinson’s Disease using Computer Aided Tool based on EEG,” in *2020 IEEE 17th India Council International Conference, INDICON, 2020*.
- [291] M. Rahman, A. Tutul, and A. Islam, “Solving the Maze of Diagnosing Parkinson’s Disease based on Portable EEG Sensing to be Adaptable to Go In-The-Wild,” in *ACM International Conference Proceeding Series*, pp. 65–73, 2020.
- [292] S. Lee, R. Hussein, R. Ward, Z. J. Wang, and M. J. McKeown, “A convolutional-recurrent neural network approach to resting-state EEG classification in Parkinson’s disease,” *Journal of Neuroscience Methods*, vol. 361, Sept. 2021.
- [293] X. Shi, T. Wang, L. Wang, H. Liu, and N. Yan, “Hybrid Convolutional Recurrent Neural Networks Outperform CNN and RNN in Task-state EEG Detection for Parkinson’s Disease,” in *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 939–944, 2019.
- [294] M. Shaban, “Automated screening of parkinson’s disease using deep learning based electroencephalography,” in *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 158–161, 2021.
- [295] T. Anwar, N. Rehmat, and H. Naveed, “A Generic Approach for Classification of Psychological Disorders Diagnosis using EEG.,” *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, vol. 2021, pp. 2025–2029, 2021.

- [296] S. Liu, M. Li, Y. Feng, M. Zhang, M. Acquah, S. Huang, J. Chen, and P. Ren, "Brain Network Analysis by Stable and Unstable EEG Components," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 4, pp. 1080–1092, 2021.
- [297] S. Kamalraj, K. Rejith, and G. Prasanna Venkatesan, "Frequency domain analysis for the classification of Parkinson's disease patients," in *IOP Conference Series: Materials Science and Engineering*, vol. 561, 2019. Issue: 1.
- [298] K. Rejith and K. Subramaniam, "Analysis of emotional states in Parkinson's disease using entropy, energy-entropy and teager energy-entropy features," *Indian Journal of Public Health Research and Development*, vol. 9, no. 10, pp. 1099–1102, 2018.
- [299] K. Rejith and K. Subramaniam, "Classification of emotional states in Parkinson's disease patients using machine learning algorithms," *Biomedical and Pharmacology Journal*, vol. 11, no. 1, pp. 333–341, 2018.
- [300] S. K. Khare, V. Bajaj, and U. R. Acharya, "PDCNNet: An Automatic Framework for the Detection of Parkinson's Disease Using EEG Signals," *IEEE Sensors Journal*, vol. 21, pp. 17017–17024, Aug. 2021.
- [301] H. W. Loh, C. P. Ooi, E. Palmer, P. D. Barua, S. Dogan, T. Tuncer, M. Baygin, and U. R. Acharya, "GaborPDNet: Gabor Transformation and Deep Neural Network for Parkinson's Disease Detection Using EEG Signals," *Electronics*, vol. 10, July 2021.
- [302] E. Emamzadeh-Hashemi, A. Mahdizadeh, M. Mirian, S. Lee, and M. McKeown, "Deep transfer learning for Parkinson's disease monitoring by image-based representation of resting-state EEG using directional connectivity," *Algorithms*, vol. 15, no. 1, 2022.
- [303] S. Lee, R. Hussein, and M. J. McKeown, "A Deep Convolutional-Recurrent Neural Network Architecture for Parkinson's Disease EEG Classification," in *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1–4, 2019.
- [304] "Lista completa de países con sus respectivos enchufes, voltajes (tensión eléctrica) y frecuencias." <https://www.worldstandards.eu/es/electricidad/enchufes-por-pais/>. Accedido: 18-10-2022.
- [305] A. Nogales, A. J. García-Tejedor, A. M. Maitín, A. Pérez-Morales, M. D. Del Castillo, and J. P. Romero, "Bert learns from electroencephalograms about parkinson's disease: Transformer-based models for aid diagnosis," *IEEE Access*, vol. 10, pp. 101672–101682, 2022.

- [306] B. Barz and J. Denzler, “Deep Learning on Small Datasets without Pre-Training using Cosine Loss,” in *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1360–1369, Mar. 2020. ISSN: 2642-9381.
- [307] A. Ezen-Can, “A Comparison of LSTM and BERT for Small Corpus,” Sept. 2020. arXiv:2009.05451 [cs].
- [308] F. Muslim, A. Purwarianti, and F. Z. Ruskanda, “Cost-Sensitive Learning and Ensemble BERT for Identifying and Categorizing Offensive Language in Social Media,” in *2021 8th International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA)*, pp. 1–6, Sept. 2021.
- [309] T. Ballarini, F. Růžička, O. Bezdicek, E. Růžička, J. Roth, A. Villringer, J. Vymazal, K. Mueller, M. L. Schroeter, and R. Jech, “Unraveling connectivity changes due to dopaminergic therapy in chronically treated Parkinson’s disease patients,” *Scientific Reports*, vol. 8, p. 14328, Sept. 2018.
- [310] J. L. W. Bosboom, D. Stoffers, C. J. Stam, B. W. van Dijk, J. Verbunt, H. W. Berendse, and E. C. Wolters, “Resting state oscillatory brain dynamics in Parkinson’s disease: An MEG study,” *Clinical Neurophysiology*, vol. 117, pp. 2521–2531, Nov. 2006.
- [311] R. W. Thatcher, P. J. Krause, and M. Hrybyk, “Cortico-cortical associations and EEG coherence: A two-compartmental model,” *Electroencephalography and Clinical Neurophysiology*, vol. 64, pp. 123–143, Aug. 1986.
- [312] P. Manganotti, C. Gerloff, C. Toro, H. Katsuta, N. Sadato, P. Zhuang, L. Leocani, and M. Hallett, “Task-related coherence and task-related spectral power changes during sequential finger movements,” *Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control*, vol. 109, pp. 50–62, Feb. 1998.
- [313] P. Klobošiaková, R. Mareček, J. Fousek, E. Výtvarová, and I. Rektorová, “Connectivity Between Brain Networks Dynamically Reflects Cognitive Status of Parkinson’s Disease: A Longitudinal Study,” *Journal of Alzheimer’s Disease*, vol. 67, pp. 971–984, Jan. 2019.
- [314] L.-L. Gao, J.-R. Zhang, P. Chan, and T. Wu, “Levodopa Effect on Basal Ganglia Motor Circuit in Parkinson’s Disease,” *CNS Neuroscience & Therapeutics*, vol. 23, no. 1, pp. 76–86, 2017.

- [315] M. Cassidy and P. Brown, "Task-related EEG-EEG coherence depends on dopaminergic activity in Parkinson's disease," *NeuroReport*, vol. 12, pp. 703–707, Mar. 2001.
- [316] J.-M. Melgari, G. Curcio, F. Mastrolilli, G. Salomone, L. Trotta, M. Tombini, L. di Biase, F. Scrascia, R. Fini, E. Fabrizio, P. M. Rossini, and F. Vernieri, "Alpha and beta EEG power reflects L-dopa acute administration in parkinsonian patients," *Frontiers in Aging Neuroscience*, vol. 6, 2014.
- [317] S. Evangelisti, F. Pittau, C. Testa, G. Rizzo, L. L. Gramegna, L. Ferri, A. Coito, P. Cortelli, G. Calandra-Buonaura, F. Bisquoli, C. Bianchini, D. N. Manners, L. Talozzi, C. Tonon, R. Lodi, and P. Tinuper, "L-Dopa Modulation of Brain Connectivity in Parkinson's Disease Patients: A Pilot EEG-fMRI Study," *Frontiers in Neuroscience*, vol. 13, 2019.
- [318] S. N. Haber, "The place of dopamine in the cortico-basal ganglia circuit," *Neuroscience*, vol. 282, pp. 248–257, Dec. 2014.
- [319] B. D. Berman, J. Smucny, K. P. Wylie, E. Shelton, E. Kronberg, M. Leehey, and J. R. Tregellas, "Levodopa modulates small-world architecture of functional brain networks in Parkinson's disease," *Movement Disorders*, vol. 31, no. 11, pp. 1676–1684, 2016.
- [320] F. G. Hillary, C. A. Roman, U. Venkatesan, S. M. Rajtmajer, R. Bajo, and N. D. Castellanos, "Hyperconnectivity is a fundamental response to neurological disruption," *Neuropsychology*, vol. 29, pp. 59–75, Jan. 2015.
- [321] S. Waninger, C. Berka, M. Stevanovic Karic, S. Korszen, P. D. Mozley, C. Henchcliffe, Y. Kang, J. Hesterman, T. Mangoubi, and A. Verma, "Neurophysiological biomarkers of parkinson's disease," *Journal of Parkinson's Disease*, vol. 10, pp. 471–480, 2020.
- [322] M. Chaturvedi, J. G. Bogaarts, V. V. Kozak (Cozac), F. Hatz, U. Gschwandtner, A. Meyer, P. Fuhr, and V. Roth, "Phase lag index and spectral power as QEEG features for identification of patients with mild cognitive impairment in Parkinson's disease," *Clinical Neurophysiology*, vol. 130, pp. 1937–1944, Oct. 2019.
- [323] V. J. Geraedts, M. Koch, M. F. Contarino, H. A. M. Middelkoop, H. Wang, J. J. van Hilten, T. H. W. Back, and M. R. Tannemaat, "Machine learning for automated EEG-based biomarkers of cognitive impairment during Deep Brain Stimulation screening in patients with Parkinson's Disease," *Clinical Neurophysiology*, vol. 132, pp. 1041–1048, May 2021.

- [324] S. Hassin-Baer, O. S. Cohen, S. Israeli-Korn, G. Yahalom, S. Benizri, D. Sand, G. Is-sachar, A. B. Geva, R. Shani-HersHKovich, and Z. Peremen, "Identification of an early-stage Parkinson's disease neuromarker using event-related potentials, brain network analytics and machine-learning," *PloS one*, vol. 17, no. 1, pp. e0261947–e0261947, 2022.
- [325] J. Zhang, Y. Gao, X. He, S. Feng, J. Hu, Q. Zhang, J. Zhao, Z. Huang, L. Wang, G. Ma, Y. Zhang, K. Nie, and L. Wang, "Identifying Parkinson's disease with mild cognitive impairment by using combined MR imaging and electroencephalogram," *European Radiology*, vol. 31, pp. 7386–7394, Oct. 2021.
- [326] C. Chu, Z. Zhang, J. Wang, S. Liu, F. Wang, Y. Sun, X. Han, Z. Li, X. Zhu, and C. Liu, "Deep learning reveals personalized spatial spectral abnormalities of high delta and low alpha bands in EEG of patients with early Parkinson's disease," *Journal of Neural Engineering*, vol. 18, Dec. 2021.
- [327] G. Guo, S. Wang, S. Wang, Z. Zhou, G. Pei, and T. Yan, "Diagnosing Parkinson's Disease Using Multimodal Physiological Signals," in *Human Brain and Artificial Intelligence. HBAI 2021. Communications in Computer and Information Science*, vol 1369. (Y. Wang, ed.), (Singapore), pp. 125–136, Springer Singapore, 2021.
- [328] S. K. Khare, V. Bajaj, and U. R. Acharya, "Detection of Parkinson's disease using automated tunable Q wavelet transform technique with EEG signals," *Biocybernetics and Biomedical Engineering*, vol. 41, pp. 679–689, June 2021.
- [329] A. de Oliveira, M. de Santana, M. Andrade, J. Gomes, M. Rodrigues, and W. dos Santos, "Early diagnosis of Parkinson's disease using EEG, machine learning and partial directed coherence," *Research on Biomedical Engineering*, vol. 36, no. 3, pp. 311–331, 2020.
- [330] S. A. A. Shah, L. Zhang, and A. Bais, "Dynamical system based compact deep hybrid network for classification of Parkinson disease related EEG signals," *Neural Networks*, vol. 130, pp. 75–84, Oct. 2020.
- [331] V. Jurcak, D. Tsuzuki, and I. Dan, "10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems," *NeuroImage*, vol. 34, no. 4, pp. 1600–1611, 2007.