

Instituto Tecnológico y de Estudios Superiores de Occidente

Reconocimiento de validez oficial de estudios de nivel superior según acuerdo secretarial 15018, publicado en el Diario Oficial de la Federación del 29 de noviembre de 1976.

Departamento de Matemáticas y Física
Maestría en Ciencia de Datos



Caracterización del funcionamiento cerebral en hombres y mujeres en estado de vigilia a partir de un registro electroencefalográfico (EEG).

**TRABAJO RECEPCIONAL que para obtener el GRADO de
Maestro en Ciencia de Datos**

Presenta:

Miguel Alfonso Guevara Hernández

Director:

Maestro Byron Michael Motta Bonilla

Tlaquepaque, Jalisco, 9 de diciembre de 2025

Caracterización del funcionamiento cerebral en hombres y mujeres en estado de vigilia a partir de un registro electroencefalográfico (EEG).

Miguel Alfonso Guevara Hernández

Resumen

Se ha demostrado que tanto hombres como mujeres presentan patrones de funcionamiento cerebral típicos de su género y que sus cerebros usan diferentes estrategias dependiendo de las situaciones y contextos. Muchas de estas diferencias funcionales del cerebro se han reportado en relación con la ejecución de tareas o procesos cognitivo-emocionales, sin embargo, no se ha determinado si esta diferenciación funcional se puede presentar en un estado fisiológico de reposo, en vigilia con ojos abiertos. El registro de la actividad eléctrica cerebral o electroencefalograma (EEG), es una técnica no invasiva que permite caracterizar el funcionamiento de diferentes áreas de la corteza cerebral tanto en estado de reposo como durante conductas y procesos cognitivos o emocionales específicos. El EEG consiste en oscilaciones de voltaje originadas por corrientes iónicas intra y extraneuronales en una gran población de neuronas que se activan sincrónicamente, con variaciones de voltaje en rango de amplitud desde $-100 \mu\text{V}$ hasta $+100 \mu\text{V}$ y en frecuencias por debajo de 100 ondas por segundo o Hz (casi siempre entre 0.5 y 50 Hz). Los parámetros a medir del EEG son la potencia absoluta (PA), la potencia relativa (PR) y el grado de correlación EEG (rEEG) entre áreas corticales, de ahí que la cantidad de datos EEG a analizar es excesivamente grande y, por ende, complicada y difícil de analizar. La aplicación de modelos de optimización para el manejo y análisis de datos EEG implica una gran ventaja para los procesos rápidos de diagnóstico clínico y obtención de resultados en la investigación básica, por lo que el objetivo principal de este trabajo recepcional fue aplicar diferentes modelos de optimización para determinar el funcionamiento cerebral de jóvenes participantes en base a la categorización de su actividad EEG registrada en estado de vigilia con ojos abiertos. Los objetivos específicos incluyeron el análisis de la PA, la PR y la correlación (rEEG) de las frecuencias rápidas en el rango de beta (de 13 a 30 Hz), aplicando para cada tipo de parámetro EEG tres modelos de optimización: 1) regresión cuadrática por mínimos cuadrados, 2) regresión logística y 3) clasificación por K-medias. Los análisis exploratorios mostraron que de los tres parámetros EEG, fue con la PR que se encontró una mejor distribución normal y con el primer método (regresión por mínimos cuadrados) los resultados más estables y acordes a las hipótesis neurofisiológicas. Se concluye que la aplicación del modelo de optimización, regresión por mínimos cuadrados, es el más adecuado para caracterizar, con datos de PR de las frecuencias rápidas, la funcionalidad EEG que presentan hombres y mujeres en condición de reposo ojos abiertos. De los 37 participantes masculinos registrados, 67.6 por ciento fueron clasificados con un tipo de funcionalidad cerebral dominante tipo A y un 32.4 por ciento con tipo B; asimismo de los 37 sujetos femeninos, 70.3 por ciento se clasificaron con funcionalidad cerebral dominante tipo B y un 29.7 por ciento presentaron una funcionalidad cerebral dominante tipo A, indicando que el sexo no determina la funcionalidad cerebral en un estado de reposo con ojos abiertos.

Tabla de Contenidos

	Página
1 Introducción	13
1.1. Contexto	13
1.2. Planteamiento del problema	14
1.3. Justificación	16
1.4. Objetivos	16
1.4.1. Objetivo general	16
1.4.2. Objetivos específicos	16
2 Metodología	19
2.1. Descripción de los datos	19
2.2. Análisis exploratorio	22
3 Descripción de modelos	27
3.1. Descripción de los modelos a utilizar	27
3.1.1. Regresión cuadrática por Mínimos Cuadrados	27
3.1.2. Regresión Logística	29
3.1.3. Clasificación por K-medias	30
3.2. Descripción de las métricas	30
3.3. Explicación del código paso a paso	31
3.4. Experimentos o simulaciones realizadas	33
4 Resultados y discusión.	35
4.1. Resultados de las pruebas realizadas usando todas las variables de PR	35
4.1.1. Primera prueba – Regresión por Mínimos Cuadrados usando datos de PR	35
4.1.2. Segunda prueba – Regresión Logística usando datos de Potencia Relativa	36
4.1.3. Tercera prueba – Clasificación de datos de Potencia Relativa por K-Medias	37
4.2. Resultados de las pruebas realizadas usando sólo las variables con poca o nula correlación de PR	38
4.2.1. Cuarta prueba – Regresión por Mínimos Cuadrados usando datos de Potencia Relativa seleccionando datos SIN correlación	40

4.2.2.	Quinta prueba – Regresión Logística usando datos de Potencia Relativa seleccionando datos SIN correlación	40
4.2.3.	Sexta prueba – Clasificación de datos de Potencia Relativa por K-Medias seleccionando datos SIN correlación	41
4.3.	Resultados de las pruebas realizadas usando todas las variables de correlación EEG	42
4.3.1.	Séptima prueba – Regresión por Mínimos Cuadrados usando datos de correlación EEG	42
4.3.2.	Octava prueba – Regresión Logística usando datos de Correlación EEG	43
4.3.3.	Novena prueba – Clasificación de datos de Correlación EEG por K-Medias	44
4.4.	Selección del mejor modelo	44
4.5.	Supuestos de la Regresión	45
4.5.1.	Primer supuesto - Linealidad	46
4.5.2.	Segundo supuesto - Normalidad	46
4.5.3.	Tercer supuesto - No Multicolinealidad	47
4.5.4.	Cuarto supuesto - Independencia	49
4.5.5.	Quinto supuesto - Homocedasticidad	49
5	Conclusiones y trabajo futuro.	51
5.1.	Conclusiones	51
5.2.	Trabajo futuro	52

Índice de figuras

	Página
2.1. Vista superior (A) y lateral (B) del cerebro humano donde se indican las áreas corticales y derivaciones (o sitios de registro) donde se colocaron los electrodos a nivel del cuero cabelludo para obtener el registro de la actividad EEG de los participantes.	20
2.2. Fotografía del cuarto de registro y ubicación del equipo que se utilizó para el registro y captura de señales EEG, en el laboratorio de neurofisiología de la Conducta Reproductiva del instituto de Neurociencias de la Universidad de Guadalajara)	21
2.3. Vista lateral del cerebro humano en donde se representan en colores las áreas corticales donde se colocaron los electrodos para obtener el registro EEG.	22
2.4. Espectros de PA de la actividad EEG registrada en las distintas áreas o derivaciones corticales.	22
2.5. Espectros de PR de la actividad EEG en rango de beta registrada en las distintas áreas o derivaciones corticales.	23
2.6. Distribución de los valores de correlación entre las señales EEG de la banda beta de áreas corticales correspondientes al hemisferio izquierdo y derecho.	24
4.1. Pairplot de distribución de las variables contra otras variables de potencia relativa.	38
4.2. Pairplot de distribución de las variables contra otras variables de potencia relativa.	39
4.3. Histograma con distribución normalizada de los residuos del modelo seleccionado.	47
4.4. Mapa de calor de la correlación de las variables del modelo seleccionado.	48
4.5. Resultados del VIF para cada una de las variables del dataset.	48
4.6. Resultados de la prueba Durbin-Watson para comprobar independencia en los residuos.	49

4.7. Varianza constante de los residuos para comprobar la homocedasticidad de los mismos.	50
--	----

Índice de tablas

	Página
2.1. Valores de PA (microvolts) de la actividad EEG en el rango de beta registrada en las distintas áreas corticales o derivaciones de los 30 sujetos durante la condición basal ojos abiertos.	25
2.2. Datos EEG en PR (porcentaje de energía) de la banda beta calculada en cada área cortical.	26
2.3. Valores de correlación entre las señales EEG en el rango de beta de las áreas corticales izquierda y derecha del cerebro correspondientes a los 30 sujetos que realizaron la prueba.	26
4.1. Resultados de las pruebas de Regresión por Mínimos Cuadrados utilizando los datos de PR de las señales EEG de la banda beta.	35
4.2. Resultados de las pruebas de Regresión Logística utilizando los datos de PR.	36
4.3. Resultados de Regresión por Mínimos Cuadrados con Potencia Relativa usando datos SIN correlación.	40
4.4. Resultados de Regresión Logística usando con Potencia Relativa seleccionando datos SIN correlación.	41
4.5. Resultados de la Regresión por Mínimos Cuadrados usando datos de correlación EEG.	42
4.6. Resultados de Regresión Logística usando datos de correlación EEG.	43
4.7. Tabla con las 9 variables de la PR y la función cerebral obtenida para cada sujeto por medio de la regresión. . .	46

*Dedicado a mis padres, sin quienes este
trabajo, jamás se habría hecho.*

1 Introducción

1.1 Contexto

Durante el día a día, todas las personas debemos ejecutar una serie de actividades y resolver conflictos o problemas por medio de distintas estrategias. Pueden existir varios caminos para llegar a una misma solución, pero el ¿Cómo llegamos a esta solución? puede variar mucho según la persona.

¿Qué es lo que define cómo actuamos, realizamos actividades y experimentamos distintas condiciones fisiológicas y emociones? La respuesta se encuentra en la funcionalidad de nuestro cerebro, la cual, para términos de este trabajo, y en base a numerosos estudios que lo fundamentan [1, 2, 3, 4] puede ser clasificada como tipo A o tipo B.

La funcionalidad cerebral, es decir, la activación y manera en que se conectan funcionalmente las áreas de nuestro cerebro, son las que determinan o establecen las distintas estrategias cerebrales (de habilidades, cognitivas o motivo/emocionales) que presentan tanto animales como humanos. En este trabajo, consideraremos la funcionalidad tipo A como aquella que prevalece en el hemisferio izquierdo del cerebro y que se asocia con funciones cognitivas como la lógica matemática e inteligencia espacial, mientras que la funcionalidad tipo B es la que prevalece en el hemisferio derecho del cerebro, y se asocia con la inteligencia artística, las manualidades, las habilidades lingüísticas, las características físicas y estados emocionales [1, 2, 3, 4].

Varios estudios a lo largo de la historia han mostrado que la funcionalidad tipo A es más dominante en sujetos del sexo masculino, mientras que la funcionalidad tipo B es más dominante en sujetos del sexo femenino [5, 6, 7] Razón por la cual, cultural y socialmente hablando, se ha generado el concepto general de que los hombres son más eficientes para la lógica matemática, menos emocionales y las mujeres son más eficientes para las artes, los aspectos físicos de personas y objetos, la orientación espacial (el entorno), y son más emocionales [6, 7, 8].

Este trabajo de obtención de grado, pretende demostrar que el sexo NO determina la funcionalidad cerebral de una persona, mostrando

matemáticamente que existe una cierta proporción de hombres con funcionalidad cerebral dominante tipo B, al igual que una cierta proporción de mujeres con funcionalidad cerebral dominante tipo A.

1.2 *Planteamiento del problema*

¿Cómo se evalúa la funcionalidad cerebral de una persona? Esta se analiza en base a los resultados de distintos estudios tanto de imagenología como electroencefalográficos (EEG) durante la ejecución de distintas tareas o estados fisiológicos. En el caso de la actividad EEG, esta se registra mediante electrodos de superficie que se colocan en distintas áreas (o derivaciones) sobre el cuero cabelludo, de los dos hemisferios del cerebro (izquierdo y derecho) y se conectan a un polígrafo donde se registran, amplifican y filtran las señales EEG para almacenarlas en archivos específicos en computadora. Estas señales EEG, son posteriormente analizadas fuera de línea mediante la Transformada Rápida de Fourier para obtener distintos parámetros, tales como la potencia absoluta (en microvolts), la potencia relativa (en porcentaje) y el grado de acoplamiento EEG entre dos áreas cerebrales, denominado correlación EEG [9, 10, 11].

Así, se ha demostrado por ejemplo, que durante el estado de reposo con ojos cerrados predomina una actividad sincronizada en el rango de la banda alfa (de 4-12 Hz μ ondas por segundo) mientras que con ojos abiertos lo que predomina es una actividad más rápida y de bajo voltaje indicadora de un estado de atención o alerta (de 14 a 30 Hz) [12, 13, 14, 15]. Asimismo, en la ejecución de tareas que demandan procesos cognitivos de alto orden, la actividad EEG se caracteriza por frecuencias que van desde beta hasta más rápidas (en el rango de 25 a 50 Hz) [16, 17]. Distintos patrones EEG pueden ser observados y categorizados de acuerdo a las condiciones fisiológicas de los individuos y se ha mostrado que presentan importantes diferencias sexuales [5, 6, 7], lo que ha llevado a pensar que hombres y mujeres poseen un funcionamiento cerebral típico y característico [5, 18, 19].

Sin embargo, es importante notar que el funcionamiento cerebral es muy variable y depende de una multitud de factores, tales como el estado fisiológico (sueño, vigilia, emocionalidad, patologías, etc), la edad, la hora del día, el grado de atención, y el tipo de tarea que el participante esté realizando. En el caso de las tareas, se suma el efecto de la actividad motora (movimientos de las manos, los ojos, etc.) y de las condiciones ambientales y experimentales, de manera que en muchos casos, los resultados de actividad EEG en hombres y mujeres son controversiales.

Se ha encontrado que, en la ejecución de tareas, ya sea manuales o cognitivas, no sólo prevalece la funcionalidad de un hemisferio u

otro en cada sexo, sino que en ambos, hombres y mujeres, hay un patrón de actividad EEG que puede implicar ambos hemisferios, e inclusive conformarse redes o circuitos funcionales entre áreas corticales bilaterales (intrahemisféricas o interhemisféricas) [20, 18, 21, 22, 23, 24, 25].

Así, se ha reportado que dependiendo de la actividad o problema a resolver se activan los dos hemisferios del cerebro, pero: para una prueba lógico matemática predominará el funcionamiento del hemisferio izquierdo, mientras que para una prueba artística, manual o emocional, predominará el funcionamiento del hemisferio derecho, sin importar el sexo de la persona [1, 2, 3, 4].

Ahora bien, sabemos que en la ejecución de tareas, que implican procesos cognitivos de alto orden, como la planeación, la toma de decisiones, los problemas matemáticos, etc., participan distintas áreas subcorticales y corticales del cerebro y requieren de una serie de procesos muy complicados y por ende, los individuos pueden recurrir a distintas estrategias para resolverlos, de ahí que el funcionamiento cerebral será muy diferente en las distintas personas. Lo mismo ocurre con la actividad EEG registrada en situaciones de valoración emocional y motivacional, así como de destrezas motoras y categorización de metas, dando resultados muy variables entre hombres y mujeres.

Dado lo anterior, es que en todos los estudios clínicos y de investigación, resulta importante determinar el funcionamiento cerebral de los participantes en un estado basal de vigilia en reposo, situación en la cual los participantes no reciben información sensorial relevante, no experimentan emoción alguna, ni requieren de realizar algún proceso cognitivo, ejecución motora o postural y por tanto, permanecen sin ninguna perturbación ambiental [26]. Esta condición se considera como un estado basal que se toma como punto de referencia para comparar con el funcionamiento cerebral que ocurre en otras situaciones fisiológicas, como patologías, estados de ánimo, o bien, ejecución de tareas o procesos cognitivos [27, 28, 29, 30, 31, 32].

Entonces, ¿cómo se determina el funcionamiento cerebral en una persona? La respuesta a esta cuestión es difícil de encontrar, sobre todo si consideramos la inmensa cantidad de variables que lo pueden afectar, por lo que un primer acercamiento a esta respuesta (la cual se propone en este trabajo de obtención de grado), sería realizar una clasificación en base a la actividad electroencefalográfica de varios sujetos de ambos sexos, en un estado basal de vigilia con ojos abiertos. El estado basal de ojos abiertos fue considerado en este estudio, considerando que se ha mostrado que esta condición de reposo es la adecuada a comparar respecto a estados de ejecución de tareas, estados emocionales e incluso patologías como la depresión, ya que el sujeto está despierto, generalmente observando un punto fijo, tranquilo y sin esfuerzo mental

alguno [33, 34, 23, 31, 35].

1.3 *Justificación*

De comprobarse que se puede determinar la funcionalidad cerebral de una persona en base a su actividad EEG en un estado basal de vigilia con ojos abiertos, esto podría, en primer lugar, ampliar el conocimiento en la neurociencia básica y clínica, lo que abriría la oportunidad de realizar otros estudios en varios campos de la ciencia, tales como la medicina, la psicología, las ciencias sociales, la investigación e incluso en el ámbito social y laboral.

Además, desmentiría la concepción general de que el sexo de una persona, es lo que determina si será mejor que el sexo opuesto en determinadas actividades.

Algunos ejemplos de aplicación de estos resultados, podrían ser:

- Determinar las estrategias cerebrales que permitirán a un individuo desarrollar cierta habilidad.
- Determinar si el funcionamiento cerebral de un individuo explica sus habilidades cognitivo-emocionales.
- Determinar si el funcionamiento cerebral de un individuo predice o explica el riesgo de padecer alguna patología, tales como la epilepsia, un problema de sueño, la depresión, etc.
- Determinar si en base al funcionamiento cerebral de un individuo se explican sus rasgos de personalidad.

1.4 *Objetivos*

1.4.1 *Objetivo general*

Caracterizar la funcionalidad cerebral de un conjunto de personas, con base en su actividad EEG en estado de vigilia con ojos abiertos.

1.4.2 *Objetivos específicos*

1. Desarrollar un modelo matemático, con base en la actividad EEG, que pueda determinar la funcionalidad cerebral característica de un grupo de personas.
2. Corroborar que los resultados de la caracterización de la funcionalidad cerebral de dichas personas son verídicos bajo el modelo de regresión propuesto, utilizando otros modelos de regresión distintos al propuesto.

3. Corroborar que los resultados de la caracterización de la funcionalidad cerebral de dichas personas son verídicos bajo el modelo de regresión propuesto, comprobando los cinco supuestos de la regresión lineal.
4. Calcular el porcentaje de ocurrencia de los tipos de funcionalidad cerebral (tipo A y tipo B) en un grupo de personas.

2 Metodología

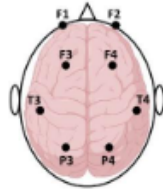
2.1 Descripción de los datos

Los datos de EEG utilizados en este proyecto fueron obtenidos de los experimentos realizados en el laboratorio de Neurofisiología de la Conducta Reproductiva, del Instituto de Neurociencias de la Universidad de Guadalajara, del cual es responsable la Dra. Marisela Hernández González, profesora-investigadora Titular C, Miembro del SNII nivel III. Por tanto, los datos son propiedad de dicho laboratorio.

Los registros EEG de 74 estudiantes voluntarios de entre 20 y 30 años de edad de la Universidad de Guadalajara, fueron distribuidos en dos grupos, uno con 37 estudiantes del sexo masculino y otro con 37 estudiantes del sexo femenino.

El día del experimento los participantes acudieron al laboratorio donde permanecieron cómodamente sentados frente a un monitor, y después de firmar la carta de consentimiento voluntario, se les colocaron electrodos sobre el cuero cabelludo de las áreas corticales: Frontopolar (F₁ y F₂), Dorsolateral (F₃ y F₄), Temporal (T₃ y T₄) y Parietal (P₃ y P₄), tanto del hemisferio izquierdo como del derecho. Así mismo, se colocaron electrodos en orejas cortocircuitadas como referencia, y uno en la frente para hacer tierra. Luego de colocar un electrodo en cada una de las 8 zonas corticales o derivaciones (4 en lado izquierdo y 4 en lado derecho), tal como se observa en la sección 2.1, se pidió a los participantes que se relajaran y cerraran los ojos durante 3 minutos para registrar el EEG durante la condición Ojos Cerrados. Posterior a ello se les solicitó abrieran los ojos y observaran un punto fijo de color blanco en la pantalla de computadora para registrar el EEG en la condición Ojos Abiertos (3 minutos).

A



B



Figura 2.1: Vista superior (A) y lateral (B) del cerebro humano donde se indican las áreas corticales y derivaciones (o sitios de registro) donde se colocaron los electrodos a nivel del cuero cabelludo para obtener el registro de la actividad EEG de los participantes.

Para el registro EEG se utilizó un polígrafo NatusOne VEEG de la marca Nicolet; que permitió amplificar y filtrar la señal, la cual fue capturada y almacenada en archivos específicos de una computadora para su posterior análisis (section 2.1). La frecuencia de muestreo fue de 512 Hz y la banda de análisis incluyó de 1 a 50 Hz. Posteriormente, y ya fuera de línea, las señales EEG fueron revisadas para eliminar cualquier artefacto por movimiento ocular o ruido, y se analizaron mediante la Transformada Rápida de Fourier para calcular la cantidad de energía o Potencia Absoluta (PA) en microvolts, y el porcentaje de tal energía o Potencia Relativa (PR).

Asimismo, mediante el Coeficiente de Correlación de Pearson, se calculó el grado de acoplamiento entre las señales EEG de dos áreas cerebrales, también denominado grado de correlación (r_{EEG}); todo utilizando el programa EEGmagic [36, 10].

Para este trabajo en particular se decidió analizar los parámetros de PA, PR y r_{EEG} solo de la banda EEG en el rango de beta (de 13.5 a 30 Hz), ya que estas frecuencias rápidas son las que predominan en estado de vigilia con ojos abiertos, así como en relación con procesos de alto orden como son las tareas cognitivas [12, 13, 15, 16].

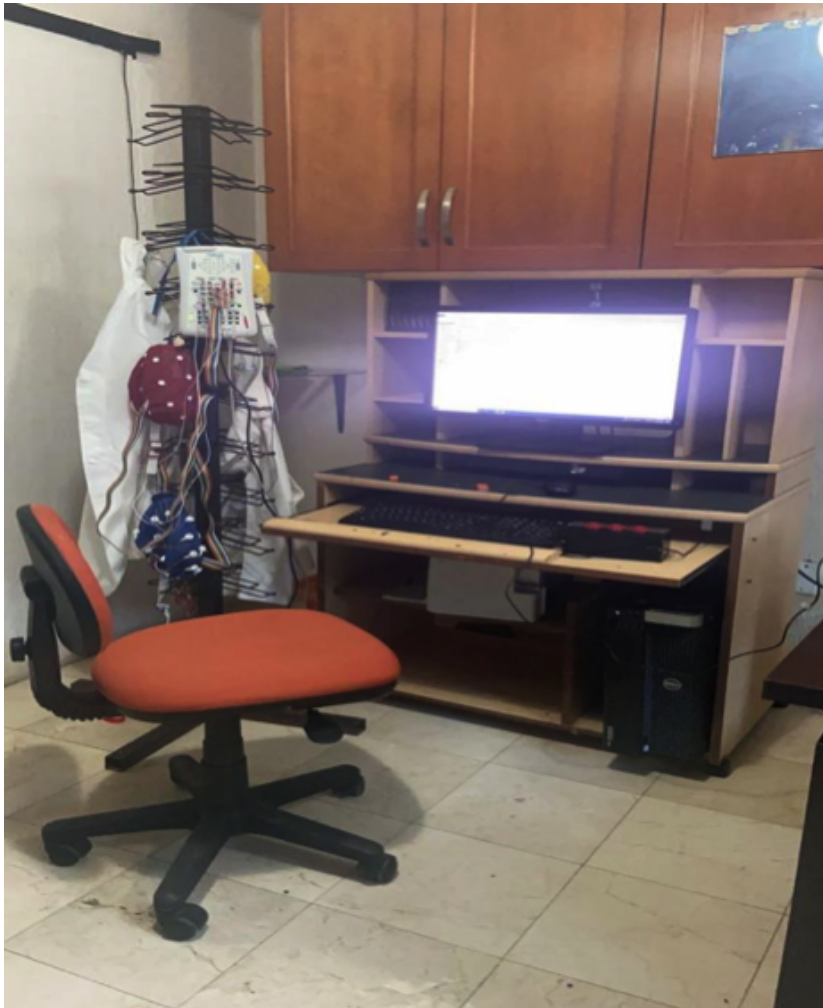


Figura 2.2: Fotografía del cuarto de registro y ubicación del equipo que se utilizó para el registro y captura de señales EEG, en el laboratorio de neurofisiología de la Conducta Reproductiva del instituto de Neurociencias de la Universidad de Guadalajara)

2.2 Análisis exploratorio

Para las primeras pruebas, se trabajó con los registros EEG de 15 hombres y 15 mujeres, dando así un total de 30 sujetos distribuidos equitativamente. En esta primera prueba se decidió trabajar con los datos de PA (en microvolts) los cuales se presentan en la section 2.2.

Las primeras ocho columnas de la section 2.2, corresponden a las zonas del cerebro las cuales, se pueden ubicar por su color de acuerdo a la section 2.2.

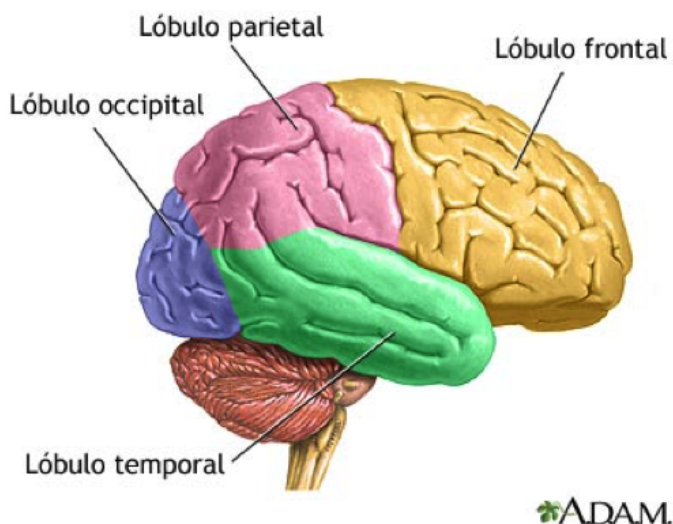


Figura 2.3: Vista lateral del cerebro humano en donde se representan en colores las áreas corticales donde se colocaron los electrodos para obtener el registro EEG.

Sin embargo, tras aplicarle un análisis exploratorio a los datos, se pudo observar que los datos dados en PA del rango de beta, no tenían una distribución normal, tal como se puede observar en la section 2.2.

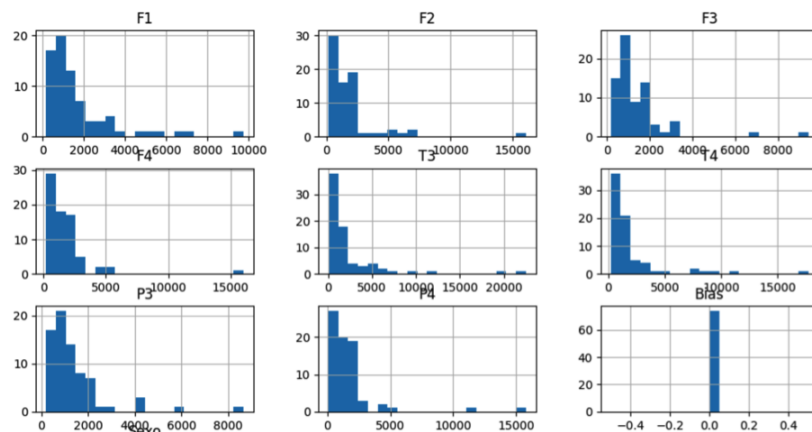


Figura 2.4: Espectros de PA de la actividad EEG registrada en las distintas áreas o derivaciones corticales.

Dado que estos datos así como están con PA, no fueron útiles para realizar una buena regresión, se decidió utilizar los datos de PR, calculando el porcentaje de energía respecto a la PA en valores del 0 al 100, tal como se puede observar en la section 2.2.

Esta tabla es muy similar a la anterior, sólo que ahora en lugar de tener los datos dados en su potencia absoluta, ahora están dados en su potencia relativa.

Esta tabla es muy similar a la anterior, sólo que ahora en lugar de tener los datos en PA, ahora están dados en PR.

Tras aplicar el análisis exploratorio, se encontró que a diferencia de los datos en PA, los datos en PR sí presentaron una distribución que se aproxima a la normal y por tanto, se decidió utilizar dichos datos para caracterización de funcionalidad cerebral por medio del EEG de frecuencias rápidas (section 2.2).

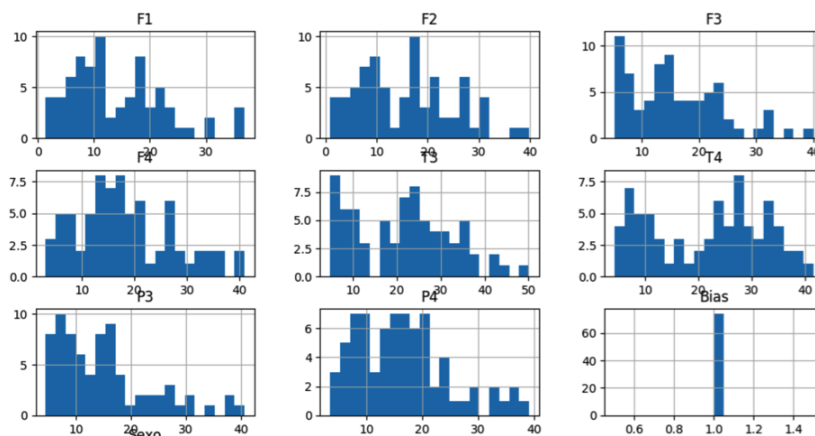


Figura 2.5: Espectros de PR de la actividad EEG en rango de beta registrada en las distintas áreas o derivaciones corticales.

Finalmente, se realizó un análisis exploratorio de los datos EEG de la banda beta en base al coeficiente de correlación de Pearson. Es decir, en base al grado de acoplamiento (r_{EEG}) entre dos áreas corticales el cual va de -1 a $+1$.

- Cercano al -1 , indica que las señales EEG de las dos áreas corticales están desfasadas (correlación negativa).
- Cercano al 1 , indica que las señales EEG de las dos áreas corticales están acopladas (correlación positiva).
- Cercano al 0 , quiere decir que no hay correlación alguna entre ambas señales EEG.

Los datos de correlación EEG en el rango de beta se describen en la section 2.2.

Tras realizar el respectivo análisis exploratorio a estos últimos datos, se comprobó que su distribución cercana a la normal, tal como se observa en la section 2.2.

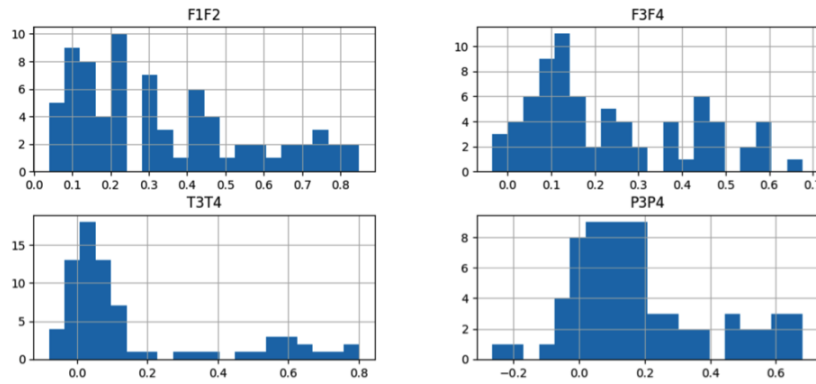


Figura 2.6: Distribución de los valores de correlación entre las señales EEG de la banda beta de áreas corticales correspondientes al hemisferio izquierdo y derecho.

Es en base a lo anterior, que se decidió utilizar ambas tablas para determinar con cuál de los dos parámetros con distribución normal, la potencia relativa o la rEEG, podían proporcionar mejores resultados.

F1	F2	F3	F4	T3	T4	P3	P4	Bias	Sexo
236.99	410.31	288.99	323.65	357.26	274.03	221.33	133.56	0	1
968.53	909.2	876.23	677.89	327.77	269.9	353.01	140.11	0	1
481.51	570.39	567.9	721.89	604.66	776.94	452.36	884.06	0	1
430.1	407.58	474.47	400	510.67	493.63	482.93	382.67	0	1
619.73	593.4	831.13	804.02	989.27	1046.9	713.29	528.91	0	1
152.41	158	186.16	178.8	236.5	258.57	2951.96	135.7	0	1
402.64	446.28	226.95	342.12	252.3	227.61	8655.22	2776.45	0	1
415.18	305.27	357.91	268.72	268.63	208.37	230.47	300.26	0	1
958.05	849.85	745.28	708.25	980.97	1009.96	4053.61	11460.73	0	1
546.59	484.68	1055.98	816.46	765.6	602.3	1351.7	912	0	1
410.94	272.23	439.68	377.96	453.57	432.86	190.83	233.18	0	1
1048.25	1169.49	596.21	1186	338.05	258.55	357.83	87.65	0	1
485.16	500.22	750.51	1813.13	832.17	686.41	4116	2289.11	0	1
1107.96	1186.15	1200.98	1111.13	1009.41	713.44	1599.58	5561.02	0	1
657.85	578.39	695.22	488.3	792.91	611.93	1091.46	1460.42	0	1
1101.27	898.3	1361.01	1168.18	1690.25	1187.53	1364.25	1219.4	0	0
1191.24	2864.94	669.13	695.03	1081.87	322.8	636.54	686.64	0	0
1882.51	1872.2	3299.78	1786.1	1337.22	1957.72	1090.98	1703.62	0	0
3180.73	2269.29	1627.62	2527.36	397.07	428.65	504.83	444.16	0	0
627.48	551.36	699.61	524.6	3614.24	394.09	1106.59	523.87	0	0
545.7	503.79	952.25	1083.85	1626.72	1296.25	641.3	489.96	0	0
445.74	397.62	753.13	589.67	1311.02	857.51	581.54	600.09	0	0
1159.61	1082.99	857.38	796.27	1935.96	4659.78	990.45	1021.72	0	0
935.05	653.57	949.34	945.06	631.85	413.41	860.16	1140.5	0	0
544.17	570.82	557.4	594.96	1612.19	1552.46	453.73	456.92	0	0
1445.81	1904.88	750.09	1027.96	2502.77	8948.82	664.93	527.77	0	0
618.91	436.2	570.56	336.14	2845.85	2628.47	509.66	398.49	0	0
4743.75	4157.38	9402.22	2853.63	9114.55	9737.42	1982.86	1785.81	0	0
1109.94	1218.5	931.57	1177.42	22489.74	8013.73	1280.47	1126.22	0	0
640.88	512.11	765.16	511.89	600.71	301.86	1127.89	923.46	0	0

Tabla 2.1: Valores de PA (microvolts) de la actividad EEG en el rango de beta registrada en las distintas áreas corticales o derivaciones de los 30 sujetos durante la condición basal ojos abiertos.

F1: Se refiere a la potencia absoluta calculada en el área cortical frontopolar del hemisferio izquierdo del cerebro.

F2: Se refiere a la potencia absoluta calculada en el área cortical frontopolar del hemisferio derecho del cerebro.

F3: Se refiere a la potencia absoluta calculada en el área cortical dorsolateral del hemisferio izquierdo del cerebro.

F4: Se refiere a la potencia absoluta calculada en el área cortical dorsolateral del hemisferio derecho del cerebro.

T3: Se refiere a la potencia absoluta calculada en el área cortical temporal del hemisferio izquierdo del cerebro.

T4: Se refiere a la potencia absoluta calculada en el área cortical temporal del hemisferio derecho del cerebro.

P3: Se refiere a la potencia absoluta calculada en el área cortical parietal del hemisferio izquierdo del cerebro.

P4: Se refiere a la potencia absoluta calculada en el área cortical parietal del hemisferio derecho del cerebro.

Bias: Es una constante que equivale a todas las variables no controladas.

Sexo: Se refiere al sexo del sujeto en cuestión.

F1	F2	F3	F4	T3	T4	P3	P4	Bias	Sexo
7.14	25.16	7.71	14.49	4.73	4.24	6.06	15.88	1	1
22.16	21.32	15.6	15.82	10.55	10.31	13.3	21.08	1	1
5.9	6.46	7.63	7.74	5.36	5.68	16.06	15.36	1	1
3.85	2.72	6.09	4.78	6.19	6.06	16.19	12.35	1	1
6.93	7.03	12.53	13.35	7.25	6.18	15.05	10.97	1	1
7.94	8.29	8.97	8	10.03	11.87	33.8	15.82	1	1
10.32	16.22	11.95	18.19	9.76	8.9	37.92	39.11	1	1
7.14	5.24	7.95	6.19	6.61	6.22	9.35	13.24	1	1
11.35	7.6	10.49	11.67	9.57	10.6	38.26	35.49	1	1
8.93	9.73	14.24	11.29	9.92	8.84	29.19	23.17	1	1
16.67	8.88	24.99	16.01	28.14	26.51	24.62	18.58	1	0
17.57	27.95	31.96	28.36	50.11	34.05	15.81	21.14	1	0
22.33	26.52	13.59	26.96	23.3	29.09	11.15	24.06	1	0
25.78	26.96	20.31	26.77	37.54	30.77	17.53	27.08	1	0
17.82	20.9	24.05	26.01	31.72	33.38	23.31	25.92	1	0
9.32	10.41	15.15	13.76	21.97	18.82	13.11	16.74	1	0
26.63	17.1	24.67	16.02	25.54	28.35	21.47	12.39	1	0
16.21	10.64	20.43	15.04	35.48	33.71	16.73	15.95	1	0
23.66	27.39	24.08	34.76	29.69	27.15	17.09	18.35	1	0
35.29	28.92	40.02	30.29	42.66	35.26	40.63	29.48	1	0

Tabla 2.2: Datos EEG en PR (porcentaje de energía) de la banda beta calculada en cada área cortical.

F1F2	F3F4	T3T4	P3P4	Bias	Sexo
0.3386	0.2784	0.5243	-0.0014	1	1
0.3557	0.4525	0.7	-0.0738	1	1
0.7782	0.4953	0.6029	-0.0367	1	1
0.7506	0.5726	0.5851	0.069	1	1
0.7276	0.3746	0.4584	-0.0949	1	1
0.7253	0.468	0.3623	0.1024	1	1
0.437	0.4355	0.7753	0.0516	1	1
0.5472	0.4312	0.6137	0.0007	1	1
0.619	0.5815	0.6613	0.1977	1	1
0.7699	0.4577	0.7143	0.054	1	1
0.0632	0.0413	-0.0001	0.0327	1	0
0.1256	0.0539	-0.0045	0.1686	1	0
0.1081	0.085	0.0367	0.1066	1	0
0.1418	0.0917	0.034	0.0553	1	0
0.1035	0.1129	0.0703	0.1139	1	0
0.4552	0.27	0.072	0.2839	1	0
0.1756	0.1541	0.1692	0.1859	1	0
0.3833	0.2173	0.1927	0.288	1	0
0.1224	0.0416	0.1067	0.1744	1	0
0.1187	0.1138	0.118	0.1246	1	0

Tabla 2.3: Valores de correlación entre las señales EEG en el rango de beta de las áreas corticales izquierda y derecha del cerebro correspondientes a los 30 sujetos que realizaron la prueba.

F1-F2: correlación entre las señales EEG de las áreas frontopolares izquierda y derecha.

F3-F4: correlación entre las señales EEG de las áreas dorsolaterales izquierda y derecha.

T3-T4: correlación entre las señales EEG de las áreas temporales izquierda y derecha.

P3-P4: correlación entre las señales EEG de las áreas parietales izquierda y derecha.

Bias: Es una constante que equivale a todas las variables no controladas.

Sexo: Se refiere al sexo del sujeto en cuestión.

3 Descripción de modelos

Para la determinación del funcionamiento cerebral de una persona, se decidió utilizar como base un modelo de regresión cuadrática por mínimos cuadrados. La razón por la cual se decidió utilizar este modelo de regresión cuadrática, es debido a su gran historia, relevancia y antecedentes que ha tenido este modelo de regresión en los campos de investigación de las neurociencias [12, 37, 38]. Además, para pruebas adicionales y comprobar que los resultados de la regresión cuadrática son válidos, se decidió también realizar la clasificación por medio de regresión logística y por medio de clasificación por K-medias, esperando que los resultados con ambos métodos fueran muy similares a los de la regresión cuadrática por mínimos cuadrados.

A continuación, se describe cada uno de los modelos de aprendizaje que se usaran para el presente trabajo de obtención de grado:

3.1 Descripción de los modelos a utilizar

3.1.1 Regresión cuadrática por Mínimos Cuadrados

Es una técnica estadística de aprendizaje la cual, consiste en encontrar los parámetros desconocidos o coeficientes de una función de regresión muestral para formar una línea recta que se ajuste a un conjunto dado de datos, de tal forma que la suma residual de cuadrados sea tan pequeña como sea posible [39, 40, 41]. En otras palabras, hay que minimizar la suma de los cuadrados de los errores. Esto se realiza de la siguiente manera.

Supongamos que tenemos un conjunto de datos con dos variables, una variable independiente (X) y una variable dependiente (Y). Lo que buscamos, es generar una recta donde cada punto predicho de Y (\hat{Y}) forme una recta que divida los valores de una clase de la otra. Matemáticamente hablando, buscaríamos generar esta recta.

$$\hat{Y}_i = \beta_0 + \beta_1 X_i \quad (3.1)$$

Lógicamente, realizando este cálculo habrá algunos errores en la Y predicha, lo que buscamos con mínimos cuadrados es minimizando la suma de los cuadrados de todos los errores de Y predicha, utilizando la siguiente formula:

$$S(\beta_0, \beta_1) = \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_i))^2 \quad (3.2)$$

Para lograrlo, hay que encontrar los valores óptimos de β_0 y β_1 que minimicen esta fórmula, pero ¿Cómo encontramos dichos valores óptimos? Para ello, hay que calcular las derivadas parciales de la función de costo de mínimos cuadrados con respecto a β_0 y β_1 .

$$\frac{\partial S}{\partial \beta_0} = -2 \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_i)) = 0 \quad (3.3)$$

$$\frac{\partial S}{\partial \beta_1} = -2 \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_i)) x_i = 0 \quad (3.4)$$

De aquí obtenemos las ecuaciones normales:

$$\sum_{i=1}^n y_i = n\beta_0 + \beta_1 \sum_{i=1}^n x_i \quad (3.5)$$

$$\sum_{i=1}^n y_i x_i = \beta_0 \sum_{i=1}^n x_i + \beta_1 \sum_{i=1}^n x_i^2 \quad (3.6)$$

Finalmente, las soluciones para β_0 y β_1 son:

$$\beta_1 = \frac{n \sum_{i=1}^n y_i x_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \quad (3.7)$$

$$\beta_0 = \frac{1}{n} \sum_{i=1}^n y_i - \beta_1 \frac{1}{n} \sum_{i=1}^n x_i \quad (3.8)$$

Obteniendo así que:

$$\beta_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (3.9)$$

$$\beta_0 = \bar{y} - \beta_1 \bar{x} \quad (3.10)$$

Ya con los coeficientes encontrados, podemos calcular el error cuadrático medio o R_2 para evaluar así la calidad del modelo, y finalmente graficar la recta que buscábamos para observar la división lineal de los datos [39, 40, 41].

3.1.2 Regresión Logística

Es una técnica estadística de aprendizaje la cual se utiliza principalmente para problemas de regresión con una variable dependiente dicotómica, es decir, cuya variable objetivo (Y) es siempre 1 o 0 [39, 41].

Consiste en utilizar la función logística para formar una línea curva en forma de S en el conjunto de datos, a la cual se ajustan todas las predicciones de la regresión utilizando la función logística para asegurar que estas siempre tengan valores entre 0 y 1.

En primer lugar, para predecir si un valor será igual a 1, se utiliza la siguiente fórmula.

$$P(y = 1|\mathbf{x}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}} \quad (3.11)$$

Mientras que para predecir si un valores será igual a 0, se utiliza la siguiente fórmula.

$$P(y = 0|\mathbf{x}) = 1 - P(y = 1|\mathbf{x}) \quad (3.12)$$

Como se puede observar, al igual que con la regresión por mínimos cuadrados, lo que hay que hacer para resolver este problema es encontrar los valores de β_0 y β_1 , para lo cual, primero hay que determinar la función de costo que utilizaremos, siendo esta, la función de log-verosimilitud.

$$\mathcal{L}(\beta) = \sum_{i=1}^n [y_i \log(P(y_i = 1|\mathbf{x}_i)) + (1 - y_i) \log(1 - P(y_i = 1|\mathbf{x}_i))] \quad (3.13)$$

Ya que tenemos nuestra función de costo, hay que derivar la log-verosimilitud con respecto a cada β , utilizando la fórmula del gradiente descendente.

$$\frac{\partial \mathcal{L}(\beta)}{\partial \beta_j} = \sum_{i=1}^n (y_i - P(y_i = 1|\mathbf{x}_i)) x_{ij} \quad (3.14)$$

Finalmente, los coeficientes β se irán actualizando cada vez mediante el algoritmo de gradiente descendente, hasta llegar a los coeficientes óptimos, utilizando α como tasa de aprendizaje del modelo.

$$\beta_j \leftarrow \beta_j + \alpha \frac{\partial \mathcal{L}(\beta)}{\partial \beta_j} \quad (3.15)$$

Una vez tenemos los coeficientes de β , procedemos a resolver para cada punto del dataset, determinando así si pertenece a la clase 1 o clase 0 [39, 41].

3.1.3 Clasificación por K-medias

La clasificación K-medias es un algoritmo de aprendizaje automatizado no supervisado que sirve para agrupar un conjunto de datos sin determinar en un número específico de grupos o clústeres [42, 43].

Consiste en seleccionar un par de puntos al azar dentro del mapa de conjunto de datos para que funcionen como centroides de cada clúster.

A cada punto del conjunto de datos se le asigna su centroide más cercano, calculando la distancia euclidiana del mismo hacia cada centroide y ya que a cada punto se le ha asignado un centroide, se generan nuevos centroides calculando la media de todos los puntos que se le asignaron.

Asignación de los puntos a los centroides más cercanos:

$$y_i^{(t)} = \arg \min_k \|x_i - C_k^{(t-1)}\|^2 \quad (3.16)$$

Donde $y_i^{(t)}$ es el índice del cluster al que pertenece el punto x_i en la iteración t .

Actualización de los centroides:

$$C_k^{(t)} = \frac{1}{|S_k^{(t)}|} \sum_{x_i \in S_k^{(t)}} x_i \quad (3.17)$$

Donde $S_k^{(t)}$ es el conjunto de puntos asignados al cluster k en la iteración t .

Este proceso se repite una y otra vez, hasta que cada centroide ya no cambie más su posición de manera significativa [42, 43].

$$\|C_k^{(t)} - C_k^{(t-1)}\| < \epsilon, \quad \text{para todo } k \quad (3.18)$$

Donde ϵ es un umbral de convergencia pequeño.

3.2 Descripción de las métricas

De funcionar correctamente el modelo de regresión cuadrática, la proporción de los resultados que se esperarían obtener serían los siguientes:

- 66 por ciento de los sujetos masculinos con funcionalidad tipo A (y predicha > 0.5).
- 33 por ciento de los sujetos masculinos con funcionalidad tipo B (y predicha < 0.5).
- 33 por ciento de los sujetos femeninos con funcionalidad tipo A (y predicha > 0.5).

- 66 por ciento de los sujetos femeninos con funcionalidad tipo B (y predicha < 0.5).

Estos resultados que se esperaría obtener son principalmente hipotetizados en primer lugar, considerando los estudios donde se muestra una variabilidad importante en la activación y el grado de conexiones funcionales que presentan hombres y mujeres sobre todo en relación a la ejecución de tareas cognitivas como memoria de trabajo [6, 7, 18, 44, 45], procesamiento de estímulos relevantes; por ejemplo estímulos con contenido displacentero como la estimulación fónica [46] o placentero, como los estímulos con contenido sexual [19] y muy pocos durante estados de reposo [29, 47, 48] segundo, en base a las encuestas nacionales que muestran que en México, el 28 por ciento de egresados de carreras técnicas como la tecnología, ingeniería y matemáticas, son mujeres, tal que según el informe de la UNICEF de diciembre de 2023, el 38 por ciento de las mujeres mexicanas estudian carreras técnicas, y tercero, en base a las aproximaciones que he obtenido con los cálculos realizados a la fecha con muestras pequeñas.

3.3 Explicación del código paso a paso

Ya con los datos a trabajar listos para ser procesados, se empezó a trabajar en el código para la regresión y clasificación de los sujetos a una funcionalidad cerebral.

Este código fue realizado en Python con uso de las librerías `numpy` y `pandas`.

El primer paso que se realizó fue importar las librerías y datos que se necesitarían para la regresión. Los datos se importaron a un dataframe de `pandas` desde un libro de Excel.

Ya con los datos dentro del dataframe, lo siguiente que se hizo fue crear una función llamada “calcular resultados” la cual tuviera como parámetros de entrada el dataframe a utilizar y un array donde se guardarán los resultados de cada regresión.

Esta función serviría para:

1. Dividir los datos del dataframe en 70 por ciento de entrenamiento y 30 por ciento de testeo aleatoriamente, usando la función “train test split” de Sklearn.
2. Separar las columnas de entrada (X) y salida (Y) tanto de los datos de entrenamiento como de testeo.
3. Ejecutar el modelo de regresión seleccionado.
4. Guardar los resultados de la regresión en la lista de resultados.

5. Contar el mínimo de veces por las que cada fila del dataframe completo ha estado en los datos de testeo.

Al final, esta función retorna el mínimo de veces.

Ya con la función “calcular resultados” creada, lo siguiente que se hizo fue crear un bucle el cual, ejecutará la función “calcular resultados” de forma indefinida hasta que el mínimo de veces por las que cada fila del dataframe completo haya estado en los datos de testeo sea mayor a 6.

Por lo tanto, cada vez que se ejecute la función “calcular resultados”, habrá un 70 por ciento de datos de entrenamiento diferentes a los de la ejecución anterior. Por lo cual, los coeficientes resultantes de la regresión serán distintos cada vez que se ejecute la función, y por lo tanto, los resultados de la multiplicación de los coeficientes por los datos de testeo, podrían también ser diferentes.

Por ejemplo, supongamos que el sujeto 7 es un sujeto de sexo masculino cuya funcionalidad cerebral se desea determinar si es tipo A o tipo B.

Para ello, contamos con los datos de otros 29 sujetos, dando así un total de 30 sujetos (15 sexo masculino y 15 sexo femenino).

Al ejecutarse la función “calcular resultados” por primera vez, la función “train test split” de Sklearn, coloca al sujeto 7 entre los 6 sujetos de testeo, por lo cual, tras obtener los coeficientes de la regresión cuadrática con los actuales 24 sujetos de training, se determinó que con dichos coeficientes, el sujeto 7 tiene funcionalidad cerebral tipo A.

Para este punto, 6 de los 30 sujetos, ya formaron parte 1 vez de los datos de testeo, mientras que los otros 24, han formado parte 0 veces de los datos de testeo. Dado que el mínimo actual de 24 sujetos es igual a 0, se vuelve a ejecutar la función “calcular resultados”.

En esta segunda ejecución, el sujeto 7 vuelve a ser colocado entre los 6 sujetos de testeo, pero ahora, los 24 sujetos que fueron seleccionados para el entrenamiento, son diferentes a los 24 que fueron seleccionados en la prueba anterior. Por lo tanto, los coeficientes dados de la regresión, también serán diferentes a los anteriores. Sin embargo, a pesar de ahora tener coeficientes diferentes, el resultado de la clasificación para el sujeto 7, volvió a ser funcionalidad cerebral tipo A.

Para este punto el sujeto 7 ya ha formado parte 2 veces de los datos de testeo, otros más han formado parte 1 vez de los datos de testeo, y otros más han formado parte 0 veces de los datos de testeo.

Dado que el mínimo actual de algunos sujetos todavía sigue siendo igual a 0, se vuelve a ejecutar la función “calcular resultado”.

Supongamos que en esta tercera ejecución, el sujeto 7 ahora es seleccionado para formar parte de los sujetos de entrenamiento, por lo tanto, el número de veces que ha formado parte de los sujetos de

testeo, se mantiene 2.

Lo mismo pasa para la cuarta y quinta ejecución, hasta la sexta ejecución vuelve a formar parte de los sujetos de testeo. Para esta sexta ejecución, los coeficientes de la regresión han vuelto a cambiar, ya que los 24 sujetos de entrenamiento son diferentes a los de la primera y segunda ejecución. Sin embargo, aún con estos coeficientes diferentes, el sujeto 7 vuelve a ser clasificado con funcionalidad cerebral tipo A.

Hasta ahora, el sujeto 7 ha formado parte de los sujetos de testeo 3 veces, de las cuales, las 3 veces ha sido clasificado con funcionalidad tipo A, sin importar que los coeficientes de la regresión hayan sido diferentes en cada ejecución.

Por otra parte, digamos que el sujeto 8 aún no ha formado parte ninguna vez de los sujetos de testeo (siempre ha estado entre los sujetos de entrenamiento). Por lo tanto, el mínimo actual de este sujeto sigue siendo 0, y por lo tanto, la función “calcular resultado” se vuelve a ejecutar.

La función “calcular resultado” se vuelve a ejecutar una y otra vez hasta que finalmente, cada uno de los sujetos en el dataframe, ha formado parte de los sujetos de testeo un mínimo de 6 veces.

Ya que todos los sujetos del dataframe han formado parte de los sujetos de testeo un mínimo de 6 veces, se detiene el bucle y se cuenta, cuantas veces fue clasificado cada sujeto en las funcionalidades cerebrales.

Retomando el ejemplo anterior, para el momento en que el bucle finalmente se detiene, observamos que el sujeto 7 formó parte de los sujetos de testeo 10 veces, de las cuales, 8 veces fue clasificado en funcionalidad tipo A y 2 veces fue clasificado en la funcionalidad tipo B. Por lo tanto, determinamos que la funcionalidad cerebral dominante de dicho sujeto es la funcionalidad tipo A.

3.4 Experimentos o simulaciones realizadas

Ahora sí, dejando los ejemplos de lado y ya explicado el código realizado, se procedió a realizar las distintas pruebas. Las cuales fueron:

1. Regresión por Mínimos Cuadrados usando datos de Potencia Relativa.
2. Regresión Logística usando datos de Potencia Relativa.
3. Clasificación de datos de Potencia Relativa por K-Medias.
4. Regresión por Mínimos Cuadrados usando datos de Potencia Relativa seleccionando datos SIN correlación.
5. Regresión Logística usando datos de Potencia Relativa seleccionando datos SIN correlación.

6. Clasificación de datos de Potencia Relativa por K-Medias seleccionando datos SIN correlación
7. Regresión por Mínimos Cuadrados usando datos de correlación EEG.
8. Regresión Logística usando datos de Correlación EEG
9. Clasificación de datos de Correlación EEG por K-Medias.

Cabe mencionar que para todas estas pruebas, se utilizaron los datos totales descritos anteriormente, conformados por las señales EEG (en el rango de beta) de 37 sujetos del sexo masculino y 37 sujetos del sexo femenino.

4 Resultados y discusión

4.1 Resultados de las pruebas realizadas usando todas las variables de PR

4.1.1 Primera prueba – Regresión por Mínimos Cuadrados usando datos de PR

Para esta primera prueba, se ejecutó el bucle de la función “calcular resultados” un total de 42 veces, hasta que finalmente cada uno de los 74 sujetos del dataframe, formó parte de los sujetos de testeo un mínimo de 6 veces, dando como resultado final los datos de la Sección 4.1.1.

ID_Sujeto	Sexo	Promedio	Fun_Ceb	
0	0	H	0.831867	A
1	1	H	0.774152	A
2	2	H	0.877054	A
3	3	H	0.801830	A
4	4	H	0.714258	A
...
69	69	F	0.545062	A
70	70	F	0.331227	B
71	71	F	0.200111	B
72	72	F	0.291720	B
73	73	F	0.085746	B

74 rows x 5 columns

Tabla 4.1: Resultados de las pruebas de Regresión por Mínimos Cuadrados utilizando los datos de PR de las señales EEG de la banda beta.

- El 67.6 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo A.
- El 32.4 por ciento de los sujetos masculinos se clasificaron con

funcionalidad tipo B.

- El 29.7 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo A.
- El 70.3 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo B.

Viendo estos resultados, se observó que para esta primera prueba, la regresión dio como resultado los valores esperados para la caracterización de funcionalidad cerebral.

4.1.2 Segunda prueba – Regresión Logística usando datos de Potencia Relativa

Siguiendo la metodología respectiva, se obtuvieron los resultados descritos en la Sección 4.1.2.

ID_Sujeto	Sexo	Promedio	Fun_Ceb
0	H	0.923077	A
1	H	1.000000	A
2	H	1.000000	A
3	H	1.000000	A
4	H	1.000000	A
...
69	F	0.916667	A
70	F	0.000000	B
71	F	0.000000	B
72	F	0.000000	B
73	F	0.000000	B

74 rows x 5 columns

Tabla 4.2: Resultados de las pruebas de Regresión Logística utilizando los datos de PR.

- El 70.3 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo A.
- El 29.7 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo B.
- El 32.4 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo A.

- El 67.6 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo B.

4.1.3 Tercera prueba – Clasificación de datos de Potencia Relativa por K-Medias

En base al análisis de K-medias, se encontró que:

- El 67.6 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo A.
- El 32.4 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo B.
- El 29.7 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo A.
- El 70.3 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo B.

Posterior a la aplicación de cualquiera de las tres técnicas de clasificación, se observó que se obtuvieron casi los mismos resultados o al menos muy parecidos para cada uno de los sujetos de la prueba, lo cual significó un avance bastante positivo.

4.2 Resultados de las pruebas realizadas usando sólo las variables con poca o nula correlación de PR

Posteriormente, se procedió a aplicar las mismas pruebas pero ahora utilizando únicamente variables que estén no correlacionadas, dado que esta suele ser una buena práctica cuando se está trabajando con regresión. Así que para ello, se realizaron los siguientes esquemas de análisis visual para determinar ¿Cuáles son las variables que están menos correlacionadas?

En la section 4.2, se puede observar un pairplot donde se grafica la distribución de cada variable de la tabla de datos de PR (section 2.2) contra cada una de las otras variables de la misma section 2.2. En color azul están todos los sujetos de sexo masculino y en color naranja todos los sujetos de sexo femenino.



Figura 4.1: Pairplot de distribución de las variables contra otras variables de potencia relativa.

Por otra parte en la section 4.2, observamos un mapa de calor (Heatmap) donde se grafican los valores de correlación de cada una de las variables de potencia relativa contra el resto de variables de la section 4.2.

- Entre más intenso sea el color rojo, más cercana a 1 será la correlación entre las 2 variables, es decir, más correlacionadas de manera positiva estarán dichas variables.
- Entre más intenso sea el color azul, más cercana a -1 será la correlación entre las 2 variables, es decir, más correlacionadas de manera negativa estarán dichas variables.
- Entre más intenso sea el color blanco, más cercana a 0 será la correlación entre las 2 variables, es decir, menos correlacionadas estarán dichas variables.

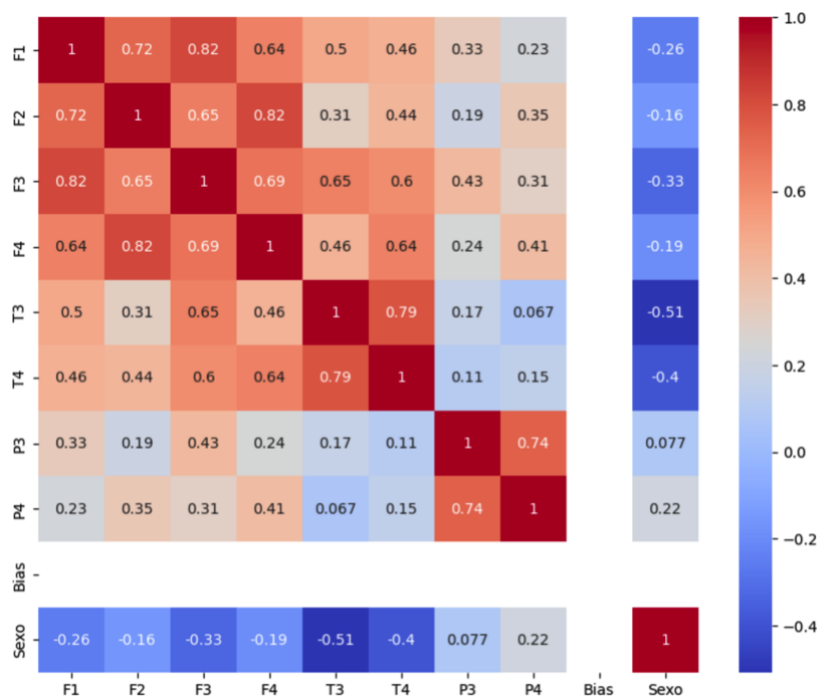


Figura 4.2: Pairplot de distribución de las variables contra otras variables de potencia relativa.

Viendo los resultados obtenidos por el pairplot y el heatmap para las variables de potencia relativa, se determinó que las variables F2, F4 y P3 son las que mayores índices de correlación poseen y por lo tanto, podrían ser no útiles para la regresión.

Considerando lo anterior, se procedió a volver a realizar las mismas tres pruebas anteriores, pero ahora utilizando únicamente las variables sin alta correlación, siendo estas: F1, F3, T3, T4 y P4.

4.2.1 Cuarta prueba – Regresión por Mínimos Cuadrados usando datos de Potencia Relativa seleccionando datos SIN correlación

Posterior a la realización de la prueba mencionada, se obtuvieron los datos de la Sección 4.2.1.

ID_Sujeto	Sexo	Promedio	Fun_Ceb
0	H	0.895687	A
1	H	0.787375	A
2	H	0.869546	A
3	H	0.797279	A
4	H	0.664284	A
...
69	F	0.518005	A
70	F	0.298588	B
71	F	0.184053	B
72	F	0.309085	B
73	F	0.180815	B

74 rows x 5 columns

Tabla 4.3: Resultados de Regresión por Mínimos Cuadrados con Potencia Relativa usando datos SIN correlación.

- El 73.0 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo A.
- El 27.0 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo B.
- El 27.0 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo A.
- El 73.0 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo B.

4.2.2 Quinta prueba – Regresión Logística usando datos de Potencia Relativa seleccionando datos SIN correlación

- El 73.0 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo A.
- El 27.0 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo B.

ID_Sujeto	Sexo	Promedio	Fun_Ceb
0	H	1.000000	A
1	H	1.000000	A
2	H	1.000000	A
3	H	1.000000	A
4	H	1.000000	A
...
69	F	0.578947	A
70	F	0.000000	B
71	F	0.000000	B
72	F	0.000000	B
73	F	0.000000	B

74 rows x 5 columns

Tabla 4.4: Resultados de Regresión Logística usando con Potencia Relativa seleccionando datos SIN correlación.

- El 29.7 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo A.
- El 70.3 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo B.

4.2.3 Sexta prueba – Clasificación de datos de Potencia Relativa por K-Medias seleccionando datos SIN correlación

Los resultados de PR sin correlación con K-medias indicaron que:

- El 67.6 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo A.
- El 32.4 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo B.
- El 29.7 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo A.
- El 70.3 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo B.

Los resultados de esta cuarta, quinta y sexta prueba, si bien no fueron malos, tampoco fueron tan buenos como lo fueron los resultados de las primeras tres pruebas.

Viendo lo anterior, y posterior a la consulta con los investigadores responsables de los datos en el Instituto de Neurociencias de la Universidad de Guadalajara, la indicación a seguir fue que, cuando se hace una regresión de este tipo de datos, los datos más valiosos son aquellos que tienen más correlación y por lo tanto, no deberían quitarse de nuestro análisis de regresión. Por lo tanto, estas últimas tres pruebas realizadas (cuarta, quinta y sexta) se descartaron como método a utilizar para la clasificación.

4.3 *Resultados de las pruebas realizadas usando todas las variables de correlación EEG*

Finalmente, para las últimas tres pruebas restantes, se decidió volver a utilizar todas las variables, sin importar si están correlacionadas o no, pero ahora utilizando los datos del coeficiente de correlación de Pearson para ver si los resultados de la regresión mejoraban, empeoraban o quedaban igual que los anteriores.

4.3.1 *Séptima prueba – Regresión por Mínimos Cuadrados usando datos de correlación EEG*

Los resultados correspondientes a esta séptima prueba, aparecen en la Sección 4.3.1.

ID_Sujeto	Sexo	Promedio	Fun_Ceb
0	H	0.825976	A
1	H	0.936562	A
2	H	1.155646	A
3	H	0.923171	A
4	H	1.199085	A
...
69	F	0.416304	B
70	F	0.444068	B
71	F	0.431263	B
72	F	0.444618	B
73	F	0.485165	B

74 rows x 5 columns

Tabla 4.5: Resultados de la Regresión por Mínimos Cuadrados usando datos de correlación EEG.

- El 67.7 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo A.
- El 32.3 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo B.
- El 18.9 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo A.
- El 81.1 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo B.

4.3.2 Octava prueba – Regresión Logística usando datos de Correlación EEG

Los resultados correspondientes a esta octava prueba, aparecen en la Sección 4.3.2.

ID_Sujeto	Sexo	Promedio	Fun_Ceb
0	H	1.000000	A
1	H	1.000000	A
2	H	1.000000	A
3	H	1.000000	A
4	H	1.000000	A
...
69	F	0.000000	B
70	F	0.428571	B
71	F	0.210526	B
72	F	0.588235	A
73	F	0.714286	A

74 rows x 5 columns

Tabla 4.6: Resultados de Regresión Logística usando datos de correlación EEG.

- El 62.2 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo A.
- El 37.8 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo B.
- El 29.7 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo A.

- El 70.3 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo B.

4.3.3 *Novena prueba – Clasificación de datos de Correlación EEG por K-Medias*

Los resultados de K-medias fueron los siguientes:

- El 40.6 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo A.
- El 59.4 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo B.
- El 08.2 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo A.
- El 91.8 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo B.

Como se puede observar, los resultados de la séptima y octava prueba, resultaron ser no tan malos, y de hecho se asemejan un poco con los resultados de las primeras 2 pruebas. Sin embargo la novena y última prueba, siendo la clasificación por K-medias, reportó resultados nada favorables en cada una de sus aplicaciones, dando resultados muy alejados de lo que se esperaba.

Debido a estos resultados, se decidió mejor quedarse con los resultados de las primeras dos pruebas realizadas con los datos de PR, que se mantuvieron como los mejores resultados, más coherentes con la hipótesis planteada y mejor explicables desde el punto de vista neurofisiológico.

4.4 *Selección del mejor modelo*

Para determinar cuál sería el mejor modelo de todos, y, por lo tanto, el seleccionado para ser utilizado en futuras implementaciones, se seleccionaron los modelos de las tres primeras pruebas, es decir, aquellos que utilizaron todas las variables de la potencia relativa.

De estos tres, el primero en ser descartado fue el de K-medias, dado su bajo desempeño tanto en la sexta prueba (con PR), como en la novena (con valores de correlación), mientras que en la tercera prueba los resultados fueron ligeramente más bajos respecto a las pruebas 1 y 2.

Al descartar la clasificación por K-medias, se tuvo que seleccionar entre el modelo de regresión por mínimos cuadrados y el modelo por regresión logística.

Como bien se observa con los datos obtenidos, ambos modelos ofrecen resultados bastante buenos y similares, muy apegados a los resultados esperados. Por lo anterior, después de exponer estos resultados (de ambos modelos) a los doctores del instituto de neurociencias, se determinó que el modelo a ser utilizado sería el de mínimos cuadrados con los valores de PR.

Tal decisión fue tomada sobre todo considerando, en primer lugar, el hecho de que el modelo de regresión de mínimos cuadrados es el que más ha sido utilizado en el campo de las neurociencias para analizar las señales electroencefalográficas, ofreciendo con ello una mayor confiabilidad. Y en segundo lugar, porque los resultados obtenidos coinciden con la hipótesis neurofisiológica propuesta en este trabajo, en la cual se planteó que el sexo de un individuo no determina su funcionalidad cerebral en estado de reposo.

Ya seleccionado el modelo a utilizar, finalmente se decidió comprobar que la regresión por mínimos cuadrados fuera el mejor modelo para este trabajo, comprobando que estos cumplieran con los cinco supuestos de la regresión.

4.5 *Supuestos de la Regresión*

Para comprobar que estos resultados fueran verídicos, los resultados obtenidos pasaron por los cinco supuestos de la regresión, los cuales son condiciones o requisitos que se deben de cumplir para asegurar que los resultados de un análisis de regresión, sean válidos y confiables.

Estos cinco supuestos de la regresión, son:

1. Linealidad: Indica que existe una relación lineal entre la variable dependiente y las variables independientes.
2. Normalidad: Implica que los errores siguen una distribución normal.
3. No multicolinealidad: Indica que las variables independientes, no deben de estar altamente correlacionadas entre sí.
4. Independencia: Implica que los errores o residuos, son independientes entre sí.
5. Homocedasticidad: Indica que la varianza de los errores o residuos, es constante en todos los niveles de las variables independientes.

De cumplirse estos cinco supuestos, se asegura la validez y confiabilidad de los resultados obtenidos con la regresión [49].

Como datos de entrada X, tomamos los mismos datos de PR de la banda EEG beta de los sujetos, mientras que como datos de salida Y, tomamos los resultados obtenidos por la regresión cuadrática de mínimos cuadrados, tal como se observa en la sección 4.5.

	F1	F2	F3	F4	T3	T4	P3	P4	Bias	Fun_Ceb
0	7.14	25.16	7.71	14.49	4.73	4.24	6.06	15.88	1	1
1	22.16	21.32	15.60	15.82	10.55	10.31	13.30	21.08	1	1
2	5.90	6.46	7.63	7.74	5.36	5.68	16.06	15.36	1	1
3	3.85	2.72	6.09	4.78	6.19	6.06	16.19	12.35	1	1
4	6.93	7.03	12.53	13.35	7.25	6.18	15.05	10.97	1	1
...
69	9.32	10.41	15.15	13.76	21.97	18.82	13.11	16.74	1	1
70	26.63	17.10	24.67	16.02	25.54	28.35	21.47	12.39	1	0
71	16.21	10.64	20.43	15.04	35.48	33.71	16.73	15.95	1	0
72	23.66	27.39	24.08	34.76	29.69	27.15	17.09	18.35	1	0
73	35.29	28.92	40.02	30.29	42.66	35.26	40.63	29.48	1	0

74 rows x 10 columns

Tabla 4.7: Tabla con las 9 variables de la PR y la función cerebral obtenida para cada sujeto por medio de la regresión.

Ya que cargamos nuestros datos, procedemos a comprobar cada uno de los supuestos de la regresión.

4.5.1 *Primer supuesto - Linealidad*

El primer supuesto a comprobar fue el de la linealidad de los resultados. Para ello, se observaron los resultados obtenidos contra sus residuos. Viendo que existe una clara línea recta divisora entre ambos resultados posibles (A o B), se comprueba que existe linealidad en los resultados y con ello se cumple el primer supuesto de la regresión.

4.5.2 *Segundo supuesto - Normalidad*

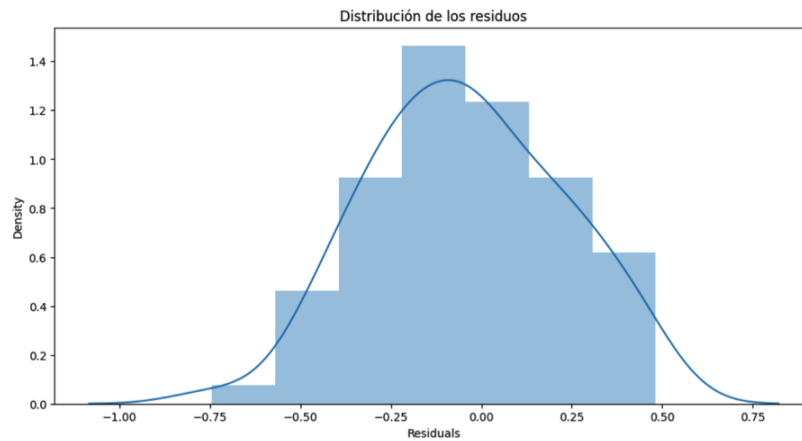
El segundo supuesto que se comprobó fue el de la normalidad, es decir, que los residuos del modelo sigan una distribución normal. Para ello, se realizó un test de Anderson-Darling el cual evalúa si un cierto conjunto de datos se ajusta a una distribución determinada. En este caso, se utilizó dicho test para comprobar si los residuos siguen una distribución normal y así, probar el segundo supuesto de la regresión.

Si el test obtiene una calificación de p-value menor a 0.05, quiere decir que la distribución no está normalizada.

Tras realizar el test en los datos, se obtuvo un p-value con valor de 0.89, así como una distribución normal representada en la Sección 4.5.2.

Supuesto 2: Los residuos están normalmente distribuidos

Usar la prueba de Anderson-Darling para comprobar la distribución normal de los r . Si el valor de p-value de la prueba es menor a 0.05 quiere decir que los residuos están normalmente distribuidos.



Supuesto cumplido
 Sesgo de los residuos: -0.01271750033743326
 Los residuos NO están sesgados

Figura 4.3: Histograma con distribución normalizada de los residuos del modelo seleccionado.

Como se puede observar tanto en la Sección 4.5.2 como en el p-value de 0.89, los residuos del modelo sí están normalizados, por lo tanto, se cumple con el segundo supuesto de la regresión.

4.5.3 Tercer supuesto - No Multicolinealidad

El tercer supuesto que se comprobó, fue el de la multicolinealidad, es decir, que las variables independientes del modelo, no estén altamente correlacionadas entre sí. Para ello, se realizó un mapa de calor (heatmap) en el cual, se pudiera ver gráficamente la colinealidad de todas las variables del modelo.

Con esta gráfica se realizó una prueba de factores de inflación de la varianza (VIF) en la cual, para cada una de las variables se comprueba su valor de colinealidad.

- Si alguna de las variables presenta un resultado mayor a 10, quiere decir que posiblemente haya un caso de multicolinealidad en esa variable.
- Si alguna de las variables presenta un resultado mayor a 100, quiere decir que seguro hay casos de multicolinealidad en dichas variables.
- Finalmente, si se obtiene un resultado menor o igual a 10, quiere decir que no existe ningún caso de multicolinealidad y por lo tanto, se cumple el tercer supuesto de la regresión.

Tras haber calculado el VIF para cada una de las variables del modelo, se obtuvieron los resultados de la Sección 4.5.3 y la Sección 4.5.3:

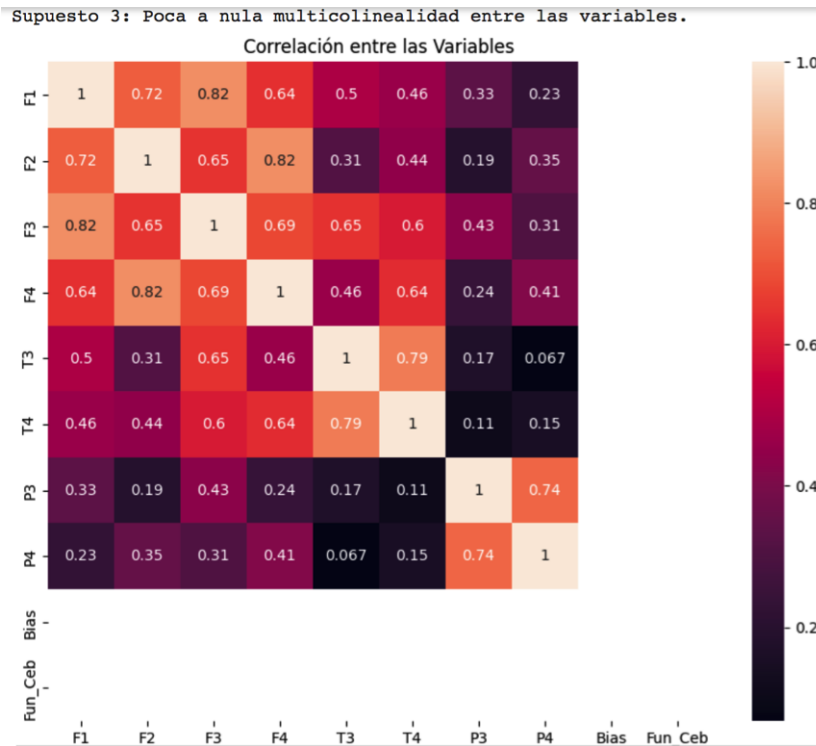


Figura 4.4: Mapa de calor de la correlación de las variables del modelo seleccionado.

```
Variance Inflation Factors (VIF).
> 10: Multicolinealidad posiblemente presente.
> 100: Multicolinealidad entre las variables.
-----
F1: 4.167985812014901
F2: 4.551869176461771
F3: 5.023092612680282
F4: 4.788704331892651
T3: 3.4866911710133297
T4: 3.6871557975875118
P3: 3.4108555131845457
P4: 3.2990106816228
Bias: 8.809368993072207

0 casos de posible multicolinealidad
0 casos de segura multicolinealidad

Supuesto cumplido
```

Figura 4.5: Resultados del VIF para cada una de las variables del dataset.

Tanto en el mapa de calor, como en los resultados del VIF, ninguna de las variables presenta resultados mayores a 10, por lo tanto, no existe ningún caso de multicolinealidad y se cumple el tercer supuesto de la regresión.

4.5.4 Cuarto supuesto - Independencia

El cuarto supuesto que se comprobó fue el de la Independencia, es decir, que los residuos del modelo sean independientes entre sí. Para ello, se realizó un test de Durbin-Watson, la cual, es una prueba estadística para detectar si existe autocorrelación entre los residuos del modelo.

- Si la prueba de Durbin-Watson obtiene un resultado menor a 1.5, quiere decir que existe una autocorrelación positiva, y por lo tanto, no se cumple el supuesto.
- Si obtiene un resultado mayor a 2.5, quiere decir que existe una correlación negativa y por lo tanto, tampoco se cumple el supuesto.
- Finalmente, si se obtiene un resultado mayor o igual a 1.5, y menor o igual a 2.5, entonces se considera que no existe correlación entre los residuos y por lo tanto, se cumple el cuarto supuesto de la regresión.

Tras haber realizado la prueba de Durbin-Watson, utilizando los residuales del modelo, se obtuvieron los resultados de la Sección 4.5.4:

```
Supuesto 4: No Autocorrelación

Realizar prueba de Durbin-Watson
Valores entre 1.5 < d < 2.5 muestran que no hay correlación entre los residuos.
0 to 2< hay correlación positiva.
>2 to 4 hay correlación negativa
-----
Durbin-Watson: 1.879007612581618
Poca a nula autocorrelación

Supuesto cumplido
```

Figura 4.6: Resultados de la prueba Durbin-Watson para comprobar independencia en los residuos.

Al ser 1.879 un resultado mayor a 1.5 y menor a 2.5, se considera que no existe autocorrelación entre los residuos del modelo y por lo tanto, se cumple el cuarto supuesto de la regresión.

4.5.5 Quinto supuesto - Homocedasticidad

Finalmente, el quinto supuesto que se comprobó fue el de la Homocedasticidad, es decir, que la varianza de los residuos sea constante en todos los niveles de las variables independientes. Para ello, se graficaron todos los residuos del modelo y se dibujó una línea naranja justo en el 0, dando como resultado la Sección 4.5.5.

En la Sección 4.5.5 se observa como la varianza de los residuos es constante a lo largo de todo el modelo y por lo tanto, se cumple con el quinto y último supuesto de la regresión.

Al haberse comprobado los cinco supuestos de la regresión, comprobamos que el modelo de regresión cuadrática de mínimos

Supuesto 5: Homoscedasticidad de los residuos

Los residuos deben tener una varianza relativamente constante.

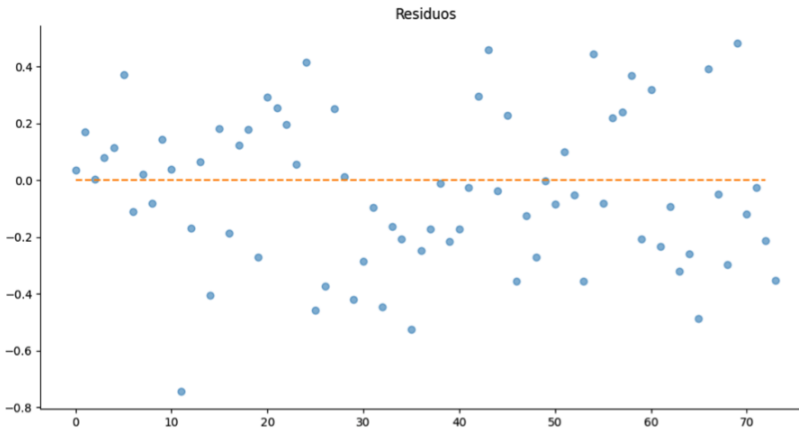


Figura 4.7: Varianza constante de los residuos para comprobar la homoscedasticidad de los mismos.

cuadrados es un modelo válido para clasificar la funcionalidad cerebral de los sujetos en este trabajo de obtención de grado y que, por lo tanto, los resultados de la clasificación de funcionalidad cerebral de dichos sujetos, es válida y confiable.

5 Conclusiones y trabajo futuro

5.1 Conclusiones

En el campo de las neurociencias, el estudio del funcionamiento cerebral constituye uno de los aspectos claves para poder conocer y entender cuáles áreas cerebrales participan en la inducción y modulación de las diferentes conductas, estados emocionales, procesos cognitivos y patologías que presentan los individuos. A lo largo de la historia, diferentes técnicas han sido desarrolladas para evaluar dicha funcionalidad cerebral. Una de ellas es la actividad EEG, la cual ha sido usada desde hace 100 años para estudiar toda una gama de aspectos, desde la caracterización funcional asociada a ejecución motora, hasta las bases neurales de las emociones, de procesos cognitivos como la memoria y distintas patologías [50]. En humanos, el EEG es registrado en base al sistema internacional 10-20 descrito por Jasper en 1958. Este sistema implica la colocación de electrodos en 19 áreas o derivaciones del cuero cabelludo, correspondientes a: 7 subregiones frontales (Fp1, Fp2, F2, F3, Fz, F7 y F8), 3 subregiones parietales (P3, P4 y Pz), 4 subregiones temporales (T3, T4, T5 y T6), y 2 subregiones occipitales (O1 y O2), así como 3 centrales (C3, C4 y Cz). El EEG de estas derivaciones está a su vez dividido en 5 bandas básicas: delta, theta, alfa, beta y gamma, aunque algunos investigadores subdividen alfa en alfa1 y alfa2, así como beta en beta1 y beta2. Si de estas bandas EEG se calculan los parámetros de PA y PR entonces tenemos una cantidad de 133 valores o datos (19 derivaciones x 7 bandas EEG). Si además calculamos las correlaciones de las derivaciones del hemisferio derecho e izquierdo, y si aunado a ello, se requiere comparar la actividad EEG de diferentes condiciones experimentales, entonces la cantidad de datos aumenta exponencialmente. Es por ello que, en la investigación neurocientífica se recurre a métodos de optimización de datos que permitan elegir variables para: A) caracterizar el funcionamiento cerebral asociado a cierto estado fisiológico, tarea o condición patológica, B) establecer métodos de predicción y C) establecer métodos de clasificación en base al funcionamiento cerebral. El amplio desarrollo de métodos de manejo de datos ha abierto la posibilidad de analizar y extraer información a

partir de una gran cantidad de variables como es el caso de los datos electroencefalográficos. Los datos para la realización de este trabajo recepcional, fueron obtenidos de la colaboración académica con el Laboratorio de Neurofisiología de la Conducta Reproductiva. En base a la literatura, se sabe que en la condición de vigilia en reposo ojos abiertos hay una prevalencia de actividad rápida de bajo voltaje que va de los 14 a los 30 Hz, denominada banda beta, que la funcionalidad cerebral presenta una lateralización hemisférica y que esta lateralidad presenta diferencias sexuales, así, la hipótesis que se decidió evaluar en este trabajo fue demostrar que el sexo de una persona no determina su funcionalidad cerebral. A tres parámetros EEG en el rango de beta (PA, PR y rEEG) se aplicaron las pruebas de regresión por mínimos cuadrados, de regresión logística y de K- medias. Los datos de PR mostraron una adecuada distribución normal y, usando todas las variables (estuvieran correlacionadas o no), se encontró que fue la prueba de regresión por mínimos cuadrados la que proporcionó los resultados más acordes a las hipótesis propuestas y que coinciden mejor con los antecedentes neurofisiológicos de lateralidad funcional y diferencias sexuales. Ya que el 67.7 por ciento de los sujetos masculinos se clasificaron con funcionalidad tipo A y 70.3 por ciento de los sujetos femeninos se clasificaron con funcionalidad tipo B, se demuestra la hipótesis de que el sexo de una persona no determina su funcionalidad cerebral en un estado basal de ojos abiertos.

5.2 Trabajo futuro

¿Qué uso práctico se le va a dar al modelo seleccionado en el instituto?

En el laboratorio de Neurofisiología de la Conducta Reproductiva se realizan estudios de funcionalidad cerebral tanto en animales como en humanos, bajo diferentes condiciones experimentales y estados fisiológicos, lo que resulta en la obtención de una vasta cantidad de datos que en numerosas ocasiones obliga a los investigadores a solo analizar una parte de ellos, dejando almacenada información valiosa que tal vez permitiría obtener una visión más completa y exacta de las variables a estudiar y los procesos neurofisiológicos implicados. A la fecha, en ese laboratorio se han limitado a análisis de correlación entre máximo 2 variables EEG, pero con el conocimiento de los resultados de este proyecto, será posible que usen el modelo de regresión de mínimos cuadrados para poder clasificar y caracterizar no solo parámetros de la actividad EEG, sino también analizar datos de diferente índole, tales como datos de ejecución motora, de tareas cognitivas, de pruebas neuropsicológicas, e inclusive, clasificar una combinación de ellos. La aplicación del modelo de regresión por mínimos cuadrados al análisis

de la actividad EEG abrirá nuevas posibilidades de investigación y aplicación en las diferentes líneas experimentales que se realizan en el laboratorio de Neurofisiología de la Conducta Reproductiva.

Bibliografía

- [1] H. Brown and S. Kosslyn, "Cerebral lateralization," *Current Opinion in Neurobiology*, vol. 3, no. 2, 1993.
- [2] R. Davidson and G. Schwartz, "Patterns of cerebral lateralization during cardiac biofeedback versus the self-regulation of emotion: sex differences," *Psychophysiology*, vol. 13, no. 1, pp. 62–68, 1976.
- [3] R. Davidson, "Eeg measures of cerebral asymmetry: conceptual and methodological issues," *International Journal of Neuroscience*, vol. 39, pp. 71–89, 1988.
- [4] R. Davidson, P. Ekman, S. Saron, J. Senulis, and W. Friesen, "Approach-withdrawal and cerebral asymmetry: emotional expression and brain physiology," *Journal of Personality and Social Psychology*, vol. 58, no. 2, pp. 330–341, 1990.
- [5] A. Arnold, "Sexual differences in the brain," *American Scientist*, vol. 68, no. 2, pp. 165–173, 1980.
- [6] E. Bell, M. Willson, A. Wilman, S. Dave, and P. Silverstone, "Males and females differ in brain activation during cognitive tasks," *NeuroImage*, vol. 30, pp. 529–538, 2006.
- [7] K. Evans and E. Hampson, "Sex differences on prefrontally-dependent cognitive tasks," *Brain Cognition*, vol. 93, pp. 42–53, 2015.
- [8] J. McGlone, "Sex differences in human brain asymmetry. a critical survey.," *Behavioral and Brain Sciences*, vol. 3, no. 2, pp. 215–227, 1980.
- [9] M. Guevara and M. Hernández-González, *Registro y análisis automatizado de señales bioeléctricas cerebrales durante la ejecución sexual*. Universidad de Guadalajara, 2006.
- [10] M. Guevara and M. Hernández-González, "Eegmagic: programa para analizar señales electroencefalográficas," *Revista Mexicana de Ingeniería Biomédica*, vol. 30, pp. 41–53, 2009.

- [11] M. Guevara, M. Hernández-González, and A. Sanz-Martín, *Registro y análisis automatizado de señales bioeléctricas cerebrales durante la ejecución sexual*. Universidad de Guadalajara, 2015.
- [12] A. Abdul-latif, I. Cosic, D. Kumar, B. Polus, and C. D. Costa, "Power changes of eeg signals associated with muscle fatigue: the root mean square analysis of eeg bands.," *IEEE*, 2004.
- [13] A. Engel and P. Fries, "Beta band oscillations signalling the status quo?," *Current opinion in neurobiology*, vol. 20, no. 2, pp. 156–165, 2010.
- [14] J. Cantero, M. Atienza, R. Salas, and C. Gomez, "Alpha eeg coherence in different brain states: an electrophysiological index of the arousal level in human subjects," *Neuroscience Letters*, vol. 271, pp. 167–170, 1999.
- [15] M. Hernán-Díaz, F. M. Cid, J. Otárola, R. Rojas, O. Alarcón, and L. Cañete, "Eeg beta band frequency domain evaluation for assessing stress and anxiety in resting, eyes closed, basal conditions," *Procedia computer science*, vol. 162, pp. 974–981, 2019.
- [16] G. Piantoni, K. Kline, and D. Eagleman, "Beta oscillations correlate with the probability of perceiving rivalrous visual stimuli," *Journal of vision*, 10(13), 18, vol. 10, no. 13, p. 18, 2010.
- [17] A. Wróbel, "Beta activity : A carrier for visual attention," *Acta neurobiologiae experimentalis*, 60(2), 2, vol. 60, no. 2, p. 2, 2000.
- [18] S. Shirzadi, M. Dadgostar, Z. Einalou, S. Erdogan, and A. Akin, "Sex based differences in functional connectivity during a working memory task: an fnirs study," *Frontiers Psychology*, vol. 15, 2024.
- [19] H. Rupp and K. Wallen, "Sex differences in response to visual sexual stimuli: A review," *Archives of sexual behavior*, vol. 37, pp. 206–218, 2008.
- [20] Y. Nir, R. Mukamel, I. Dinstein, E. Privman, M. Harel, L. Fisch, H. Gelbard-Sagiv, S. Kipervasser, F. Andelman, and M. Neufeld, "Interhemispheric correlations of slow spontaneous neuronal fluctuations revealed in human sensory cortex," *Natural Neuroscience*, vol. 11, no. 9, p. 8, 2008.
- [21] A. Shmuel and D. Leopold, "Neuronal correlates of spontaneous fluctuations in fmri signals in monkey visual cortex: implications for functional connectivity at rest," *Human Brain Mapping*, vol. 29, no. 7, pp. 751–761, 2008.

- [22] A. Aertsen, G. Gerstein, M. Habib, and G. Palm, "Dynamics of neuronal firing correlation: modulation of "effective connectivity"," *Journal of Neurophysiology*, vol. 61, no. 5, pp. 900–917, 1989.
- [23] M. Greicius, "Resting-state functional connectivity in neuropsychiatric disorders," *Current Opinion Neurology*, vol. 21, no. 4, pp. 424–430, 2008.
- [24] K. Friston, C. Frith, P. Liddle, and R. Frackowiak, "Functional connectivity: the principal-component analysis of large (pet) data sets," *Journal of Cerebral Blood Flow Metabolism*, vol. 13, no. 1, pp. 5–14, 1993.
- [25] H. H. Pol, H. Schnack, R. Mandl, W. Cahn, D. Collins, A. Evans, and R. Kahn, "Focal white matter density changes in schizophrenia: reduced inter-hemispheric connectivity," *Neuroimage*, vol. 21, no. 1, pp. 27–35, 2004.
- [26] J. Gómez-Ramirez, S. Freedman, D. Mateos, J. Pérez-Velazquez, and T. Valiente, "Eyes closed or eyes open? exploring the alpha desynchronization hypothesis in resting state functional connectivity networks with intracranial eeg," *bioRxiv*, pp. 1–27, 2017.
- [27] D. Gusnard, M. Raichle, and M. Raichle, "Searching for a baseline: functional imaging and the resting human brain," *Natural Review Neuroscience*, vol. 2, no. 10, pp. 685–694, 2001.
- [28] C. Honey, O. Sporns, L. Cammoun, X. Gigandet, J. Thiran, R. Meuli, and P. Hagmann, "Predicting human resting-state functional connectivity from structural connectivity," *Proceeding Natural Academic of Science*, vol. 106, no. 6, pp. 2035–2040, 2009.
- [29] N. Jausovec and K. Jausovec, "Resting brain activity: differences between genders," *Neuropsychologia*, vol. 48, no. 13, pp. 3918–3925, 2010.
- [30] S. Cacioppo, F. Bianchi-Demicheli, D. d. P. Bischof, C. Michel, and T. Landis, "Hemispheric specialization varies with eeg brain resting states and phase of menstrual cycle," *PLoS ONE*, vol. 8, no. 4, 2013.
- [31] S. Liu, X. Liu, D. Yan, S. Chen, Y. Liu, X. Hao, W. Ou, Z. Huang, F. Su, F. He, and D. Ming, "Alterations in patients with first-episode depression in the eyes-open and eyes-closed conditions: A resting-state eeg study," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 30, pp. 1019–1029, 2022.

- [32] D. Mulc, J. Vukojevic, E. Kalafatic, M. Cifrek, D. Vidovic, and A. Jovic, "Opportunities and challenges for clinical practice in detecting depression using eeg and machine learning," *Sensors*, vol. 1, 2024.
- [33] R. Buckner and J. Vincent, "Unrest at rest: default activity and spontaneous network correlations," *Neuroimage*, vol. 37, no. 4, pp. 1091–1096, 2007.
- [34] R. Buckner, J. Andrews-Hanna, and D. Schacter, "The brain's default network: anatomy, function, and relevance to disease," *Annual New York Academy of Science*, vol. 1124, pp. 1–38, 2008.
- [35] D. Kan, P. Croarkin, C. Phang, and P. Lee, "Eeg differences between eyes-closed and eyes-open conditions at the resting stage for euthymic participants," *Neurophysiology*, vol. 49, no. 6, pp. 432–440, 2017.
- [36] M. Guevara and M. Corsi-Cabrera, "Eeg coherence or eeg correlation?," *International Journal of Psychophysiology*, vol. 23, no. 3, pp. 145–153, 1996.
- [37] A. Shakeel, T. Tanaka, and K. Kitajo, "Time-series prediction of the oscillatory phase of eeg signals using the last mean square algorithm- based ar model," *Applied Sciences*, 10, 3616, vol. 10, p. 3616, 2020.
- [38] R. Anderson and M. Sandsten, "Classification of eeg signals based on mean-square error-optimal time-frequency features," *26th European signal processing conference*, 2018.
- [39] D. Gujarati, *Econometría Básica*. McGraw Hill, 1981.
- [40] G. Maddala, *Econometría*. McGraw Hill, 1985.
- [41] D. Montgomery and E. Peck, *Introduction to Linear Analysis*. John Wiley and Sons, Inc., 1982.
- [42] H. Romesburg, *Cluster Analysis for researchers*. Lulu.com, 2024.
- [43] M. Albuquerque, *Estabilidad en el análisis de conglomerados: Análisis de conglomerados y su aplicación a las ciencias forestales (Spanish Edition)*. Ediciones Nuestro Conocimiento, 2023.
- [44] M. Modarres, D. Cochran, D. Kennedy, and J. Frazier, "Comparison of comprehensive quantitative eeg metrics between typically developing boys and girls in resting state eyes-open and eyes-closed conditions," *Frontiers Human Neuroscience*, vol. 17, 2023.

- [45] P. Ramyarangsi, S. Bennett, V. Siripornpanich, A. Nanbancha, A. Pokaisasawan, W. Chatthong, and A. Ajjimaporn, "Eeg differences in competitive female gymnastics, soccer, and esports athletes between resting states with eyes closed and open," *Scientific Reports*, vol. 14, 2024.
- [46] Y. Wada, Y. Takizawa, J. Zheng-Yan, and N. Yarnaguchl, "Gender differences in quantitative eeg at rest and during photic stimulation in normal young adults," *Clinical Electroencephalography*, vol. 25, no. 2, pp. 81–85, 1994.
- [47] R. Barry, A. Clarke, R. McCarthy, M. Selikowitz, J. Rushby, and E. Ploskova, "Eeg differences in children as a function of resting-state arousal level," *Clinical Neurophysiology*, vol. 115, pp. 402–408, 2004.
- [48] R. Barry and F. D. Blasio, "Eeg differences between eyes-closed and eyes-open resting remain in healthy ageing," *Biological Psychology*, vol. 129, pp. 293–304, 2017.
- [49] J. Fox, *Applied Regression Analysis and Generalized Linear Models*. Sage Publications, 2015.
- [50] F. Mushtaq, D. Welke, A. Gallagher, G. Pavlov, L. Kouara, J. Bosch-Bayard, and P. Valdes-Sosa, "One hundred years of eeg for brain and behaviour research," *Nature human behavior*, vol. 18, no. 8, pp. 1437–1443, 2024.