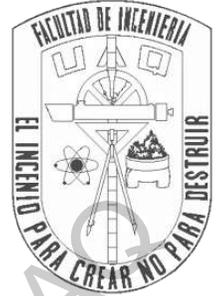




UNIVERSIDAD AUTÓNOMA DE QUERÉTARO

FACULTAD DE INGENIERÍA CAMPUS AEROPUERTO



Metodología para la identificación de predisposición al alcoholismo en personas mediante herramientas tiempo-frecuencia y machine learning usando señales EEG

TESIS

Como parte de los requisitos para obtener el título de
INGENIERO BIOMÉDICO

Línea terminal

BIOSEÑALES Y BIOINSTRUMENTACIÓN

Presenta

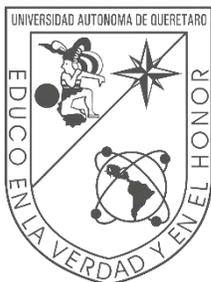
Luis Antonio Franco Vergara

Directores

Dr. Carlos Andrés Pérez Ramírez

Dr. Juan Pablo Amézquita Sánchez

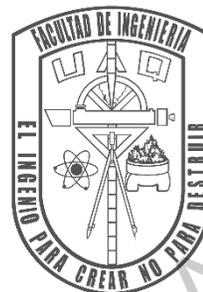
Santiago de Querétaro, Querétaro, octubre de 2021.



Universidad Autónoma de Querétaro

Facultad de Ingeniería

Ingeniería Biomédica



“Metodología para la identificación de predisposición al alcoholismo en personas mediante herramientas tiempo-frecuencia y machine learning usando señales EEG.”

Tesis

Que como parte de los requisitos para obtener el título de Ingeniero Biomédico.

Presenta:

Luis Antonio Franco Vergara

Dirigido por:

Dr. Carlos Andrés Pérez Ramírez
Dr. Juan Pablo Amézquita Sánchez

Dr. Carlos Andrés Pérez Ramírez

Presidente

Dr. Juan Pablo Amézquita Sánchez

Secretario

Dr. Martín Valtierra Rodríguez

Vocal

M.C. Arturo Hernández Medina

Suplente

Firma

Firma

Firma

Firma

Centro Universitario
Querétaro, Qro.
Octubre, 2021

RESUMEN

El alcoholismo es un problema actual de salud en nuestro país y en el mundo, provocando consecuencias fatales en aquellos que la padecen. Las herramientas vigentes de diagnóstico no son confiables para detectar efectivamente alcoholismo en todas las situaciones; por ello, se propone una metodología novedosa para la detección objetiva de personas con riesgo a sufrir alcoholismo a través del procesamiento de señales de EEG. Ésta consiste en utilizar la transformada wavelet discreta para obtener diferentes sub-bandas de frecuencia hasta el nivel seis de descomposición, permitiendo la medición del cambio geométrico de las bioseñales por medio de cinco algoritmos diferentes de dimensión fractal: Higuchi, Petrosian, Dimensión de Caja, Sevcik y Katz; generando así posibles características para el entrenamiento de modelos de aprendizaje automático de máquinas. Los datos de mayor significancia para el fenómeno estudiado fueron seleccionados por medio del método estadístico no paramétrico ANOVA de Kruskal-Wallis; con estas características se entrenaron y validaron dos modelos de machine learning: la máquina de vectores de soporte y el perceptrón multicapa. De acuerdo con la validación para detectar de alcoholismo en señales de la misma naturaleza, se logró conseguir una exactitud, precisión, sensibilidad y especificidad de 96.52%, 98.55%, 94.44% y 98.61 respectivamente con la máquina de vectores de soporte.

Palabras clave: procesamiento de bioseñales, EEG, alcoholismo, DWT, dimensión fractal, machine learning.

DEDICATORIA

Esta tesis la dedico principalmente a mis padres Sergio y Claudia, por ser el mayor apoyo en mi trayecto de vida y más grande ejemplo a seguir. Les dedico mi trabajo de tesis y futuros logros, así como ellos me han dedicado gran parte de su vida.

A mis hermanos, por ser mis más cercanos y fieles compañeros.

A mi familia materna, porque sin ellos no habría sido posible culminar mis estudios y ni llegar la persona que soy hoy.

A mi familia de La Paz BCS, porque a pesar de la distancia física siempre los siento presentes y cercanos a mí.

A mi novia Valeria, por iluminarme mis días y motivarme constantemente.

A mi director de tesis el Dr. Carlos Pérez, que gracias a su generosa ayuda y espléndido conocimiento fue posible realizar el trabajo de tesis.

A mis más cercanos amigos, por encender mi vida y su brindarme su gratificante compañía.

A la Universidad Autónoma de Querétaro, por la formación académica proporcionada, por permitirme disfrutar del basquetbol universitario competitivo y por enriquecer mi mente con tantas enseñanzas.

AGRADECIMIENTOS

Mi mayor agradecimiento a Dios, por darme la fortaleza y entusiasmo para cumplir esta meta. También por poner en mi camino a las personas correctas que sin su ayuda la conclusión de mi carrera profesional no hubiera sido posible.

De igual manera, le agradezco intensamente a mis padres, no solo por ser mi mayor inspiración, sino también por inculcar en mí el esfuerzo y perseverancia para culminar mis metas, por sus enseñanzas y consejos en cada paso de mi caminar, por ilustrarme con el ejemplo a enfrentar la adversidad con valentía porque Dios está siempre conmigo.

A mi tía Elsa y mi tío Toño, por extenderme su mano incondicionalmente cada vez que los tiempos se vuelven turbulentos, por alentarme a ser siempre una mejor persona y por acompañarme en mis sueños y aspiraciones.

A mis abuelos Elena y Félix, por siempre recibirme y acogerme afectuosamente con los brazos abiertos, escuchándome y guiándome con su gran sabiduría, cuidándome en cada instante.

A mi tío Félix, por su generosa y espléndida compañía en las joviales experiencias que hemos vivido y vamos a seguir teniendo.

A mis hermanos Sergio y Mariana, por ser mis primeros y más importantes leales amigos, con quienes he compartido y aprendido tanto. Los llevo siempre en mi corazón.

A mi novia Valeria, por apoyarme y alentarme completamente, porque su cálido y desmedido afecto me aviva, me motiva a seguir luchando por mis sueños.

A mi mejor amigo Diego, por impulsarme a ser siempre la mejor versión de mismo y ser mi más querido confidente dentro y fuera de la cancha de baloncesto.

A mis amigos de la universidad, por acompañarme y apoyarme cada semestre de la carrera, por las experiencias y momentos compartidos y, sobre todo, por ser parte fundamental de mi crecimiento como estudiante.

A mis amigos del equipo de baloncesto y mi coach el Dr. Arturo Sánchez, por darme un segundo hogar y una segunda familia, por enseñarme el trabajo en equipo y a ayudarme a forjar carácter.

A mis sinodales el Dr. Juan Pablo Amézquita, el Dr. Martín Valtierra y el M.C. Arturo Medina por sus valiosas observaciones, provechosos consejos y grandes enseñanzas en este proyecto de tesis.

ÍNDICE DE TABLAS

Tabla 2.1. Manifestaciones neurológicas secundarias al consumo excesivo del alcohol (Planas-Ballvé et al., 2017).	17
Tabla 2.2. Resumen de los Kernels más comunes que satisfacen el teorema de Mercer (S. Haykin, 2008).	42
Tabla 3.1. Similitud y número de prueba entre pacientes considerados para el trabajo	48
Tabla 3.2. Sub-bandas de EEG correspondientes a su descomposición tras Transformada Wavelet.	49
Tabla 4.1. Valores de Fractalidad seleccionados tras la prueba estadística no paramétrica ANOVA de Kursk-Wallis.	61
Tabla 4.2. p-values de cada uno de los algoritmos de fractalidad con mayor importancia para el método de clasificación, considerando electrodo y descomposición.	62
Tabla 4.3. Principales diagramas de cajas y bigotes correspondientes a los mejores p-values.	63
Tabla 4.4. Desempeño del Perceptrón Multicapa.	65
Tabla 4.5. Desempeño de la Máquina de Vectores de Soporte	66
Tabla 4.6. Tabla comparativa de Metodologías.	67

ÍNDICE DE FIGURAS

Figura 2.1. Estructura básica de una motoneurona.....	14
Figura 2.2. Señales de EEG.....	19
Figura 2.3. Representación básica de dónde proviene el EEG y su registro	20
Figura 2.4. Vista sagital izquierda de las marcas sobre el cráneo.....	23
Figura 2.5. Nomenclatura del sistema 10-20.....	24
Figura 2.6. Algunas de las funciones de ondoleta utilizadas en la transformada wavelet	26
Figura 2.7. Comparativa entre DWT y WPD o WPT.....	27
Figura 2.8. Procedimiento de descomposición de una señal en sus diferentes sub-bandas por medio de una transformada wavelet discreta.	28
Figura 2.9. Fractales matemáticos ideales.....	29
Figura 2.10. Representación gráfica del valor p y diagramas de cajas.	35
Figura 2.11. Representación gráfica de una red neuronal	38
Figura 2.12. Ilustración de un hiperplano óptimo que establece el límite entre datos linealmente separables	39
Figura 2.13. Interpretación geométrica de las distancias algebraicas de los datos hacia el hiperplano óptimo.....	40
Figura 3.1. Metodología.	44
Figura 3.2. Representación de las imágenes de Snodgrass y Vanderwart.....	46
Figura 3.3. Representación muestra del contenido de cada archivo.....	47
Figura 3.4. Representación gráfica de la descomposición en Aproximaciones (A) y Detalles (D)	50
Figura 3.5. Representación gráfica del acomodo de los datos.....	51
Figura 3.6. Diagrama de cajas y tabla de resultados obtenida tras el análisis estadístico no paramétrico Kruskal-Wallis.....	53
Figura 3.7. Acomodo de los datos para entrenar los modelos de aprendizaje de máquinas	54
Figura 3.8. A) Matriz de confusión.....	57
Figura 4.1. Representación gráfica de la descomposición de las señales de EEG60	

ÍNDICE GENERAL

CAPÍTULO 1. INTRODUCCIÓN	1
1.1. Introducción.....	2
1.2. Antecedentes	3
1.3. Descripción del problema.....	7
1.4. Justificación.....	9
1.5. Hipótesis y Objetivos.....	9
1.5.1. Hipótesis.....	9
1.5.2. Objetivo General.....	10
1.5.3. Objetivos Particulares.....	10
1.6. Planteamiento General.....	11
CAPÍTULO 2. FUNDAMENTACIÓN TEÓRICA.....	12
2.1. Alcoholismo y Cerebro	13
2.1.1. El Cerebro y el Sistema Nervioso	13
2.1.2. Principales Niveles Funcionales del Sistema Nervioso Central.....	15
2.1.3. Efectos del Alcohol en el Cerebro y Sistema Nervioso.....	16
2.2. Predisposición Genética al Alcoholismo.....	17
2.3. Señales de EEG.....	19
2.3.1. Adquisición de señales de EEG.....	21
2.3.2. Electrodo.....	22
2.3.3. Sistema Internacional de Colocación de Electrodo 10-20.....	22
2.3.4. Nomenclatura del sistema 10-20	23
2.4. Transformada Wavelet Discreta	25
2.5. Fractales	29
2.5.1. Algoritmo de Dimensión Fractal de Katz.....	30
2.5.2. Algoritmo de Dimensión Fractal de Higuchi.....	31
2.5.3. Algoritmo de Dimensión Fractal de Recuento de Cajas	32
2.5.4. Algoritmo de Dimensión Fractal de Sevcik	32
2.5.5. Algoritmo de Dimensión Fractal de Petrosian.....	33
2.6. ANOVA No Paramétrico.....	33
2.7. Clasificadores Basados en el Aprendizaje De Máquina.....	35
2.7.1. Deep Neural Networks.....	37

2.7.2. Máquinas de Vectores de Soporte.....	39
CAPÍTULO 3. METODOLOGÍA.....	43
3.1. Metodología Propuesta	44
3.2. ETAPA I: Lectura	45
3.2.1. Base de Datos	45
3.2.2. Selección de Sujetos	47
3.3. ETAPA II: Preprocesamiento	49
3.4. ETAPA III: Procesamiento.....	50
3.5. ETAPA IV: Método Estadístico No Paramétrico	52
3.6. ETAPA V: Clasificación	54
3.6.1. Perceptrón Multicapa.....	54
3.6.2. Máquina de Vectores de Soporte	55
3.7. Evaluación diagnóstica de los modelos.....	56
CAPÍTULO 4. EXPERIMENTACIÓN Y RESULTADOS	58
4.1. Resultados Obtenidos	59
4.2. Discusión de Resultados.....	67
CAPÍTULO 5. CONCLUSIONES Y PROSPECTIVAS.....	71
5.1. Conclusiones.....	72
5.2. Trabajos a Futuro	73
BIBLIOGRAFÍA	74

Dirección General de Bibliotecas UAQ

CAPÍTULO 1. INTRODUCCIÓN

1.1. Introducción

El incremento del consumo de alcohol en la población mundial no es reciente, pues ha formado parte de nuestras costumbres y tradiciones desde hace varios años; a pesar de su normalización, el consumo irresponsable sigue siendo un problema actual (Infocop, 2018). El alcoholismo es una enfermedad que se presenta alrededor a nivel mundial, pues de acuerdo con la OMS (*Organización Mundial de la Salud*), 3 millones de muertes al año son consecuencia del consumo excesivo de alcohol, lo que representa un 5.3 % de las defunciones totales. En el año 2016, el consumo total de alcohol por habitante en la población mundial de más de 15 años se mantuvo en 6,4 litros (World Health Organization, 2018)

México no se queda atrás, pues en el país se registró que el 33.6% de la población representativa consume alcohol de manera excesiva (Villatoro-Velázquez et al., 2017). El consumo de esta sustancia causa discapacidad y muerte entre los grupos poblacionales más jóvenes, siendo la principal causa de accidentes de tránsito (Ahumada-Cortez et al., 2017). Desafortunadamente, México ocupa el séptimo lugar en el mundo en la lista de países que miden la cantidad de muertos por accidentes de tránsito y la segunda posición en América Latina, según la tasa de fallecidos por cada 100,000 habitantes, que es del 24% (CEVECE, 2012).

El alcoholismo es una enfermedad primaria y crónica, en la que se involucran factores genéticos, psicosociales y ambientales para su desarrollo (Kershenobich, 2010). El consumir alcohol de manera excesiva puede perjudicar la salud de diferentes maneras; a pesar de no existir síntomas precisos, hay cambios en el comportamiento de la persona. Quien padece de ella suele necesitar beber alcohol en cualquier momento y sin control, por lo que puede realizar cualquier actividad a cambio de conseguir alcohol, además de mostrar desinterés por su salud y aspecto físico (IMSS, 2015).

Dimensionando la problemática social que conlleva el alcoholismo, es necesario poder diagnosticarlo eficazmente. Actualmente, su detección se realiza

por medio de cuestionarios como CAGE¹ (*Cut down – Annoyed – Guilty – Early morning drink*, en inglés), CBA (*Cuestionario Breve de Alcoholismo*) y AUDIT² (*Alcohol Use Disorders Identification Test*, en inglés) (Sáiz et al., 2002). A pesar de ser instrumentos útiles, los médicos filtran menos de la mitad de sus pacientes (Fleming, 1997). Además, los pacientes pueden retener información por vergüenza o estigmatización, lo que puede provocar que se les diagnostique equivocadamente, que demoren en tomar un tratamiento oportuno o que tomen las medidas preventivas necesarias. Por ello, es necesario desarrollar una metodología eficaz para la detección objetiva de alcoholismo, misma que empleando el procesamiento de señales y técnicas de aprendizaje de máquina avanzadas, se puede cubrir dicha necesidad.

1.2. Antecedentes

Durante varios años, investigadores alrededor del mundo han propuesto diversas metodologías para detectar alcoholismo a través del procesamiento de señales de Electroencefalografía (*EEG*), ya que la obtención de estas señales biomédicas es no invasiva y es una manera vigente de representar y estudiar la actividad cerebral. Inicialmente a nivel internacional, las características buscadas en este tipo de señales eran la potencia de espectro o PSD (*Power Spectrum Density*, en inglés) en sujetos control y alcohólicos. Los trabajos que a continuación se citan utilizan la misma base de datos pública, que posee tres versiones con diferente número de sujetos control y alcohólicos: pequeña (2-2), grande (10-10) y la base de datos completa (120 pruebas). Por un lado, Sun et al. utilizaron la DWT (*Discrete Wavelet Transform*, en inglés) y FFT (*Fast Fourier Transform*, en inglés) para pre-procesar y estimar la potencia de espectro de cada banda seleccionada de ambas clases de señales, determinando que la potencia perteneciente a la banda teta y delta del cerebro disminuyen considerablemente en personas alcohólicas (Sun et al., 2006). Faust y colaboradores también estimaron el espectro de potencia de las señales por medio de tres métodos: Welch, Tule-Waker y Burg, en señales

¹ Cuestionario de alcoholismo, sus siglas corresponden a “Reducir – Molesto – Culpable – Beber por la mañana” en español.

² Cuestionario de Identificación de los Trastornos debidos al Consumo de Alcohol en español.

de sujetos control, alcohólicos y epilépticos. Basándose en los picos de potencia y asociándolos a su frecuencia correspondiente obtuvieron tres parámetros el primero representa la media y varianza del pico de potencia, el segundo simboliza la media y la varianza de la frecuencia donde se encuentra el pico y el tercero corresponde al pico dividido por su frecuencia correspondiente, teniendo en total nueve características diferentes, las cuales eran distintas en cada clase de señal. Con el análisis de la Curva ROC (*Receiver Operating Characteristic*, en inglés) concluyeron que el método de Burg proporciona los parámetros más distinguibles (O. Faust et al., 2008).

Con el desarrollo de las técnicas de aprendizaje de máquina (*machine learning*, en inglés), diversas estrategias de procesamiento de señales han sido propuestas. Faust y sus colaboradores emplearon un clasificador Bayesiano, utilizando como características de las señales la energía y energía relativa, obteniendo una exactitud de 87.9%. Para ello descompusieron la señal con WPT (*Wavelet Packet Transform*, en inglés) hasta el nivel 3. Calcularon la energía de cada nivel sumando la energía de cada aproximación y detalle. La energía relativa de cada nivel se calculó dividiendo la energía del nivel entre la energía total. Con una validación cruzada de 10 iteraciones, se entrenó el clasificador Bayesiano (Oliver Faust et al., 2013). Rachman y colaboradores propusieron una nueva metodología que emplea ICA (*Independent Component Analysis*, en inglés) para remover el ruido de parpadeos y sonido, y SWT (*Stationary Wavelet Transform*, en inglés) para obtener las bandas de frecuencia del nivel 3 al 6. A estas descomposiciones se les calculó el máximo, mínimo, promedio y desviación como características para ser ingresadas a una PNN (*Probabilistic Neural Network*, en inglés). El ancho de la función de activación de la PNN se fijó en 0.6. La exactitud obtenida de dicho método fue del 85% (Rachman et al., 2017). Bae y colaboradores presentan un método novedoso, al utilizar redes efectivas con la causalidad de Granger entre los canales de EEG. A partir de esto calcularon el coeficiente de agrupamiento, asertividad, grado medio de vecindad y centralidad de intermediación de nodo. Seleccionaron las características más aptas con el “Algoritmo de Alivio” (*Relief Algorithm*, en inglés) y entrenaron con estos datos una máquina de soporte

vectorial como clasificador con cuatro validaciones cruzadas, logrando una exactitud del 90% (Bae et al., 2017).

En este mismo tenor, Taran y Bajaj plantean la combinación de características probabilísticas (diferencia media absoluta, rango de intercuartil, y covarianza) y no lineales (entropías) de las componentes de la señal EEG obtenidos usando la EMD (*Empirical Mode Decomposition*, en inglés), clasificando entre clases con LS-SVM (*Least-Square Support Vector Machine*, en inglés), alcanzando una exactitud del 93.75%. A partir de los resultados, los autores concluyen que los ritmos de señal EEG de sujetos alcohólicos son menos complejos en comparación con los ritmos de señal EEG de sujetos normales, lo que demuestra que el alcohol actúa como depresor del sistema nervioso central y ralentiza su proceso neuronal (Taran & Bajaj, 2018). En los siguientes años, Rodrigues y colaboradores utilizaron la transformada wavelet con cinco niveles de descomposición, empleando diferentes familias de ondoletas (Biortogonal, Coiflet, Daubechies y Symlets) para extraer características estadísticas de las descomposiciones (mínimo, máximo, potencia, media, desviación estándar, media absoluta, razón de la media absoluta). Experimentaron con la base de datos de 10 sujetos control y 10 sujetos alcohólicos con diferentes algoritmos de aprendizaje de máquinas. Se emplearon todas las subbandas resultantes de la descomposición y los porcentajes del total de la base de datos para el entrenamiento-prueba fueron: 25-75, 50-50 y 75-25. Los resultados mostraron valores de exactitud 97.93% y 88% utilizando máquina de vectores de soporte y perceptrón multicapa respectivamente, implementando la familia de ondoletas db4 (Rodrigues et al., 2019). Anuragi y Sisodiades componen las señales en 10 componentes con la EWT (*Empirical Wavelet Transform*, por sus siglas en inglés), extrayendo de cada una las siguientes características: media, desviación estándar, varianza, asimetría, curtosis, entropía de Shannon y la entropía logarítmica; las últimas dos para caracterizar con precisión la ubicación y variabilidad de las señales de EEG. Empleando la prueba *t-student* seleccionaron las características más significativas, quedando así la media, desviación estándar, varianza y curtosis. La metodología se probó con la base de datos que comprende dos sujetos, un alcohólico y un sujeto control. Con los clasificadores de Bayes y una

máquina de soporte vectorial, se obtienen exactitudes del 90.46% y 85.83%, respectivamente (Anuragi & Sisodia, 2020).

A nivel nacional, se encontró escasa literatura relevante para la presente investigación. Juan Ramón de la Fuente y David Kershenobich en 1987 son los únicos que alertan de la necesidad de instrumentos objetivos para la detección de alcoholismo. Revisan algunos aspectos conceptuales y metodológicos en relación a la detección y el diagnóstico de sujetos que beben alcohol de forma excesiva. Presentan una problemática en el diagnóstico temprano de alcoholismo, pues estipulan que no es una tarea fácil con las herramientas que en su momento se empleaban, las cuales son test y cuestionarios que actualmente se usan; de igual manera, los sujetos que consumen alcohol de forma excesiva no son confiables cuando se trata de averiguar las cantidades de alcohol que ingieren. Una de las limitaciones a las que se enfrentaban era que un buen número de sujetos con problemas de consumo excesivo de alcohol no eran detectados y por lo tanto no recibían la atención médica requerida. Los médicos generales diagnostican únicamente a uno de cada diez pacientes alcohólicos. Mencionan que emplear instrumentos de tamizaje (screening) podrían identificar el problema con mayor claridad y es necesaria la búsqueda de indicadores más objetivos; para ello proponen marcadores biológicos, sin embargo, es un método invasivo y, por ejemplo, el marcador biológico de volumen corpuscular medio (VCM) tan solo tiene un 50% de efectividad (Fuente & Kershenobich, 1987).

En cuanto nivel local, de acuerdo con el alcance de la investigación realizada, actualmente no hay antecedentes locales de detección de alcoholismo por medio de procesamiento de EEG o por cualquier otro medio. Sin embargo, las señales de EEG sí han sido objeto de estudio de diferentes investigaciones. Por un lado, Cruz (2017) presenta una metodología para identificar automáticamente los estados mentales en la imaginación motora a partir de EEG, utilizando dimensión fractal y EMD con redes neuronales para dicho propósito. Por otro lado, Salazar (2018) identifica los estados mentales para detectar intenciones de movimiento en la imaginación motora combinando entropías y EMD con redes neuronales;

seleccionando las IMF (*Intrinsic Mode Function*, en inglés) 5 y 7 tras la descomposición, logró una eficiencia de 70.58%. De igual manera, Pérez (2018) propone una metodología para la detección anticipada de eventos epilépticos en EEG utilizando WPT y los siguientes fractales: dimensión de Katz, dimensión de Higuchi y Dimensión de Caja. Por medio de ANOVA se determinó que el nivel 5, el nodo correspondiente al ancho de banda entre 10 y 12 Hz, era el más discriminante. Con una red neuronal logra una eficiencia de 95%. Relacionado a sistemas de diagnóstico, Piña (2019) expone una metodología para la predicción de infartos utilizando señales de electrocardiografía. Combinando las dimensiones fractales de Higuchi, de Caja, de Katz y de Sevcik con lógica difusa logra una exactitud del 91.58% para detectar eventos cardiacos previos a los infartos.

A partir de la información presentada, se concluye que debido a que las señales de EEG son no estacionarias y presentan caos (Mohanchandra et al., 2016), deben utilizarse técnicas de descomposición avanzadas en el dominio tiempo-frecuencia (Sharma & Pachori, 2017), ya que éstas están diseñadas para trabajar con señales con características como las EEG. También es importante mencionar que la exactitud de las metodologías propuestas es un área de oportunidad, ya que al integrar diversas estimaciones para medir los cambios geométricos en señales caóticas puede resultar en una mejora de la exactitud de la metodología resultante. Así mismo, se puede afirmar que es necesario ampliar los recursos en esta línea de investigación en todos los niveles: internacional, nacional y local. Por esta razón, desarrollar una metodología basada en el procesamiento de la señal EEG, medición de los cambios geométricos de la señal junto con técnicas de aprendizaje de máquina, proporcionaría un método de detección objetiva de alcoholismo; no solo siendo el primer trabajo a nivel nacional que ampliaría el conocimiento en esta área, sino también sentando las bases para futuras investigaciones.

1.3. Descripción del problema

El abuso del alcohol provoca cerca de tres millones de muertes al año a nivel mundial, lo que representa el 5.3% de las defunciones anuales totales (World Health

Organization, 2018). México ocupa el séptimo lugar en accidentes automovilísticos a nivel mundial, donde la mayor parte de incidencia de accidentes en jóvenes está relacionado al consumo de alcohol (CEVECE, 2012), ya que los porcentajes de consumo excesivo aumentaron 5.6 % en 5 años (2011-2016) (Villatoro-Velázquez et al., 2017). Dado que solo 1 de cada 10 pacientes alcohólicos logran ser diagnosticados por médicos generales (Fuente & Kershenobich, 1987), es necesario desarrollar una herramienta que permita diagnosticar objetivamente el alcoholismo.

Las señales fisiológicas son aquellas que se pueden obtener del cuerpo humano. En particular, las señales de EEG reflejan la dinámica de la actividad eléctrica generada por una gran cantidad de neuronas en el cerebro. Esto genera que se tengan fluctuaciones no periódicas y caóticas (Mohanchandra et al., 2016). Algunos investigadores las han estudiado con medidas de complejidad, pero no suelen incluir más de dos. Thomas sugiere que cada medida de complejidad detecta información única en la señal y por eso no debería limitarse el número de ellas (Burns & Rajan, 2015). Los modelos estocásticos, caóticos y ruidosos y caóticos pueden tratar matemáticamente estos patrones, donde se conceptualizan las fuertes interacciones no lineales entre unos pocos factores dan forma a las fluctuaciones. Entre los enfoques estocásticos, los modelos fractales dan la mejor descripción de la realidad (Eke et al., 2002). En los antecedentes revisados, ninguna metodología ha propuesto utilizar fractales para caracterizar la señal, siendo estos los más adecuados para analizar señales de esta naturaleza y de baja carga computacional.

En este sentido, y tomando en cuenta las consideraciones arriba mencionadas, en esta tesis se plantea el desarrollar una metodología que emplee como herramienta de descomposición la DWT, donde las bandas de frecuencia que se obtienen se analizan con características que puedan cuantificar los cambios geométricos que sufre la señal EEG; para que empleando clasificadores basados en algoritmos de aprendizaje de máquina puedan detectar la predisposición al alcoholismo objetivamente, mejorando la exactitud reportada en trabajos similares,

con el fin de evitar las consecuencias fatales que puede desencadenar la enfermedad.

1.4. Justificación

Dada la gran incidencia que tiene el alcoholismo en el país, en particular en la población joven (CEVECE, 2012), el desarrollar una metodología no invasiva de detección objetiva de alcoholismo asistiría en localizar aquellos que padezcan de esta enfermedad, promoviendo su atención oportuna y para así brindarles el tratamiento adecuado o en su defecto, tomar las medidas preventivas necesarias que, por ende, se vería reflejado en los incidentes por abuso del alcohol.

Por otra parte, la mayoría de las metodologías presentadas en la sección de los antecedentes utilizan algoritmos estadísticos para caracterizar las señales de EEG; sin embargo, como se plantea en la sección anterior, pueden no ser la manera más adecuada de interpretar y caracterizar señales fisiológicas con una naturaleza compleja y caótica. Esto lleva a considerar que al separar la señal EEG en diferentes bandas de frecuencia, acopladas a un algoritmo de cuantificación de cambios geométricos, en particular la dimensión fractal, permitirían estudiar de mejor manera la complejidad de la señal; lo que conllevaría en obtener indicadores efectivos que pueden ser interpretados por una red neuronal o máquina de vectores de soporte, que son algoritmos de aprendizaje de máquina efectivos para generar clasificadores. La elección de una red neuronal se justifica ya que el algoritmo es efectivo para clasificar patrones con cambios mínimos (Neuron et al., 2018) .

1.5. Hipótesis y Objetivos

1.5.1. Hipótesis

Mediante la integración de una herramienta de descomposición de señales tiempo-frecuencia, algoritmos de extracción de patrones no lineales y un clasificador basado en algoritmos de aprendizaje de máquina, se desarrolla una metodología que determine si una persona está en riesgo de padecer alcoholismo o no usando señales EEG, con una exactitud superior a las metodologías existentes en el estado del arte.

1.5.2. Objetivo General

Desarrollar una metodología para la detección de personas en riesgo de sufrir alcoholismo empleando herramientas de descomposición de señales tiempo-frecuencia, algoritmos de extracción de patrones no lineales y un clasificador basado en algoritmos de aprendizaje de máquina usando señales EEG.

1.5.3. Objetivos Particulares

1. Desarrollar un código de lectura de las señales EEG, empleando el software MATLAB, a fin de poder cargarlas y ser procesadas en etapas posteriores.
2. Programar la DWT para descomponer la señal EEG en sus diferentes bandas de frecuencia, a fin de encontrar aquella que tenga la información relevante para poder ser usada en la detección de personas con riesgo de sufrir alcoholismo.
3. Programar los algoritmos de estimación de la dimensión fractal, para poder cuantificar los cambios que se tienen en cada una de las bandas de frecuencia de las señales EEG.
4. Realizar un análisis exhaustivo, empleando la media, varianza y desviación estándar, de las características encontradas, a fin de determinar cuál banda de frecuencia es la que contiene los mayores cambios.
5. Programar algoritmos de aprendizaje de máquinas para poder generar un sistema que permita determinar las personas que están en riesgo de sufrir alcoholismo.
6. Evaluar la eficacia de cada uno de los clasificadores desarrollados, empleando diferentes métricas de rendimiento, para determinar el más exacto.

1.6. Planteamiento General

A continuación, se presenta un planteamiento general de cómo se llevará a cabo la tesis (Figura 1.1), la cual en general comprende de tres etapas distintas:

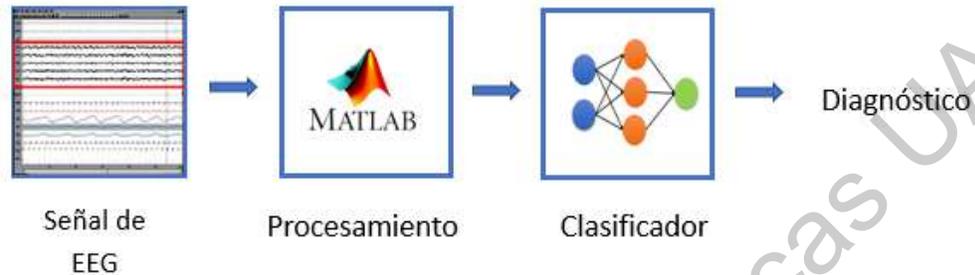


Figura 1.1. Planteamiento General.

- **Lectura y acondicionamiento de la señal EEG:** Es necesario leer las señales eléctricas que se van a estudiar, cuyo origen recae en la actividad cerebral de los sujetos control y sujetos con predisposición a sufrir alcoholismo. Estas señales provienen de una base de datos pública la cual se detalla más adelante en el Capítulo III del presente trabajo (<https://archive.ics.uci.edu/ml/datasets/EEG+Database>).
- **Procesamiento de la señal:** Como siguiente paso, se llevará a cabo el procesamiento digital de las señales EEG, tras su lectura y entendimiento. El objeto de esta etapa es encontrar características propias de cada tipo de señal que nos permitan diferenciar entre ambas. En este sentido, se emplearán algoritmos tiempo-frecuencia avanzados que permitan descomponer las señales EEG ya sea en sus componentes fundamentales o en las bandas de frecuencia determinadas por el Estado del Arte. Las características más significativas se seleccionarán por medio de métodos estadísticos.
- **Clasificador:** Por último, una vez encontradas las características, se utilizarán para validar un clasificador basado en técnicas de aprendizaje de máquina e inteligencia artificial, cuyo propósito es obtener un diagnóstico de forma objetiva y automática.

Dirección General de Bibliotecas UAQ

CAPÍTULO 2. **FUNDAMENTACIÓN TEÓRICA**

En el capítulo anterior se introdujo al tema de investigación, el cual es desarrollar una metodología para la detección objetiva de personas con predisposición genética a sufrir alcoholismo, por medio de herramientas de descomposición de señales tiempo-frecuencia, extracción de patrones no lineales (fractales) y algoritmos de aprendizaje de máquina.

En la primera sección de este capítulo se abordará brevemente la fisiología del sistema nervioso, los efectos nocivos del alcoholismo para la función cerebral, la teoría básica detrás de la electroencefalografía y los métodos de adquisición de estas señales. En seguida, se describirán teóricamente tanto las herramientas utilizadas para la descomposición de las señales como los algoritmos de extracción de patrones no lineales, finalizando con la explicación de los algoritmos de clasificación de aprendizaje de máquinas.

2.1. Alcoholismo y Cerebro

Abusar del alcohol puede traer complicaciones a lo largo del tiempo, causando daño permanente y alteraciones en diferentes sistemas del cuerpo humano, como lo pueden ser el tracto digestivo, el hígado, el corazón, el sistema vascular, el sistema óseo y muscular, el sistema inmunológico, endocrinológico y hematológico y sistema nervioso central y periférico.

La neurodegeneración y atrofia cerebral y/o atrofia cerebelosa puede ser provocada por la ingesta excesiva debida a la dependencia del alcohol, por ello a continuación se aborda a grandes rasgos la función del cerebro y su relación con el sistema nervioso.

2.1.1. El Cerebro y el Sistema Nervioso

Por su complejidad al realizar procesos de pensamiento y acciones de control, no existe sistema artificial capaz de compararse con el sistema nervioso del cuerpo humano (Guyton & Hall, 2016). A través de impulsos nerviosos el sistema nervioso recibe millones de fragmentos de información cada minuto, procedentes de los distintos nervios y órganos sensitivos repartidos en toda la extensión del cuerpo para que el organismo sea capaz de generar una respuesta adecuada. Toda

esta información llega al sistema nervioso central por medio de señales nerviosas generadas a partir de estímulos sensitivos. Esta información es procesada y transmitida por la unidad principal que conforma el tejido nervioso, la neurona. Son células únicas con extensiones alargadas y delgadas, cuyas comunicación con células vecinas es a través de neurotransmisores (Silverthorn et al., 2013). Su estructura básica se puede apreciar en la Figura 2.1.

El sistema nervioso consta del sistema nervioso central (SNC por sus siglas en español), que incluye el encéfalo y médula espinal, y del sistema nervioso periférico (SNP por sus siglas en español), que comprende receptores sensoriales, nervios sensoriales y ganglios (Costanzo, 2011). Cuando se observa desde un nivel macroscópico es posible distinguir dos tipos de tejidos en este sistema: la materia gris y la materia blanca.

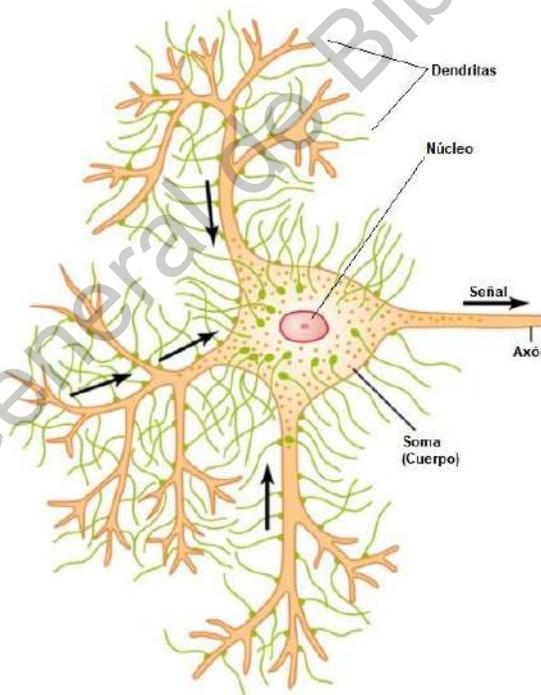


Figura 2.1. Estructura básica de una motoneurona. Se observan sus principales partes que son: las dendritas, el soma y el axón (Guyton & Hall, 2016).

La materia gris consiste en cuerpos de células nerviosas amielínicas, dendritas y terminales axónicas. Los cuerpos celulares se ensamblan de manera organizada en el cerebro y médula espinal formando grupos con funciones

similares. A cada uno de estos grupos se les denomina núcleos y se identifican por nombres específicos; por ejemplo, el núcleo geniculado lateral, encargado de procesar información visual. Por otro lado, a la materia blanca pertenecen pocos cuerpos celulares de axones mielinizados. Las vainas de mielina le dan su color característico pálido. Los grupos de axones que conectan diferentes regiones del SNC son conocidos como tractos y estos son equivalentes tanto en el SNC como en el SNP (Silverthorn et al., 2013).

2.1.2. Principales Niveles Funcionales del Sistema Nervioso Central

Los niveles del sistema nervioso central que presentan características funcionales específicas son tres: el nivel medular, el nivel encefálico inferior o subcortical y el nivel encefálico superior o cortical.

2.1.2.1. Nivel medular

La médula espinal no es solo un conducto por el que viajan las señales nerviosas de la periferia, es un centro secundario de control. Los niveles superiores del sistema nervioso envían señales a los centros de control de la médula para que estos ejecuten las funciones necesarias. Por ejemplo, acciones como los movimientos de la marcha, los reflejos ante fuentes de dolor, mantenerse de pie y erguido ante la gravedad, el control autónomo de vasos sanguíneos locales y otros son algunas de las ordenes que puede ejecutar.

2.1.2.2. Nivel encefálico inferior o subcortical

La mayor parte de las actividades inconscientes del organismo son reguladas por este nivel, compuesta por las regiones inferiores del encéfalo, el bulbo raquídeo, la protuberancia, el mesencéfalo, el hipotálamo, el tálamo, el cerebelo y los ganglios basales. Incluso patrones emocionales pueden darse aun cuando gran parte de la corteza cerebral ha dejado de funcionar.

2.1.2.3. Nivel encefálico superior o cortical

En particular, esta última estructura o nivel del sistema nervioso central funciona como un enorme almacén de recuerdos y trabaja en conjunto con los niveles anteriores, y a su vez los anteriores con este. Gracias a la gran cantidad de

información almacenada los niveles inferiores son capaces de efectuar acciones con mayor precisión.

2.1.3. Efectos del Alcohol en el Cerebro y Sistema Nervioso

El alcohol es la droga depresora del sistema nervioso central capaz de crear tolerancia y dependencia psicofísica y adicción. Se le considera una sustancia poco potente ya que los efectos adversos de su ingesta se manifiestan al ingerir gramos de etanol, a diferencia de otras drogas potentes que en miligramos los efectos son evidentes (Guerri, 2000). Algunos síntomas de intoxicación aguda alcohólica son: dificultad para realizar tareas que conlleven habilidad, euforia y locuacidad; pérdida del autocontrol, falta de coordinación, lentitud mental y alteración de la percepción; amnesia, confusión, diplopía, disartria, hipotermia, náuseas, vómitos, depresión respiratoria y coma.

La estructura química del etanol lo hace compatible a diferentes entornos, tanto a grupos polares como no polares, lo que lo hace soluble a entornos acuosos y lipídicos. Esto lo convierte en un tóxico celular con el potencial de afectar varios tejidos y órganos. Por ello, se le relaciona su consumo excesivo con más de sesenta 60 enfermedades (Planas-Ballvé et al., 2017).

2.1.3.1. Mecanismos fisiopatológicos del Alcoholismo

La propiedad del etanol de ser soluble en entornos acuosos o lipídicos le permite atravesar libremente la barrera hematoencefálica y en pocos minutos afecte al sistema nervioso (Guerri, 2000). Es así como esta sustancia se relaciona con bastantes afectaciones neurológicas como las mencionadas en la Tabla 2.1. De igual manera, altera el correcto funcionamiento de los sistemas de neurotransmisión (Díaz-soto & Calderín-Miranda, 2020).

Tabla 2.1. Manifestaciones neurológicas secundarias al consumo excesivo del alcohol (Planas-Ballvé et al., 2017).

Sistema Nervioso Central	Sistema Nervioso Periférico
<i>Agudas</i>	
<ul style="list-style-type: none"> • Intoxicación aguda • Síndrome de abstinencia y delirium tremens • Encefalopatía de Wernicke 	<ul style="list-style-type: none"> • Neuropatía alcohólica aguda • Neuropatía compresiva asociada al alcohol
<i>Crónicas</i>	
<ul style="list-style-type: none"> • Síndrome de Korsakoff • Demencia relacionada con el alcohol • Enfermedad de Machiavava-Birgami • Degeneración Cerebelosa 	<ul style="list-style-type: none"> • Neuropatía alcohólica crónica • Neuropatía asociada al disulfiram

2.2. Predisposición Genética al Alcoholismo

El desarrollo del alcoholismo puede relacionarse a una diversidad de factores fisiológicos, genéticos, psicosociales y ambientales; sin embargo, se le asocia el 50% al factor genético (Rey Buitrago, 2015). No todos aquellos que ingieren alcohol se vuelven alcohólicos, ya que uno de los factores desencadenantes es la *susceptibilidad biológica*³ a padecerlo. Son numerosos los genes que se le relacionan a esta patología, pero su aporte individual es mínimo (Costas, 2015). Así mismo, se puede tener susceptibilidad genética al desarrollo de diferentes patologías derivadas del alcoholismo, como la cirrosis por ejemplo (Gaviria et al., 2016).

En 1999 Begleiter y Porjesz asociaban el alcoholismo a la posibilidad de ser heredado, pues varios estudios estimaban que los efectos genéticos aportaban un

³ Susceptibilidad Biológica hace referencia a la totalidad de condiciones con las que un individuo nace y le hacen más o menos predispuesto al desarrollo de una condición patológica (Rey Buitrago, 2015)

40-60% de la varianza en la causalidad de genética en el desarrollo del alcoholismo (Hesselbrock, 1995).

En ese entonces ya se teorizaba que el desequilibrio inhibitorio/excitatorio del sistema nervioso central promovía la dependencia al alcohol. Primeramente, se podía manifestar en las mañanas con pequeños temblores promoviendo la ingesta inmediata de etanol para mitigarlos. De igual manera, se observaban anomalías neuro-eléctricas en alcohólicos, tanto en hombres como en mujeres y sus relativos o descendencia (Begleiter & Porjesz, 1999).

Las investigaciones con los resultados más robustos sobre la influencia genética en el alcoholismo están relacionadas con SNPs⁴. Existen 7 genes de ADH y 18 genes de ALDH, pero sólo 4 clases de genes ADH (ADH1A, ADH1B, ADH1C y ADH4) y una de ALDH (ALDH2) cuya función recae en la metabolización del alcohol (Rey Buitrago, 2015).

Susumu Higuchi et al. estudiaron la influencia de las variaciones genéticas de las enzimas que participan en la metabolización del etanol, de acuerdo con el fenotipo⁵ del alcoholismo. Hallaron que las variantes ADH1B*2 generan una enzima 100 más activa, y ALDH2*2 con menor presencia; lo anterior relacionado al fenotipo que presenta enrojecimiento facial, hipertensión, dolor de cabeza y náuseas (S. Higuchi et al., 2004).

En Colombia, Méndez investigó la relación de la tipificación de polimorfismos de los genes ADH, ALDH y citocromo P450 en una población de la ciudad de Bogotá con el alcoholismo. Parte de sus hallazgos es que la combinación de genotipos ADH1B, ADH1C, ALDH2 y CYP2E1 junto con el con el genotipo homocigoto ALDH2*1 permite mayor tasa de oxidación del alcohol, dando lugar a mayores tasas

⁴ Los polimorfismos de nucleótido único (Single Nucleotide Polymorphism, en inglés) son un tipo de polimorfismo que producen una variación en un solo par de bases. Los científicos están estudiando cómo los polimorfismos de nucleótido único, o SNPs (pronunciado "snips"), en el genoma humano se correlacionan con enfermedades, con la respuesta de los fármacos, y con otros fenotipos (*Polimorfismos de Nucleótido Único (SNPs) | NHGRI*, n.d.).

⁵ El fenotipo constituye los rasgos observables de un individuo, tales como la altura, el color de ojos, y el grupo sanguíneo. La contribución genética al fenotipo se llama genotipo. Algunos rasgos son determinados en gran medida por el genotipo, mientras que otros rasgos están determinados en gran medida por factores ambientales (*Fenotipo | NHGRI*, n.d.).

de aclaramiento de etanol y acetaldehído, promoviendo un incremento en su tolerancia y consumo, siendo así mayor la dependencia hacia ella (Méndez Plazas, 2009).

De lo anterior, es evidente que gran parte de los factores que provocan el desarrollo del alcoholismo surgen a partir de la genética del organismo, para finalmente ser detonada por el entorno en el que se desenvuelve el individuo. La expresión de genes y actividad enzimática juegan un papel importante en la metabolización del etanol, sustancia que altera la homeostasis del cuerpo y altera la función normal del sistema nervioso y del cerebro del paciente.

2.3. Señales de EEG

Las señales fisiológicas del cuerpo humano son una fuente invaluable de información que refleja la actividad eléctrica de una parte específica del cuerpo. Ellas asisten a los profesionales de la salud a la detección, rehabilitación y tratamiento de pacientes. Dichas señales provienen de sensores (electrodos) implantados o colocados superficialmente y dependiendo del fenómeno estudiado su posicionamiento cambia, lo cual determina la señal fisiológica que se desea adquirir (Oliver Faust et al., 2018).

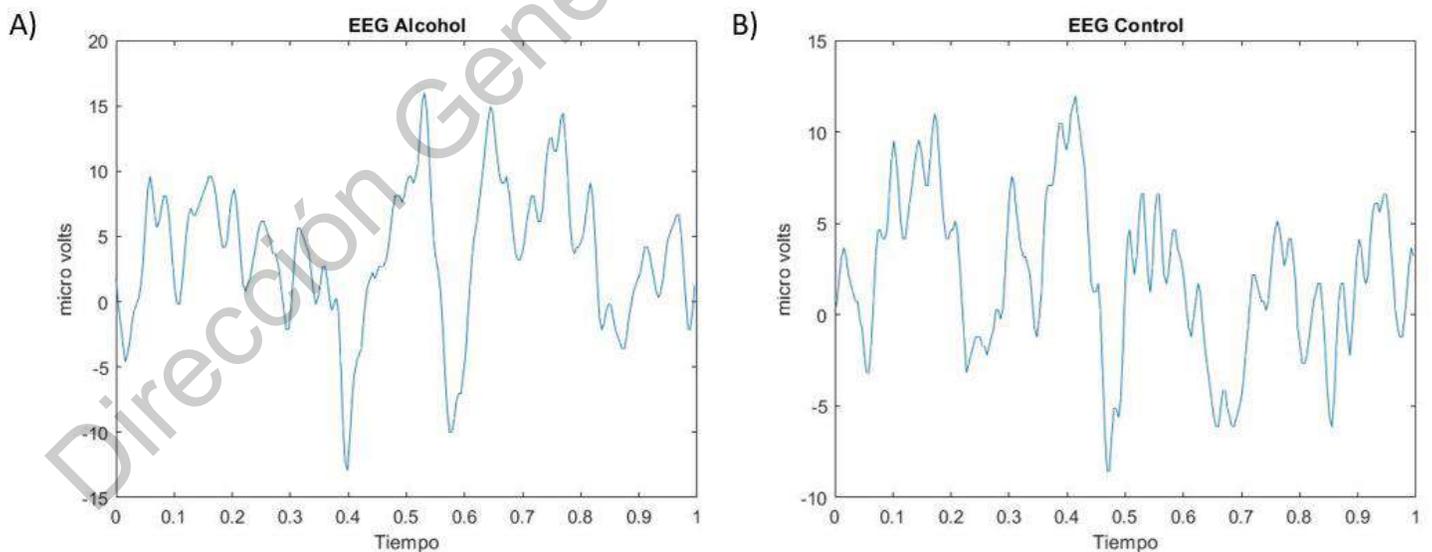


Figura 2.2. Señales de EEG. A) Se observa una señal de EEG perteneciente a un sujeto alcohólico. B) Señal de EEG perteneciente a sujeto control.

En el presente trabajo se utilizan electroencefalogramas de pacientes control y alcohólicos, ejemplificados en la Figura 2.2. Esta señal fisiológica registra las oscilaciones del potencial eléctrico cerebral proveniente principalmente de las neuronas piramidales ⁶, cuya adquisición es a través los electrodos colocados en el cuero cabelludo del individuo y proporcionan información valiosa para el estudio neurológico de trastornos epilépticos, afectaciones neuronales o alteraciones del sueño o investigaciones relacionadas con los ritmos cerebrales (P. L. Nunez et al., 2006).

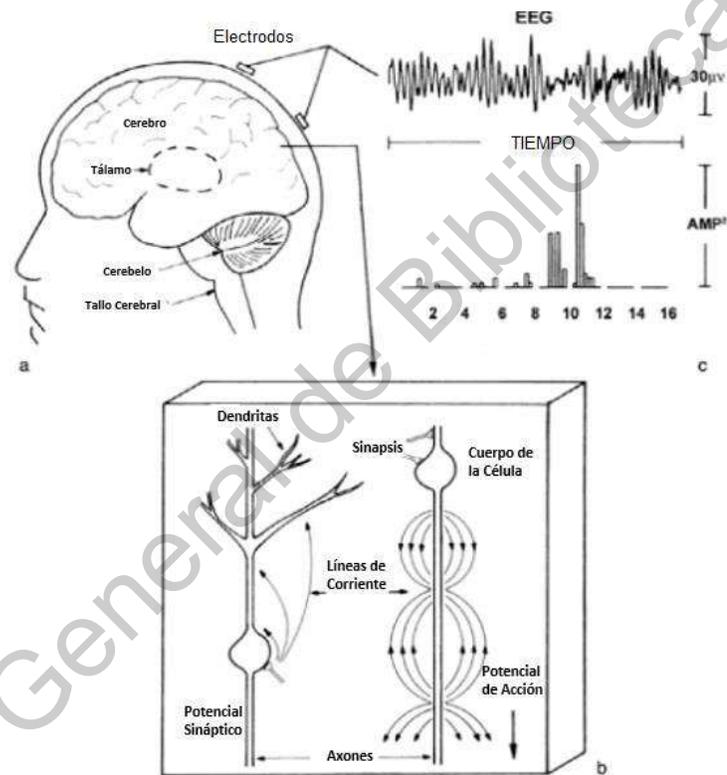


Figura 2.3. Representación básica de dónde proviene el EEG y su registro (P. L. Nunez et al., 2006).

El comportamiento del EEG se debe a las fuentes de corriente que provocan los cambios en los potenciales extracelulares. Dichas fuentes son la actividad postsináptica de las neuronas y se cree que estas se encuentran principalmente en el neocórtex (Nunez et al., 2006). En la Figura 2.3 se puede apreciar la

⁶ Tipo de neurona multipolar encargada de transmitir información lejos de su origen. Su característica principal es que poseen un soma con forma piramidal y una dendrita prolongada llamada dendrita apical (Castillero Mimenza, n.d.).

representación básica de cómo se generan y adquieren los EEG, donde en la Figura 2.3a se observan las diferentes partes que componen al cerebro humano, en la Figura 2.3b hay una sección de corteza cerebral con dos neuronas que generan una pequeña corriente generada por el potencial de acción y la sinapsis entre ellas y en Figura 2.3c se muestra cada electrodo del cuero cabelludo registra fuentes corticales promediados sobre centímetros cuadrados superficie. También se muestra un periodo de registro en el tiempo del ritmo alfa y su respectivo espectro de potencia. Los electrodos en la superficie del cuero cabelludo promedian los potenciales de numerosos dipolos de capas internas que son proyectados hacia el exterior. Un solo electrodo es capaz de proveer la estimación promediada de la activación sináptica de entre 100 millones y 1 billón de neuronas.

Las señales de EEG están compuestas por una serie de ondas eléctricas en el orden de los 0 a 300 μV con un ancho de banda de 0.5 a 100 Hz, las cuales son: delta (δ : 0 -4 Hz), teta (θ : 4.1 – 8 Hz), alfa (α : 8.1 – 13 Hz), beta 1 (β_1 : 13 – 25 Hz) y beta 2 (β_2 : 25 – 50 Hz) (García González et al., 1998).

2.3.1. Adquisición de señales de EEG

Le lectura de la actividad cerebral no es una tarea sencilla debido a su baja magnitud, sus frecuencias relativamente bajas y el ruido. El registro de los potenciales en el cuero cabelludo generalmente depende de tres factores: la ubicación y extensiones espaciales de los dipolos (producidos por las fuentes de corriente en la sinapsis), el volumen de conducción de la cabeza y el método empleado para adquirir la señal.

En todos los métodos de adquisición de bioseñales los electrodos juegan un papel sumamente importante, ya que no importa que tan avanzado sea el resto del sistema de medición, ninguna señal será adquirida adecuadamente a menos que los electrodos cumplan con ciertos requisitos. Los electrodos utilizados son metálicos cóncavos y pueden ser rellenados por una interfaz conductora, como la pasta electrolítica. Uno de los metales que se ha utilizado y es preferido para las aplicaciones neurofisiológicas es el cloruro de plata (AgCl) que, gracias a su ligera solubilidad, se satura y alcanza el equilibrio rápidamente, lo que lo hace un buen

electrodo superficial de piel (Stavrinidis et al., 2016). De igual manera, metales como el oro y platino son frecuentemente utilizados por su alta conductividad y biocompatibilidad. Estos pueden ser superficiales o intracraneales. Ambos métodos tienen sus ventajas y desventajas dependiendo de la aplicación, sin embargo, los superficiales son más utilizados por ser menos invasivos.

Los sistemas para la adquisición de EEG registran las diferencias de potencial entre pares de electrodos. El objetivo es registrar únicamente los potenciales debido a las fuentes de corriente, pero los artefactos⁷ o interferencias no pueden ser completamente eliminados por la instrumentación (P. L. Nunez et al., 2006) (García González et al., 1998).

2.3.2. Electrodos

El EEG es un registro multicanal donde cada canal es dado por un electrodo de registro, los electrodos trabajan en pares y registran los potenciales de la actividad cerebral respecto al electrodo de referencia, este último no es propiamente un canal. Se requiere un registro mínimo de 8 canales, pero existen diferentes arreglos de 12, 16, 64 o más canales dependiendo de la aplicación. Su posicionamiento es de suma importancia y se detallará más adelante con el sistema 10-20.

2.3.3. Sistema Internacional de Colocación de Electrodos 10-20

De acuerdo con García González et. al. y como su nombre lo indica, es un sistema internacional estandarizado para la colocación y medida de las posiciones a espacios iguales en el cuero cabelludo. Así como se presenta en la figura 2.4, el sistema usa identificadores en el cráneo como punto de referencia. Se basa en la relación comprobada que existe entre un sitio medido de electrodos y las estructuras y áreas corticales subyacentes. Se denomina 10-20 porque los electrodos se colocan espaciados al 10 o 20% de la distancia total entre un determinado par de

⁷ Son las señales no deseadas debidas al sujeto o al sistema de registro. Pueden ser de origen biológico, físico o ambos. Por ejemplo: fluctuaciones en la resistencia o el potencial de los electrodos, movimientos del sujeto, difusiones de la red, cargas estáticas, sudoración, entre otros (García González et al., 1998).

marcas en el cráneo y los porcentajes son empleados debido a la unicidad de cada cabeza humana (García González et al., 1998).

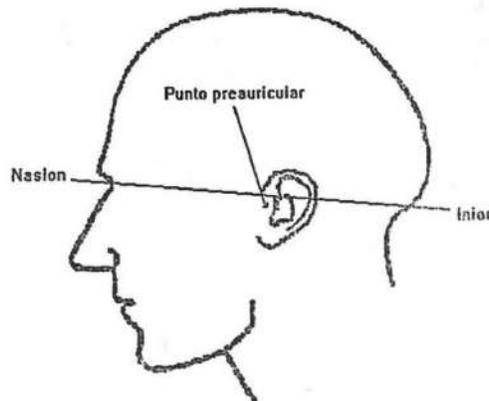


Figura 2.4. Vista sagital izquierda de las marcas sobre el cráneo (García González et al., 1998).

2.3.4. Nomenclatura del sistema 10-20

Este sistema utiliza cuatro marcas sobre el cráneo: el nasion, el inion y los puntos preauriculares izquierdo y derecho. Como se observa en la Figura 2.4, se forma una línea recta entre el inion y el nasion; el primero es una protuberancia que se encuentra en la parte posterior del cráneo en el hueso occipital, el segundo es una depresión hallada entre la nariz y la frente. Los puntos preauriculares son depresiones que se encuentran arriba de los cartílagos que cubren la abertura del oído externo.

La nomenclatura está diseñada para que cada electrodo se coloque en un sitio descrito por una abreviatura alfabética y así la transferencia de conocimientos no sea limitada por la fisiología del paciente, idioma de la experimentación o número de electrodos empleados. Todos los caracteres alfabéticos se refieren al área del cerebro sobre la cual se localizan, con excepción de la “z” que se usa para la referencia cero de la línea media. Siguiendo un orden anterior-posterior, los electrodos sobre la línea media llevan la abreviatura: Fz (Frontal Cero), Cz (Central Cero) y Pz (Parietal Cero).

labor de la telemedicina, las BCI⁸ (*Brain Computer Interface*, en inglés), los equipos de diagnóstico, entre otros. (Escobar et al., 2020) (Wabnitz et al., 2017).

2.4. Transformada Wavelet Discreta

La teoría las Wavelets (Ondoletas, en español) provee herramientas para el tratamiento de problemas relacionados con el procesamiento y compresión de imágenes y señales, codificación de bandas, entre otros, de manera continua o discreta con una amplia gama de aplicaciones. Es de gran interés el uso de la transformada wavelet para analizar las señales no estacionarias, porque proporciona una alternativa a la STFT⁹ (*Short-time Fourier Transform*, en inglés), en la cual la transformada de Fourier es aplicada en pequeñas ventanas por toda la señal, resultando en un espectrograma. Sin embargo, es limitada por el principio de incertidumbre de Heisenberg, eso es, un aumento en la resolución del tiempo disminuye la resolución en frecuencia y viceversa (Bajaj, 2021). En cambio, la WT¹⁰ (*Wavelet Transform*, en inglés) usa “ventanas” cortas de frecuencias altas y bajas, llamadas ondoletas, evitando ser limitada por este principio. También para ciertas aplicaciones, es deseable una perspectiva de la WT como medio de descomposición de la señal en un conjunto sub-bandas (Rioul & Vetterli, 1991). Se obtienen tras la dilatación y contracción de la ondoleta madre, de la cual hay una variedad de formas como: gaussiana, Gabor, Mexican Hat y otras más que se pueden observar en la figura 2.6.

⁸ Interfaz Cerebro Computadora, en español. Son interfaces que decodifican la actividad cerebral, a través de imaginación motora, reconocimiento de patrones e inteligencia artificial, con la finalidad de realizar acción de control sobre un sistema.

⁹ Transformada de Fourier de Tiempo Reducido, en español.

¹⁰ Transformada Ondícula, en español.

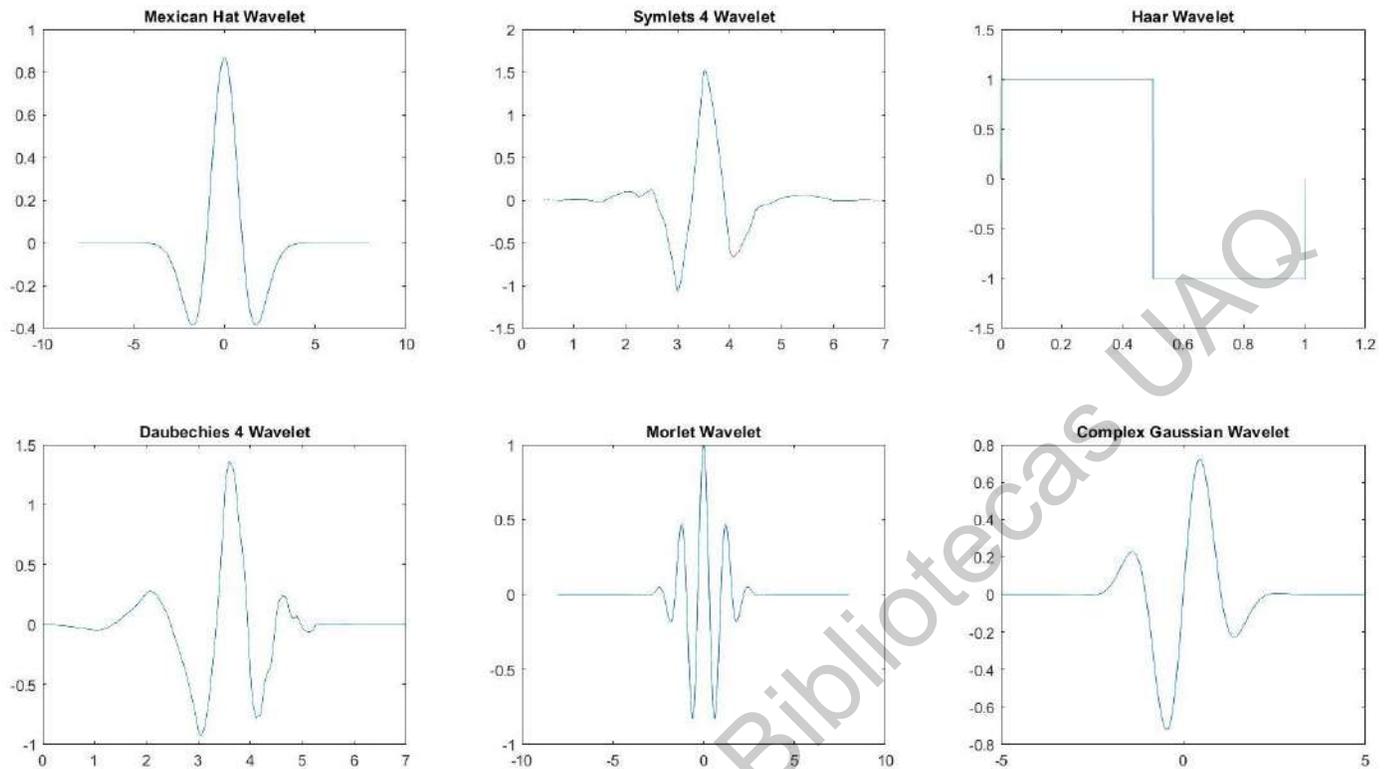


Figura 2.6. Algunas de las funciones de ondoleta utilizadas en la transformada wavelet (Bajaj, 2021).

La DWT es una transformación basada en el algoritmo de Mallat que descompone la señal, haciéndola pasar por una serie de filtros pasa altas y pasa bajas. Es cierto que la DWT presenta algunas diferencias frente a sus similares. La WTP se asemeja a la DWT con la gran diferencia de que descompone tanto para las frecuencias altas como las bajas. Como se observa en la Figura 2.7, la DWT sufre de una relativa baja resolución en la región de frecuencias altas, deficiencia que no existe en la WPT, que es capaz de capturar información de la variación de la frecuencia y magnitud de la señal en el tiempo (Gao & Yan, 2011). Sin embargo, la cantidad de datos generados por la DWT son suficientes para realizar un análisis detallado y completo de las señales de EEG para el presente trabajo.

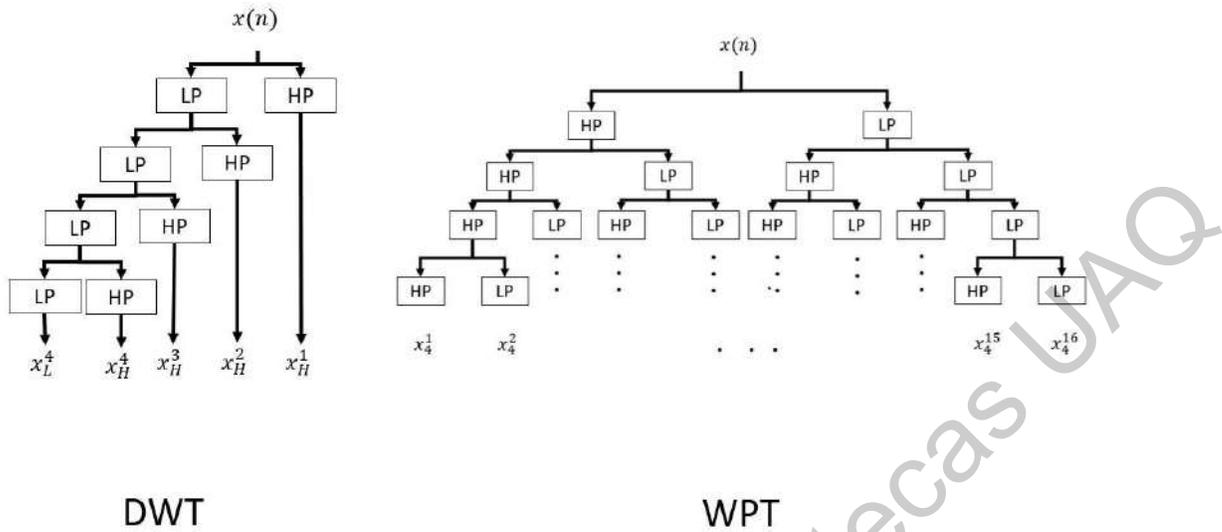


Figura 2.7. Comparativa entre DWT y WPD o WPT. LP: Low Pass (Pasa Bajas) HP: High Pass (Pasa Alta) (Bajaj, 2021).

La idea básica de la DWT hacer un análisis de señales con ondoletas es considerar una función base $\psi(t)$, llamada ondoleta madre, y generar una serie versiones de ella misma, ya sea que de expandan de manera horizontal o se dilate de manera vertical. Se define por la siguiente expresión (Brunton & Kutz, 2017):

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (2.1)$$

donde los parámetros a y b son los encargados de escalar y trasladar la función ψ respectivamente. La ondoleta cambia en longitud y frecuencia funcionando como una ventana que recorre toda la señal analizada. De acuerdo con sus características recopila la información de frecuencias altas o bajas, desempeñándose como un filtro pasa altas o pasa bajas. Cada vez que la ondoleta pasa por un fragmento de la señal base y su frecuencia coincide con la base se genera un coeficiente con valor numérico alto. Los Detalles son aquellos coeficientes que contienen frecuencias altas y las Aproximaciones contienen las frecuencias bajas. Utilizando ambos coeficientes se puede reconstruir una señal cercana a la original, una señal imperfecta. Por ello, es posible a través de la transformada wavelet descomponer una señal en sus diferentes bandas de frecuencia.

En la Figura 2.8 se representa el funcionamiento de la transformada wavelet discreta. Esta transformación toma una serie de N observaciones y produce N valores nuevos llamados coeficientes. La DWT se estructura por operaciones de filtrado que determinan el valor de estos coeficientes, los cuales contienen información de la magnitud y ubicación de eventos en diferentes escalas. La descomposición se realiza secuencialmente comenzando con las escalas menores y progresando a escalas mayores. En palabras de filtrado, se comienza con las frecuencias bajas y se prosigue con las frecuencias altas (Lindsay et al., 1996). En un inicio, dos ondoletas recorren la señal original con ciertos parámetros a y b dando como resultado las primeras Aproximaciones y Detalles. Posteriormente, la Aproximación es descompuesta en nuevos coeficientes por nuevas derivaciones de la ondoleta madre, generando así dos nuevas descomposiciones, concretamente una nueva Aproximación y un nuevo Detalle. Este proceso se repite hasta el número de niveles de descomposición, para el caso de la figura son 4 niveles.

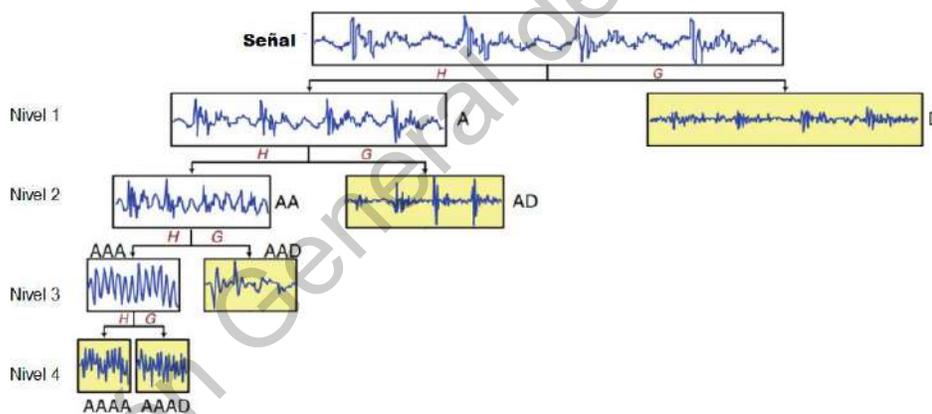


Figura 2.8. Procedimiento de descomposición de una señal en sus diferentes sub-bandas por medio de una transformada wavelet discreta. Notar que: A de Aproximación, D de Detalle, H de filtro pasa bajas y G de filtro pasa altas. Modificado de (Gao & Yan, 2011).

Las frecuencias que se encuentran en cada descomposición corresponden a la división consecutiva de las frecuencias bajas sobre 2. Si la frecuencia de muestreo fuera de 1 KHz para la señal de la Figura 2.8, en cada descomposición habría frecuencias en un rango decreciente, esto sería: 0 - 500 Hz para la Aproximación 'A' y 500 - 1000 Hz para el Detalle 'D'; 0 - 250 Hz para la

Aproximación 'AA' y 250 – 500 Hz para el Detalle 'AD'; así sucesivamente hasta el último nivel de descomposición, donde la Aproximación 'AAAA' contendría las frecuencias entre 0 y 62.5 Hz.

2.5. Fractales

De acuerdo con Di Ieva (2016), la geometría fractal es un modelo filosófico y matemático (Figura 2.9) utilizado para describir la escala de objetos físicos en el espacio, que no pueden ser caracterizados por una o algunas escalas.

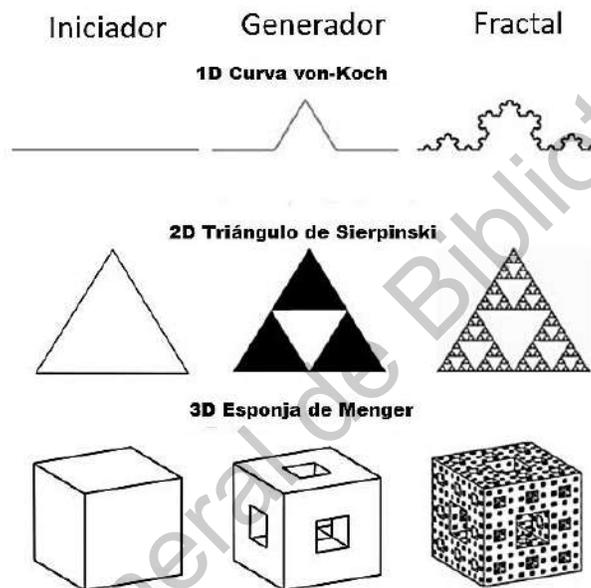


Figura 2.9. Fractales matemáticos ideales. Estas estructuras pueden ser generadas repetitivamente aplicando una sola regla de generación (generador) al objeto en escalas subsecuentes empezando en la más grande (iniciador). Este procedimiento en tan solo tres generaciones resulta en una estructura compleja auto-similar (fractal). Modificado de (Eke et al., 2002).

Un fractal también puede describir procesos dinámicos, como los observados en las series de tiempo fisiológicas. Los cambios dinámicos en las series de tiempo fisiológicas se combinan en una amplia gama de escalas de tiempo, sin que una escala sea dominante. Una dimensión fractal (D) determina cómo se acoplan los cambios más lentos o rápidos, aunque uno esté en el orden de segundos y el otro en el orden de días. Este acoplamiento da lugar a una auto semejanza de la variabilidad dinámica en el tiempo. De igual manera, existe un fractal estadístico, donde la similaridad no reside en el espacio o el tiempo de la variación de la variable

de interés, sino que el fractal está contenido en la escala de la distribución que describe las fluctuaciones aleatorias en las variables fisiológicas.

De acuerdo con el interés de la biomedicina, el uso de la dimensión fractal tiene una dualidad de beneficios: proveer un modelo único de análisis para intentar explicar la peculiaridad y singularidad de un fenómeno y, proponer un sistema universal de cuantificación, una herramienta usada para describir cosas naturales en un modo más realista.

El análisis fractal es una manera de medir un fenómeno cuando los detalles del diseño son tan importantes como la morfología bruta. El interés de este campo se reduce a una sola necesidad, medir los fenómenos peculiares que son difíciles de describir más allá de señalar que tienden a repetirse en diferentes escalas o tiene la propiedad de ser “autosimilares” (Di Ieva, 2016).

A continuación, se presentan los algoritmos de dimensión fractal con los cuales se realizó el análisis fractal de las señales en el tiempo; los cuales incluyen la dimensión fractal de: Katz, Higuchi, Counting-Box (*Conteo de Cajas*, en español), Petrosian y Sevcik.

2.5.1. Algoritmo de Dimensión Fractal de Katz

Michael J. Katz, creador del algoritmo KFD (*Katz Fractal Dimension*, en inglés), establece que, para las ondas, curvas planares que avanzan monótonamente, la dimensión fractal se encuentra en el rango entre $D = 1$ (para líneas rectas) a más o menos $D = 1.5$ (para ciertas ondas con picos altos). Las formas de onda verdaderas nunca son lo suficientemente complejas para llenar un plano; por lo tanto, nunca tendrán la fractalidad de un plano, es decir, $D = 2$.

Estima la complejidad de esta señal midiendo la relación entre la longitud total (L) de la curva y una línea recta correspondiente a la distancia Euclidiana máxima desde el primer punto (d). Esta relación es primeramente normalizada por la distancia media entre la serie sucesiva de puntos, a , donde N es el número de muestras que contiene la señal en el tiempo. El procedimiento anterior se resume en las siguientes ecuaciones (2.5) – (2.8) (Moreno-Gomez et al., 2020):

$$KFD = \frac{\ln\left(\frac{L}{a}\right)}{\ln\left(\frac{d}{a}\right)} \quad (2.2)$$

$$L = \sum_{i=2}^N |x(i) - x(i-1)| \quad (2.3)$$

$$d = \max(|x(1) - x(i)|) \text{ para } i = 2, \dots, N \quad (2.4)$$

$$a = \frac{L}{N-1} \quad (2.5)$$

2.5.2. Algoritmo de Dimensión Fractal de Higuchi

T. Higuchi presenta una técnica (HFD) para medir la dimensión fractal de una serie de puntos $(t, f(t))$ formando el gráfico de una función f definida en un intervalo (T. Higuchi, 1988).

Primero considera un arreglo finito de observaciones de series de tiempo tomadas en un intervalo regular:

$$X(1), X(2), X(3), \dots, X(N) \quad (2.6)$$

De las series de tiempo, primero se construyen nuevas series de tiempo X_k^m , definido de la siguiente manera:

$$X_k^m; X(m), X(m+k), X(m+2k), \dots, X\left(m + \left[\frac{N-m}{k}\right] * k\right) \quad (m = 1, 2, \dots, k) \quad (2.7)$$

donde $[\]$ denota la notación Gaussiana y ambos k y m , son enteros. m y k indican el tiempo de inicio y el intervalo de tiempo respectivamente. Para un tiempo con intervalo igual a k , se obtienen k conjuntos nuevos de series de tiempo. Se define el valor de la longitud de la curva:

$$L_m(k) = \frac{\left\{ \left(\sum_{i=1}^{\left[\frac{N-m}{k}\right]} |X(m+ik) - X(m+(i-1)*k)| \right)^{\frac{N-1}{\left[\frac{N-m}{k}\right]*k}} \right\}}{k} \quad (2.8)$$

donde el término $N - 1 / \left[\frac{N-m}{k} \right] * k$ representa el factor de normalización del largo de la curva en un subconjunto de series de tiempo. Higuchi define el largo de la curva para el intervalo de tiempo k , $\langle L(k) \rangle$ como el promedio del valor de k sobre los conjuntos de $L_m(k)$. Si $\langle L(k) \rangle \propto k^{-D}$, entonces la curva es fractal con la dimensión D .

2.5.3. Algoritmo de Dimensión Fractal de Recuento de Cajas

Wang propone el algoritmo de BFD (*Box-Counting Fractal Dimension*, en inglés) que calcula la dimensión fractal utilizando objetos llamados cajas que poseen la misma. Su tamaño está determinado por el área que cubrirán y el mínimo de cajas necesarias para hacerlo (Moreno-Gomez et al., 2020).

A pesar de sus limitaciones intrínsecas, se puede usar tanto para analizar la complejidad de una red vascular de un tumor como para analizar la distribución de árboles en un bosque, o análisis de la aspereza de una célula o la costa de una isla. En pocas palabras, la dimensión fractal D mide cuánto espacio llenan los datos o el objeto (Di Ieva, 2016). Se define (Moreno-Gomez et al., 2020):

$$BFD \cong \frac{\ln\left(\frac{\sum_{i=1}^{n-1} |x(i)-x(i-1)|}{\Delta t}\right)}{\ln(\Delta t)} \quad (2.9)$$

donde n es el número de muestras y Δt representa el intervalo de tiempo de muestreo.

2.5.4. Algoritmo de Dimensión Fractal de Sevcik

Sevcik propone un método (SFD) para calcular la dimensión fractal aproximada (D) a partir de un conjunto de N valores y muestreados a partir de una forma de onda entre el cero y el t_{max} con un intervalo de muestreo δ . La forma de onda fue sometida a una doble transformación lineal que la mapea en un cuadro unitario. La abscisa normalizada del cuadrado es x_i y la ordenada abscisa es y_i , ambas definidas como (Sevcik, 2010):

$$x_i^* = \frac{x_i}{x_{max}} \quad (2.10)$$

$$y_i^* = \frac{y_i - y_{min}}{y_{max} - y_{min}} \quad (2.11)$$

donde x_{max} es el máximo de x_i y, y_{min} y y_{max} son el mínimo y máximo de y_i . La dimensión fractal de la forma de onda (ϕ) es aproximada por D como:

$$\phi \approx D = 1 + \frac{\ln(L)}{\ln(2*Nr)} \quad (2.12)$$

donde L es el largo de la curva en la unidad cuadrada y $N' = N - 1$.

2.5.5. Algoritmo de Dimensión Fractal de Petrosian

Petrosian propone un algoritmo (PFD) en el que la señal de tiempo $X_i = \{x_1, x_2, \dots, x_n\}$ con longitud N es binarizada primero en una secuencia binaria z_i por (Moreno-Gomez et al., 2020):

$$f(x) = \begin{cases} 1, & x_i > \text{promedio}(x) \\ -1, & x_i \leq \text{promedio}(x) \end{cases} \quad (2.13)$$

Una vez que la señal es binarizada, el número total de símbolos adyacentes cambia, y N_Δ es estimada en la en la secuencia binaria como:

$$N_\Delta = \sum_{i=1}^{N-1} \left| \frac{z_{i+1} - z_i}{2} \right| \quad (2.14)$$

Finalmente, la dimensión fractal es calculada de la siguiente manera:

$$PFD = \frac{\log_{10} N}{\log_{10} N + \log_{10} \left(\frac{N}{N + 0.4N_\Delta} \right)} \quad (2.15)$$

2.6. ANOVA No Paramétrico

Dado que es necesario determinar el grado de independencia de los valores de los fractales calculados en el paso previo, es necesario realizar pruebas estadísticas. Algunas pruebas, las paramétricas, asumen que la población de la cual la muestra es extraída es normal o tiene una distribución normal, condición que debe ser cumplida para que la prueba de hipótesis sea válida (Quispe Andía et al., 2019). En ocasiones, esto no es posible de cumplir, por lo que es necesario usar otro tipo de algoritmos; en este sentido, las pruebas no paramétricas asumen que la población no sigue esta distribución. Estas son útiles cuando no se cumple el supuesto de normalidad y el tamaño de la muestra es pequeño (Minitab, n.d.).

El ANOVA (*Analysis Of Variance*, en inglés) es un método paramétrico para probar que las medias de dos o más poblaciones son iguales. La hipótesis establece que todas las medias de la población son iguales, mientras que la hipótesis nula establece que al menos una es diferente (Minitab, n.d.).

La prueba de Kruskal-Wallis es un método no paramétrico alternativo al ANOVA. Se puede considerar una generalización de la prueba de Mann-Whitney para cualquier variable de tipo categórica. A diferencia del ANOVA, en este método se contrasta si las diferentes muestras están equidistribuidas y que por lo tanto pertenecen a una misma distribución y en términos simples, se puede considerar que compara medianas (Joaquín Amat, 2016).

El procedimiento matemático por el cual se obtiene el resultado de la prueba de Kruskal-Wallis se describe a continuación (SEFO, 2019):

$$K = (N - 1) \frac{\sum_{i=1}^g n_i (\bar{r}_i - \bar{r})^2}{\sum_{i=1}^g \sum_{j=1}^{n_i} (r_{ij} - \bar{r})^2} \quad (2.16)$$

donde:

- n_i es el número de observaciones del grupo i
- r_{ij} es el rango (entre todas las observaciones) de la observación j en el grupo i .

$$\bar{r}_i = \frac{\sum_{j=1}^{n_i} r_{ij}}{n_i} \quad (2.17)$$

$$\bar{r} = \frac{N+1}{2} \text{ es el promedio de } r_{ij} \quad (2.18)$$

Note que el denominador de la expresión K es exactamente:

$$\frac{(N-1)N(N+1)}{12} \quad (2.19)$$

Luego,

$$K = \frac{12}{N(N+1)} \sum_{i=1}^g n_i (\bar{r}_i - \bar{r})^2 \quad (2.20)$$

Se puede realizar una corrección para los valores repetidos K por $1 - \frac{\sum_{i=1}^G (t_i^3 - t_i)}{N^3 - N}$ donde G es el número de grupos de diferentes rangos repetidos, y t_i es el número de observaciones repetidas dentro del grupo i que tiene observaciones repetidas para un determinado valor. Esta corrección hace cambiar a K muy poco al menos que existan un gran número de observaciones repetidas.

El p -value (valor p , en español) es aproximado por $\Pr(X_{g-1}^2 \geq K)$. Si algún n_i es pequeño (>5) la distribución de K puede ser distinta de la chi-cuadrado. Si el valor de p es pequeño, se puede concluir que las poblaciones tienen diferentes distribuciones; si es grande, las poblaciones tienen distribuciones que no difieren, pero no quiere decir necesariamente que sean iguales. Un ejemplo de esto se observa en la Figura 2.10.

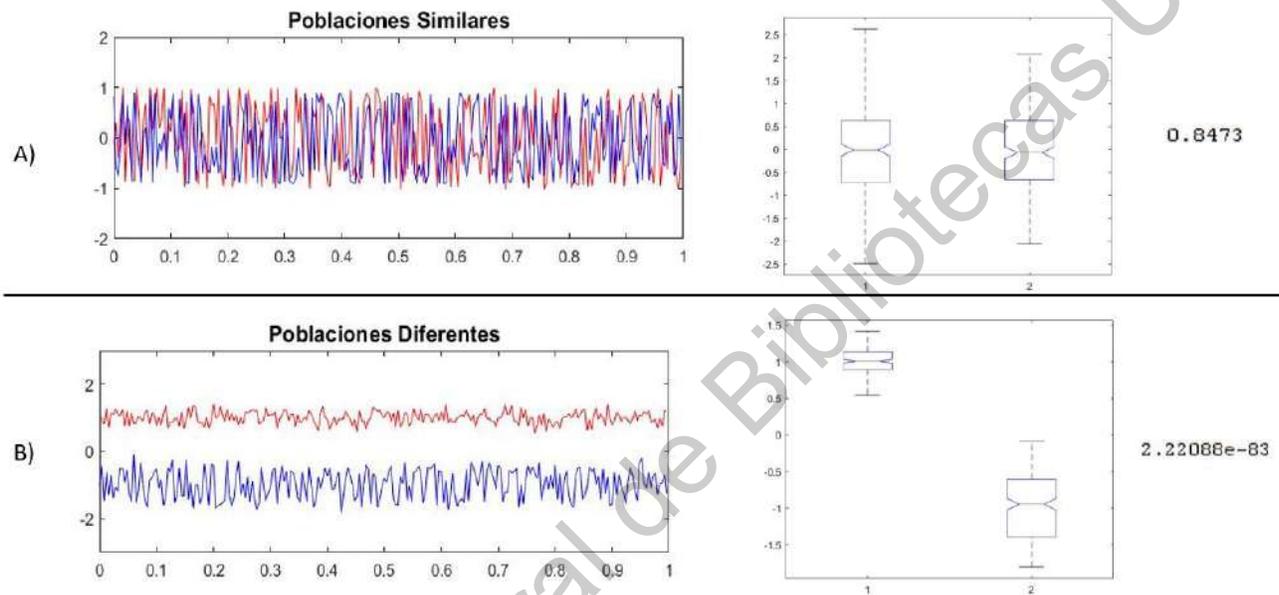


Figura 2.10. Representación gráfica del valor p y diagramas de cajas. A) Distribuciones de datos similares, por lo tanto, el valor de p es alto. B) Distribuciones diferentes, que presentan bajo valor de p .

En la Figura 2.10 se explica gráficamente cómo se comporta el valor de p y el diagrama de cajas de acuerdo con la distribución de los datos de las poblaciones a compararse. En pocas palabras, entre mayor sea la diferencia de distribuciones menor será el valor de p , con tendencia a cero, y mayor será la distancia entre cajas de poblaciones.

2.7. Clasificadores Basados en el Aprendizaje De Máquina

El aprendizaje de máquina es una sub-área de la inteligencia artificial y el aprendizaje profundo una sub-área del aprendizaje de máquinas (Pandey et al., 2019). La IA ha evolucionado a lo largo de los años, de ser una programación lógica a una tecnología de reconocimiento de patrones y Big Data (Macrodatos).

En la disciplina de machine learning (ML) se crean sistemas inteligentes con la capacidad de aprender automáticamente y el poder de encontrar patrones o hacer predicciones con enormes cantidades de datos. Estos sistemas son algoritmos que basan su funcionamiento principalmente en cálculo multivariable, álgebra lineal, optimización y métodos estadísticos que, dependiendo de la aplicación, varía su funcionamiento, como pueden ser los algoritmos de: regresión logística, redes neuronales, máquinas de vectores de soporte, árboles de decisión, clasificador de Bayes, por mencionar algunos. Estos algoritmos crean modelos matemáticos a partir de un entrenamiento realizado con un conjunto de datos, para que así sean capaces de realizar predicciones o decisiones sin que sean explícitamente programados para ello.

Se utilizan 3 tipos de conjuntos de datos para crear el modelo de regresión o clasificación. El primero es el conjunto de entrenamiento que le permite al modelo aprender sobre las “características” del grupo de datos y reconocer patrones. El segundo conjunto de datos es el de validación y se utiliza para afinar los parámetros iniciales con los que trabajará el modelo. El tercer es el conjunto de prueba, el cual está formado por datos no antes vistos por el modelo y con este se mide el desempeño del modelo para interpretar nuevos datos.

Usualmente, se divide la base de datos en los tres conjuntos anteriores. Una idea general y sin ser cierta para todos los casos es la siguiente: se le otorga al conjunto de datos de entrenamiento entre un 60-80% del total, el mayor porcentaje; para el conjunto de validación entre 5-20% y 15-30% para el de prueba. Todos estos porcentajes dependen de la aplicación que se desea realizar.

Dentro del aprendizaje hay dos categorías:

- Aprendizaje supervisado: en este tipo de aprendizaje los datos están etiquetados de acuerdo con su clase. El algoritmo solamente produce respuestas correctas nuevas y puede haber regresión (predicción) y clasificación.
- Aprendizaje no supervisado: los datos no están etiquetados y se busca una estructura o agrupación de los datos.

A continuación, se abordará el aprendizaje profundo con redes neuronales y ML con máquinas de vectores de soporte que son las técnicas de aprendizaje automático y supervisado utilizadas en este trabajo de tesis.

2.7.1. Deep Neural Networks

Los modelos de redes neuronales profundas (Deep Neural Networks, en inglés) son herramientas potentes de la inteligencia artificial y machine learning, específicamente del aprendizaje profundo (Deep Learning, en inglés), que su diseño se inspiró en la forma de funciona del cerebro humano, basado en neuronas y sus extensas redes de conexión. Teniendo la capacidad de aprender de ejemplos y datos, son empleadas en una amplia variedad de aplicaciones, desde la industria de cualquier tipo hasta la medicina y sistemas de salud. (S. Haykin, 2008)-

La arquitectura general de una red neuronal profunda se puede apreciar en la Figura 2.11; está constituida por unidades interconectadas llamadas neuronas y se ordenan en capas, siendo la primera la capa de entrada y la última la capa de salida. Entre éstas se encuentran las capas ocultas y son aquellas que realizan una transformación no lineal a los datos, permitiendo reconocer y extraer patrones en la información. Cada neurona en recibe entradas numéricas reales y produce un único resultado numérico real. No existe un número definido de capas ocultas y neuronas, estas se van adecuando de acuerdo con la aplicación pertinente; sin embargo, un mayor número de elementos demanda mayor carga computacional, pues se calcula una solución más compleja.

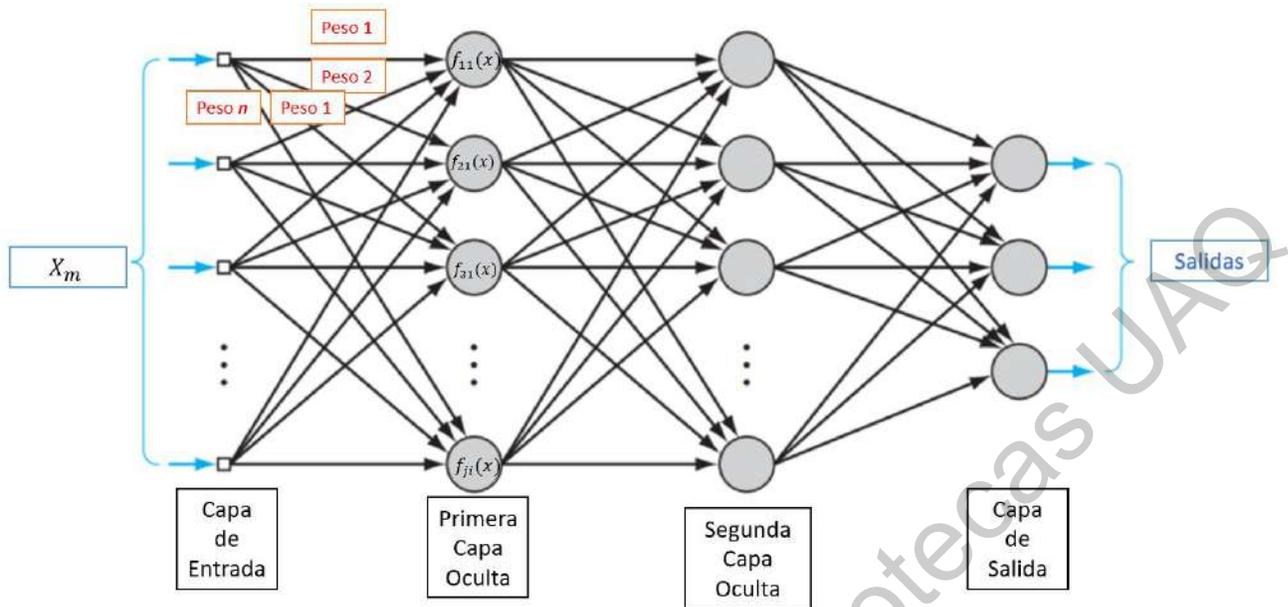


Figura 2.11. Representación gráfica de una red neuronal, donde el ejemplo de entrenamiento X_m es la capa de entrada. Posteriormente se modifica por los pesos y se procesa por las neuronas en cada capa, dando como resultado una capa de salida que, para el caso de la tesis, será una clasificación binaria. Gráfico modificado de (S. Haykin, 2008).

Partiendo de la misma figura, se observa la capa de entrada que corresponde a la observación X_m de la base de datos empleada, donde m indica el número de observación que se está ingresando (filas) a la red neuronal y que contiene n características (columnas), formando así una base de datos con dimensiones $m \times n$. Cada una de las características de la observación m ingresa a cada neurona j de la primera capa oculta, teniendo así n pesos por número j neuronas; formando una matriz de $j_i \times n_i$ pesos por capa. Cada neurona otorga un único resultado real tras el ingreso de las n características modificadas por los pesos a la función de activación $f_{ji}(X_m)$, donde i representa el número de capa oculta, teniendo así una “nueva característica modificada” en cada neurona de cada capa oculta de la red neuronal. Esto hace analogía al funcionamiento de una neurona cerebral, donde la información ingresa por las n dendritas (flechas) hacia la neurona y esta transmite el estímulo procesado a la siguiente neurona. Dependiendo del número de neuronas procesadoras en la capa anterior, será el número de dendritas para las neuronas de la capa posterior.

Una vez que la información se procesa hacia delante de la red neuronal, en la última capa los resultados se comparan con la etiqueta propia de la observación procesada. El error generado se propaga ahora hacia el interior de la red en sentido contrario, comenzando en la última capa y en dirección hacia la primera; el algoritmo generador de la retroalimentación se llama algoritmo de retropropagación, el cual permite que cada neurona reconozca y aprenda ciertas características de los datos, y así se obtengan mejores resultados (S. O. Haykin, 2011).

2.7.2. Máquinas de Vectores de Soporte

El principal objetivo de la máquina de vectores de soporte es encontrar un plano único en el que el margen de separación entre clases de datos se maximice. Bajo esta condición, el plano de decisión, como el de la Figura 2.12, se llama hiperplano óptimo y una idea básica de cómo llegar a él se detalla a continuación (S. Haykin, 2008).

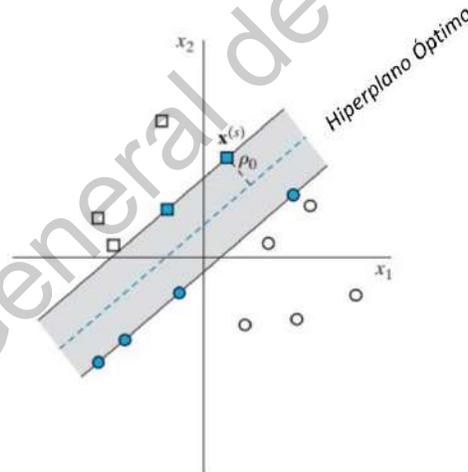


Figura 2.12. Ilustración de un hiperplano óptimo que establece el límite entre datos linealmente separables. Los datos de color azul son aquellos que corresponden a vectores de soporte. ρ_0 corresponde al margen de separación óptimo con respecto al hiperplano (S. Haykin, 2008).

Considerando un conjunto de datos $[(x_i, d_i)]_{i=1}^N$ donde x_i es una observación y d_i el valor esperado (etiqueta) se asume que aquellas observaciones con clase $d_i = +1$ o $d_i = -1$ son linealmente separables por el plano:

$$w_o^T x + b_o = 0 \quad (2.21)$$

donde w_o y b_o denotan los valores óptimos del vector de pesos (función similar a los de redes neuronales) y sesgo, respectivamente. La función discriminante es:

$$g(x) = w_o^T x + b_o \quad (2.22)$$

que otorga la medida algebraica de la distancia de x al hiperplano óptimo, de otra manera:

$$x = x_p + r \frac{w_o}{\|w_o\|} \quad (2.23)$$

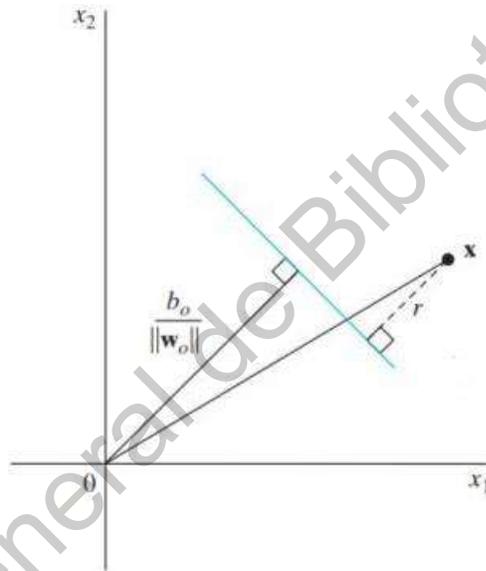


Figura 2.13. Interpretación geométrica de las distancias algebraicas de los datos hacia el hiperplano óptimo (línea azul) para dos dimensiones (S. Haykin, 2008).

donde x_p es la proyección normal de x en el hiperplano y r es la distancia algebraica deseada; r es positiva si x está en el lado positivo del hiperplano y negativa si está en el lado contrario, como se muestra en la Figura 2.13. Por definición, $g(x_p) = 0$ se obtiene lo siguiente:

$$g(x) = w_o^T x + b_o = r \|w_o\| \quad (2.24)$$

equivalentemente,

$$r = \frac{g(x)}{\|w_o\|} \quad (2.25)$$

Particularmente, la distancia desde el origen hasta el hiperplano óptimo está dada por $b_o/\|w_o\|$. Si $b_o > 0$, el origen está en el lado positivo del hiperplano; si $b_o < 0$, está en el lado negativo y si $b_o = 0$, el hiperplano óptimo pasa por el origen.

El problema en cuestión es encontrar los parámetros w_o y b_o que permitan hallar el hiperplano óptimo en un conjunto de datos de entrenamiento $T = \{(x_i, d_i)\}$. De acuerdo con la Figura 2.13, los parámetros deben satisfacer las siguientes características:

$$w_o^T x_i + b_o \geq +1 \quad \text{para } d_i = +1 \quad (2.26)$$

$$w_o^T x_i + b_o \leq -1 \quad \text{para } d_i = -1 \quad (2.27)$$

Los puntos particulares que cumplen con estas características se les llama vectores de soporte, pues son estos los puntos que forman el plano de separación entre clases, de ahí el nombre “máquina de vectores de soporte”. En otras palabras, los vectores de soporte son aquellos puntos que se encuentran cerca del hiperplano y por ello son los más difíciles de clasificar. Como tales, tienen una relación directa con la ubicación óptima de la superficie de decisión.

Si se considera un vector de soporte $x^{(s)}$ para el cual $d^{(s)} = +1$, entonces por definición se tiene que:

$$g(x^{(s)}) = w_o^T x^{(s)} + b_o = \mp 1 \quad \text{para } d^{(s)} = +1 \quad (2.28)$$

De (2.24), la distancia algebraica desde el vector de soporte hasta el hiperplano es

$$r = \frac{g(x^{(s)})}{\|w_o\|} \quad (2.29)$$

$$= \begin{cases} +\frac{1}{\|w_o\|} & \text{si } d^{(s)} = +1 \\ -\frac{1}{\|w_o\|} & \text{si } d^{(s)} = -1 \end{cases} \quad (2.30)$$

donde el signo positivo indica que el vector de soporte se encuentra en el lado positivo del hiperplano y el signo negativo lo contrario. Se denota ρ como el valor óptimo para el margen de separación entre clases que constituye la observación de entrenamiento de T. Así, de la ecuación 2.28 se obtiene que:

$$\rho = 2r = \frac{2}{\|w_o\|} \quad (2.31)$$

2.7.2.1. Kernels

Las máquinas de vectores de soporte utilizan una serie de funciones matemáticas definidas como kernel. Su función principal es tomar los datos como entrada y transformarlos a una forma deseada. Con más detalle, devuelven el producto escalar entre dos puntos realizado en un nuevo espacio dimensional de características distinto al original; así, la superficie o hiperplano de decisión no-lineal pueda transformarse en una ecuación lineal en un mayor número de espacios dimensionales (Awasthi, 2020). Utilizar un kernel permite hallar solución para cualquier dimensión, lo que se conoce como “kernel trick” (Amat, 2017).

Los kernel deben cumplir con el teorema de Mercer, el cual asegura que el algoritmo converja. Algunos de los kernels utilizados que cumplen con este teorema son (S. Haykin, 2008):

Tabla 2.2. Resumen de los Kernels más comunes que satisfacen el teorema de Mercer (S. Haykin, 2008).

Kernel Polinomial	$(x^t x_i + 1)^P$ <p>donde P es la potencia especificada por el usuario previamente.</p>
Función de Base Radial (RBF, Radial Basis Function)	$\exp\left(-\frac{1}{2\sigma^2}\ x - x_i\ ^2\right)$ <p>donde σ^2 es especificado previamente por los usuarios.</p>
Perceptrón de dos capas	$\tanh(\beta_0 x^T x_i + \beta_1)$ <p>solo algunos valores para β_0 y β_1 satisfacen el teorema</p>

Dirección General de Bibliotecas UAQ

CAPÍTULO 3. METODOLOGÍA

3.1. Metodología Propuesta

En seguida, se detalla la metodología que se siguió en el presente trabajo, de manera resumida se encuentra en la Figura 3.1.

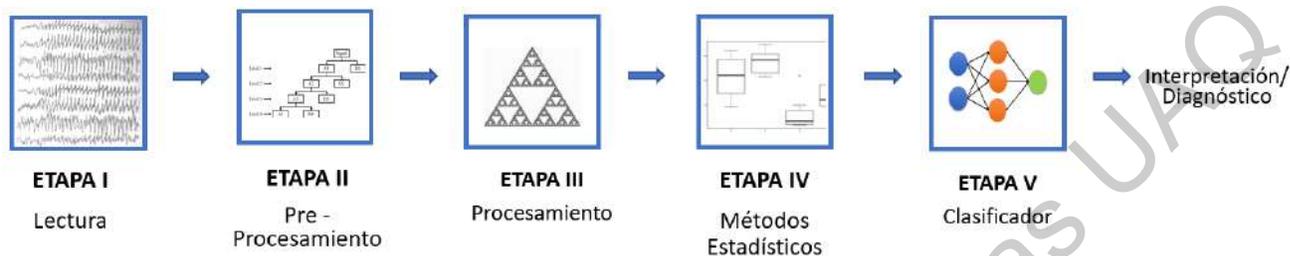


Figura 3.1. Metodología.

- 1. ETAPA I:** se llevó a cabo un procedimiento para poder leer la base de datos adecuadamente con Matlab 2020a y acceder a los datos de cada paciente. Debido a que es una base de datos incompleta, se seleccionaron aquellos pacientes con más información proveniente de los mismos estímulos. Más adelante en la sección 3.2 se detalla el proceso de selección y la base de datos utilizada.
- 2. ETAPA II:** Se realizó el preprocesamiento de las señales con transformaciones que permiten la descomposición en tiempo-frecuencia-amplitud de la señal de EEG. Específicamente, se utilizó la DWT para este fin. Empleando esta transformación con 6 niveles de descomposición y una ondoleta madre db44 se obtuvieron 12 derivaciones que se procesaron en las etapas posteriores. Se detalla y justifica el procedimiento en el apartado 3.3 de esta sección.
- 3. ETAPA III:** Se midieron los cambios de la geometría de las sub-bandas por medio de dimensión fractal. Para ello se prueban 5 algoritmos distintos, mismos que fueron descritos en la sección previa.
- 4. ETAPA IV:** Las medidas obtenidas de ambas categorías, sujetos control y sujetos con predisposición genética al alcoholismo, se comparan por medio del método estadístico no paramétrico Kruskal-Wallis con lo que se evidenciaron qué algoritmos de dimensión fractal y sub-bandas de

frecuencia aportan mayor información para diferenciar entre tipos de señales.

5. **ETAPA V:** Una vez seleccionados los algoritmos y sub-bandas estadísticamente significantes, se toman como características para entrenar modelos de aprendizaje automático y con ello generar un modelo capaz de clasificar entre clases con los datos obtenidos. Los modelos empleados en este paso fueron: SVM y MLP (*Multilayer Perceptron*, en inglés), este último es un algoritmo de aprendizaje profundo. Finalmente, se evalúa la eficacia de cada modelo para realizar la tarea de clasificación. Se calcula la especificidad, sensibilidad, exactitud y precisión.

3.2. ETAPA I: Lectura

3.2.1. Base de Datos

La base de datos utilizada para el presente trabajo de tesis proviene del repositorio de machine learning de la Universidad de California, Irvine (UCI) (Begleiter, 1999). El dueño original fue Henri Begleiter del Laboratorio de Neurodinámica del Centro de Salud de la Universidad de Nueva York. Posteriormente se donó esta base de datos para el repositorio en 1999. Los datos provienen de un estudio donde se buscó correlacionar señales de EEG con la predisposición genética de sufrir alcoholismo.

Los datos consisten en un conjunto de EEG adquirido durante una serie de pruebas donde el sujeto fue estimulado de dos maneras distintas: el estímulo uno (S1, una imagen) y estímulo dos (S1 y S2, dos imágenes); dichos estímulos se produjeron a través de imágenes de objetos (Figura 3.2). Cuando se mostraba el estímulo dos se presentaba S1 y S2 con condición de coincidencia o una condición de diferencia entre ellas. Se utilizaron 90 elementos de los 260 que conforman el conjunto de imágenes de Snodgrass y Vanderwart (Snodgrass & Vanderwart, 1980), las cuales están estandarizadas bajo una variedad de variables de importancia para las actividades de procesamiento cognitivo.

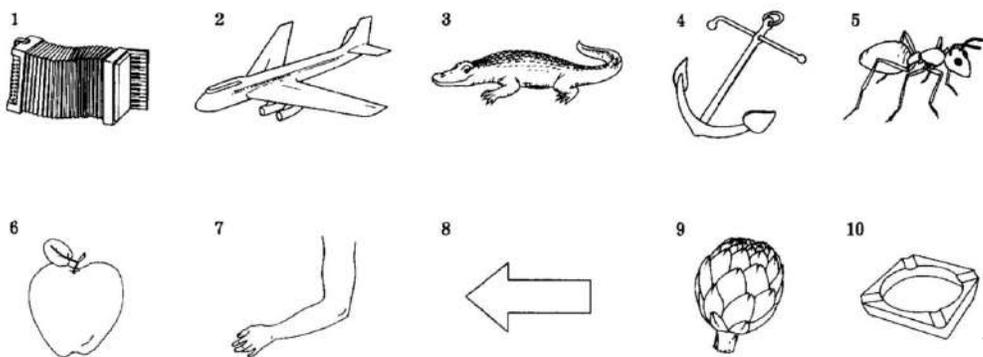


Figura 3.2. Representación de las imágenes de Snodgrass y Vanderwart, cuyo arreglo es en orden alfabético de acuerdo con su nombre más común. Modificado de (Snodgrass & Vanderwart, 1980).

En dicho repositorio se puede encontrar 3 bases de datos diferentes con distintos tamaños. La más pequeña contiene los datos de dos sujetos, uno alcohólico y otro control, con los tres paradigmas en cuestión, teniendo así 30 pruebas para cada clase. La base de datos grande contiene 10 sujetos de cada clase y 30 pruebas en total de cada uno, considerando también los 3 paradigmas. Por último, está la base de datos completa cuyo contenido contempla los dos grupos: 45 personas consideradas como control y 77 alcohólicas. Cada una de ellas se sometió a 120 pruebas, teniendo así un total de 122 sujetos con 120 pruebas cada uno. Las tres contienen medidas de 64 electrodos colocados en el cuero cabelludo utilizando la Nomenclatura Estandarizada para Posición de Electrodo 10-20, a una frecuencia de muestreo de 256 Hz (3.9 ms épocas) por 1 segundo. Cada prueba contiene la descripción de esta dentro de su archivo, como se observa en la figura 3.3. Para los fines de este trabajo se utiliza la base de datos completa.

Cabe mencionar que existen detalles aún más específicos sobre los sujetos control y alcohólicos que, lamentablemente, no han podido ser corroborados con la extensa variedad de los trabajos que utilizan esta base de datos. No obstante, se mencionan para no ser descartados: ambos grupos son conformados por sujetos masculinos que se mantuvieron abstinentes durante al menos 30 días para adquirir sus señales de EEG; el rango de edad del grupo de alcohólicos se encontraba entre los 22.3 y 49.8 años, con una media de 35.83 años; en el grupo control la edad media fue de 35.81 años y los sujetos no contaban con antecedentes familiares particulares de abuso de drogas, alcohol, antecedentes de enfermedad psiquiátrica

o trastorno neurológico; la impedancia de los electrodos estuvo siempre debajo los 5 kOms y se utilizó un electrodo nasal para conectar a tierra a los sujetos; se utilizaron dos desviaciones bipolares adicionales para registrar el electrooculograma vertical y horizontal; las señales se amplificaron con una ganancia de 10,000 mediante amplificadores EpA2 (Sensorium, Inc.) con un paso de banda entre 0,02 y 50 Hz registrándose con una computadora Concurrent 55/50 y los ensayos con movimientos oculares y corporales excesivos ($> 73.3 \mu\text{V}$) fueron rechazados en línea (Bavkar et al., 2019; Rodrigues et al., 2019). Cada prueba contiene la descripción de esta dentro de su archivo, como se observa en la figura 3.3. Para los fines de este trabajo se utiliza la base de datos completa.

```
# co2a0000364.rd
# 120 trials, 64 chans, 416 samples 368 post_stim samples
# 3.906000 msec uV
# S1 obj , trial 0
# FP1 chan 0
0 FP1 0 -8.921
0 FP1 1 -8.433
0 FP1 2 -2.574
0 FP1 3 5.239
0 FP1 4 11.587
0 FP1 5 14.028
0 FP1 6 11.587
```

Figura 3.3. Representación muestra del contenido de cada archivo de la base de datos.

3.2.2. Selección de Sujetos

La dificultad más grande presentada al momento de leer la tercera versión de la base de datos fue lo incompleta que se encuentra. Las pruebas están enumeradas del 0 al 119 y son la misma para cada sujeto; por ejemplo, los señales de EEG resultantes del sujeto 1 al ser expuesto al estímulo 1 (S1) se encuentran en la prueba 0 y la prueba 0 del resto de sujetos contiene la misma información respecto al mismo estímulo. Lamentablemente, no todos los pacientes cuentan con todas las pruebas y dentro de la inmensidad de archivos hay algunos que no contienen información o tienen errores; por esta razón se eliminaron aquellos archivos vacíos y se realizó un procedimiento de selección de sujetos efectivos para ser analizados, quienes debían cumplir con los siguientes requisitos:

- No poseer archivos vacíos.

- Tener al menos 30 ± 5 pruebas iguales con respecto a los otros sujetos.
- Tener las pruebas con los canales completos (64 electrodos)

Tomando en cuenta los puntos anteriores, el número total de sujetos contenidos en la base de datos se redujo a 30 sujetos, de los cuales 18 son sujetos considerados alcohólicos y 12 control. Las pruebas que fueron mayormente existentes en cada sujeto se muestran en la Tabla 3.1, donde las columnas son sujetos y las filas número de prueba. Como se lee en ella, los números de pruebas entre sujetos es mayormente similar, lo que permitió analizar y comparar a los sujetos bajo condiciones parecidas. Además, para ninguna de las pruebas su total alcanza 120 ejemplares, es decir, ninguna prueba existe en todos los sujetos.

Tabla 3.1. Similitud y número de prueba entre pacientes considerados para el trabajo

		SUJETOS																																								
PRUEBAS	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	6	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5							
	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	16	8	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6						
	7	7	7	7	8	7	7	7	7	7	7	7	7	7	7	7	7	7	7	17	9	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7						
	8	8	8	8	10	8	8	8	8	8	8	8	8	8	8	8	8	8	8	18	10	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8						
	9	9	9	9	11	9	9	10	9	9	9	9	9	9	10	9	9	9	9	19	11	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9					
	10	10	10	10	12	10	10	11	10	10	10	10	10	11	11	10	10	10	10	20	12	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10					
	11	11	11	11	13	13	11	12	11	11	11	11	11	12	13	11	11	11	11	21	13	11	15	11	11	11	11	11	11	11	11	11	11	11	11	11	11					
	12	12	12	12	14	14	12	13	12	12	12	12	12	13	14	12	12	12	12	22	14	12	16	12	12	12	12	12	12	12	12	12	12	12	12	12	12					
	13	13	13	13	15	15	13	14	13	13	13	13	13	14	15	13	13	13	13	23	15	15	17	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13				
	14	14	14	14	16	16	14	15	14	14	14	14	14	15	16	14	14	14	14	24	16	16	18	14	14	15	14	14	14	14	14	14	14	14	14	14	14	14				
	15	15	15	15	17	17	15	16	15	15	15	15	15	16	17	15	15	15	15	25	17	17	19	15	15	17	15	15	15	15	15	15	15	15	15	15	15	15	15			
	16	16	16	16	18	18	16	17	17	16	16	16	16	17	18	16	16	16	16	26	18	18	21	16	16	18	16	16	16	16	16	16	16	16	16	16	16	16	16			
	17	17	17	17	19	19	17	18	18	17	17	17	17	18	19	17	17	17	17	27	19	19	22	18	17	19	17	17	17	17	17	17	17	17	17	17	17	17	17			
	18	18	18	18	20	20	18	19	19	19	18	18	19	20	18	18	18	18	18	28	20	20	23	19	18	20	18	18	18	18	18	18	18	18	18	18	18	18	18			
	19	19	19	19	21	21	19	20	20	21	19	19	20	21	19	19	19	19	19	29	21	21	24	20	19	21	19	19	19	19	19	19	19	19	19	19	19	19	19			
	20	20	20	20	22	22	20	21	21	22	20	20	21	22	20	20	20	20	20	30	22	22	25	21	20	22	20	20	20	20	20	20	20	20	20	20	20	20	20			
	21	21	21	21	23	23	21	22	22	23	21	21	22	23	21	21	21	21	21	33	23	23	26	22	21	23	21	21	21	21	21	21	21	21	21	21	21	21	21	21		
	22	22	22	22	24	24	22	23	23	24	22	22	23	24	22	22	22	22	22	34	24	24	27	23	22	24	22	22	22	22	22	22	22	22	22	22	22	22	22	22		
	23	23	23	23	25	25	23	24	24	25	23	23	24	25	23	23	23	23	23	35	25	25	28	24	23	25	23	23	23	23	23	23	23	23	23	23	23	23	23	23		
	24	24	24	24	26	26	24	25	25	26	25	24	25	27	24	25	24	24	24	36	26	26	29	25	24	26	24	24	24	24	24	24	24	24	24	24	24	24	24	24		
	25	25	25	25	27	27	25	26	26	27	27	25	26	28	25	26	26	25	25	37	27	27	30	26	26	27	25	25	25	25	25	25	25	25	25	25	25	25	25	25		
	26	26	26	26	28	28	26	27	28	29	28	26	27	29	26	27	27	26	26	38	28	28	31	27	27	28	26	26	26	26	26	26	26	26	26	26	26	26	26	26		
	27	27	29	27	29	29	27	28	29	30	29	27	28	30	27	28	30	27	27	39	29	29	32	28	28	29	27	27	27	27	27	27	27	27	27	27	27	27	27	27		
	28	28	30	28	30	30	28	29	30	31	30	28	29	31	29	29	31	28	28	40	30	30	33	29	30	30	28	28	28	28	28	28	28	28	28	28	28	28	28	28		
	29	29	31	29	31	31	29	30	31	32	31	29	30	32	30	30	32	29	29	41	31	31	34	30	31	31	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	
	30	30	32	30	32	32	30	31	32	33	32	30	31	33	31	33	33	30	30	42	32	32	35	31	32	32	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	
	31	31	33	31	33	33	31	32	33	34	33	31	32	34	32	34	34	31	31	43	33	33	36	32	33	33	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	
	32	32	34	32	34	34	32	33	34	35	34	32	33	35	33	35	35	32	32	44	34	34	37	33	34	34	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32
	33	33	35	33	35	35	33	34	35	36	35	35	34	36	34	36	36	33	33	45	35	35	38	34	35	35	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33
	34	34	36	34	36	36	34	35	36	37	36	36	35	37	35	37	37	34	34	46	36	37	39	35	36	36	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34

3.3. ETAPA II: Preprocesamiento

Las señales de EEG, como ya se abordó en la teoría, cuentan con sub-bandas de frecuencia y que su magnitud depende directamente de la actividad cerebral que se esté llevando a cabo en el momento de su adquisición. Retomando su composición, las diferentes sub-bandas características son: delta (0 - 4 Hz), teta (4.1 - 8 Hz), alfa (8.1 - 13 Hz), beta 1 (13 – 25 Hz) y beta 2 (25 – 50 Hz) (García González et al., 1998).

Para preprocesar las señales de EEG es necesario considerar qué canales y qué sub-bandas de frecuencia se desean obtener. Como se ha fundamentado previamente, la actividad cerebral es inhibida por el alcohol de manera general, no por regiones específicas; por ello, se pueden considerar los 64 canales que conforman cada una de las pruebas para llevar a cabo el procesamiento de señales. En cuanto a las sub-bandas, se analizaron todas las descomposiciones posibles, derivadas de la Transformada Wavelet Discreta, con el fin de poder evaluar aquellas sub-bandas características que componen al EEG y también al resto de los datos. Seis niveles de descomposición permiten obtener en ciertas aproximaciones y detalles las frecuencias características de las sub-bandas delta, teta, alfa, beta y beta 2, como se presenta en la Tabla 3.2 y que se complementa con la Figura 3.4.

Tabla 3.2. Sub-bandas de EEG correspondientes a su descomposición tras Transformada Wavelet.

Descomposición	Frecuencia	Sub-banda	Frecuencia
Aproximación 6	0 – 4 Hz	Delta	0 – 4 Hz
Detalle 6	4.1 – 8 Hz	Teta	4.1 – 8 Hz
Detalle 5	8.1 – 16 Hz	Alfa	8.1 – 13 Hz
Detalle 4	16.1 – 32 Hz	Beta 1	13 – 25 Hz
Detalle 3	32.1 – 64 Hz	Beta 2	25 – 50 Hz

Como se observa en la Figura 3.4, aquellas descomposiciones de color azul son las que contienen las frecuencias de interés en un EEG, sin embargo, si el enfoque fuera únicamente en ellas, se podría perder información valiosa contenida en el resto de las descomposiciones.

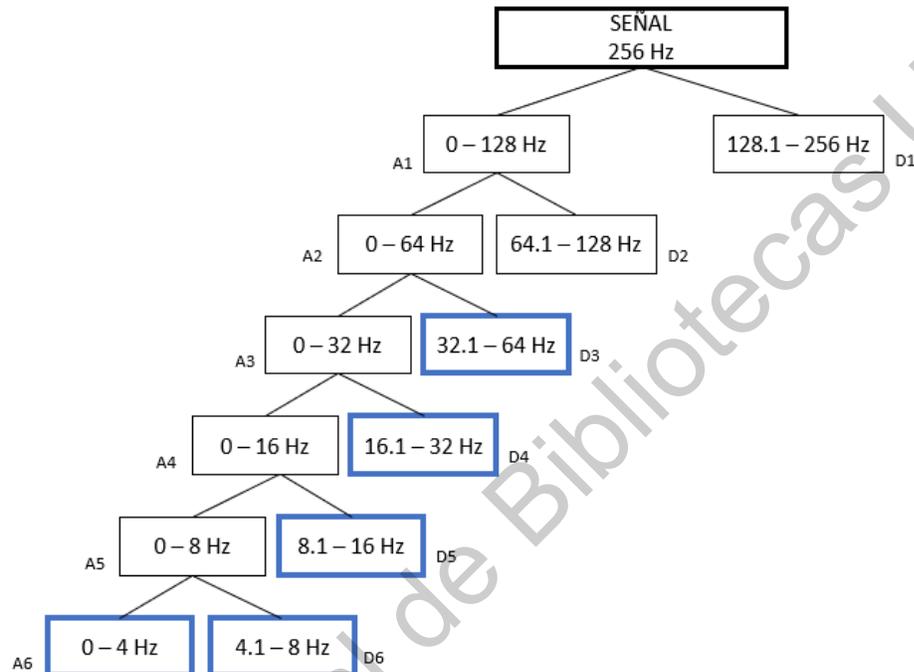


Figura 3.4. Representación gráfica de la descomposición en Aproximaciones (A) y Detalles (D) empleando la Transformada Wavelet Discreta. Aquellos recuadros marcados de azul corresponden a las sub-bandas características de un EEG.

Para llevar a cabo el cálculo de coeficientes durante transformación wavelet, se consideró una ondoleta madre db44 (Daubechies 44), pues en el estado del arte para los sistemas de diagnóstico esta ha demostrado ser una ondoleta eficiente para el diagnóstico con EEG (Al-Qazzaz et al., 2015; Gandhi et al., 2011; Rodrigues et al., 2019; Salis et al., 2013; Zhou et al., 2012).

3.4. ETAPA III: Procesamiento

Cada una de las bandas derivadas de la DWT fueron procesadas por medio de algoritmos de reconocimiento de patrones para señales no lineales, es decir, por medio de dimensión fractal. Esta tiene la capacidad de medir la complejidad geométrica, así como la “irregularidad” de una señal. Como se mencionó, mediante cinco algoritmos es posible estimarla: Dimensión Fractal de Katz (Katz, 1988),

Dimensión Fractal de Higuchi (T. Higuchi, 1988), Dimensión Fractal de Caja (Wang, 2005), Dimensión Fractal Petrosiano (Petrosian, 1995) y Dimensión Fractal de Sevcik (Sevcik, 2010). El resultado de realizar una medición sobre una señal de EEG con alguno de estos algoritmos será una cifra numérica entre 1 y 2, debido a que curva simple se encuentra en una dimensión y un plano en dos dimensiones; de ahí que también se le llame dimensión fractal, pues considera dimensiones fraccionarias.

Como se expuso previamente, de cada señal se obtienen 12 descomposiciones tras realizar la DWT y entre ellas se encuentran las 5 bandas características de un EEG. Cuando una señal se trata con un algoritmo de fractalidad el resultado es una cifra decimal real, cada una de ellas se guardó en una matriz de 3 dimensiones de la siguiente manera (Figura 3.5): prueba x canal x nivel de descomposición, generando una matriz por cada tipo de fractalidad para detalles y aproximaciones; en total 10 matrices.

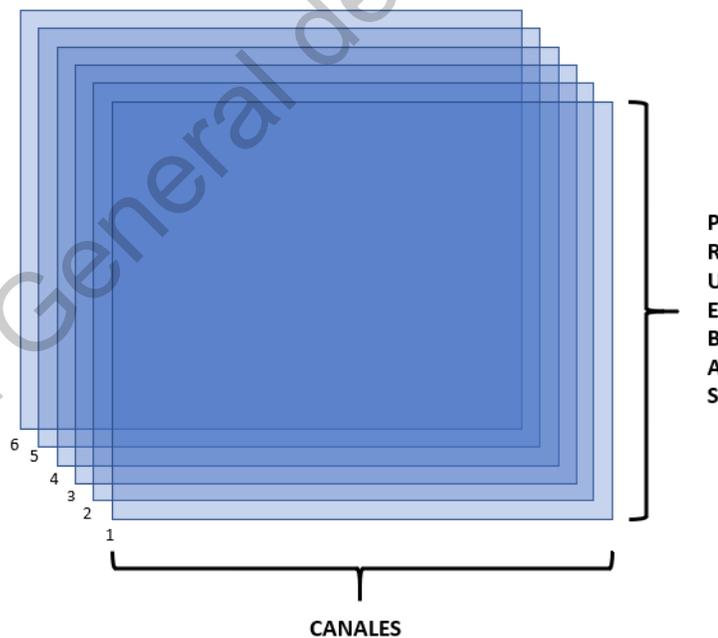


Figura 3.5. Representación gráfica del acomodo de los datos obtenidos tras implementar los algoritmos de dimensión fractal para posteriormente ser analizados por el método de Kruskal-Wallis. En las filas se indican las pruebas (30 x número de sujetos), en las columnas los canales de EEG (64 electrodos) y en cada recuadro de la figura, enumerado del 1 al 6, se indica el nivel de descomposición.

Cabe mencionar que, debido a la naturaleza de las señales de EEG, se consideran no lineales, altamente ruidosas y con gran cantidad de artefactos, la teoría del caos está dotada de descriptores cuantitativos efectivos de la dinámica del EEG y el caos subyacente en el cerebro. Por ello, el análisis no lineal para estudiar la complejidad, la aleatoriedad y el comportamiento caótico de la señal propiciaría la revelación de un comportamiento más complejo de las señales (Mohanchandra et al., 2016).

3.5. ETAPA IV: Método Estadístico No Paramétrico

Una vez que se calcularon las medidas de dimensión fractal de cada uno de los diferentes canales de EEG, se evaluaron tanto los algoritmos de fractalidad como los 64 electrodos y sus descomposiciones para saber qué combinación de ellos podría tener mayor relevancia o capacidad de ser asociada al fenómeno estudiado, ya sea control o sujeto con predisposición a sufrir alcoholismo. Para ello realizó un análisis estadístico no paramétrico basado en la varianza de los datos (ANOVA), específicamente el método de Kruskal-Wallis.

Como se mencionó previamente, el número de sujetos que cumplieron con los requisitos establecidos fue de 30, de los cuales 18 eran considerados alcohólicos y 12 eran control. Para tener poblaciones iguales de ambas clases se tomó el mismo número de sujetos control como de alcohólicos, en un principio se eligió 12 como número de sujetos para cada clase, ya que fue el máximo número de sujetos control disponibles; teniendo así una suma de 12 sujetos para cada clase, es decir, 360 pruebas tanto de alcohólicos como sujetos control.

Posteriormente, se analizó con el método estadístico propuesto canal control por canal alcohólico; considerando también sus respectivas sub-bandas y los distintos algoritmos de caracterización. El método estadístico como resultado proporciona diferentes elementos, como los que se observan en la Figura 3.6; sin embargo, para fines de selección se tomaron en cuenta únicamente los diagramas de cajas y valor de p (probabilidad), cuyo valor se encuentra marcado por un recuadro rojo. Para que el canal, descomposición y fractalidad fueran considerados como característica pertinente para el entrenamiento del modelo de aprendizaje

automático, debían poseer un valor de p muy bajo y un diagrama de cajas donde la media de las dos clases fuera muy distinta entre clases, es decir, que la línea roja dentro de las cajas se encontrara lo más separada posible. Dentro de la caja se encuentra el 50 % de los datos, y en los bigotes, tanto el superior como el inferior, el 25 % de ellos. Las cruces de color rojo son denominados valores atípicos, que son datos del conjunto que se encuentran muy distantes del valor central de la población. Para el presente estudio se buscó características con la menor cantidad de valores atípicos, a pesar de que en el método no paramétrico utilizado los estos valores pierden relevancia porque el conjunto de datos es ordenado de acuerdo con la mediana; sin embargo, se asumió que su cuantiosa presencia y distante posicionamiento del valor central pudiera no aportar información relacionada al fenómeno en cuestión.

Tabla ANOVA Krusk-Wallis					
Fuente	SS	df	MS	Chi-sq	Prob>Chi-sq
Columnas	9.45863e+06	1	9458625.8	218.65	1.78538e-49
Error	2.16453e+07	718	30146.7		
Total	3.11039e+07	719			

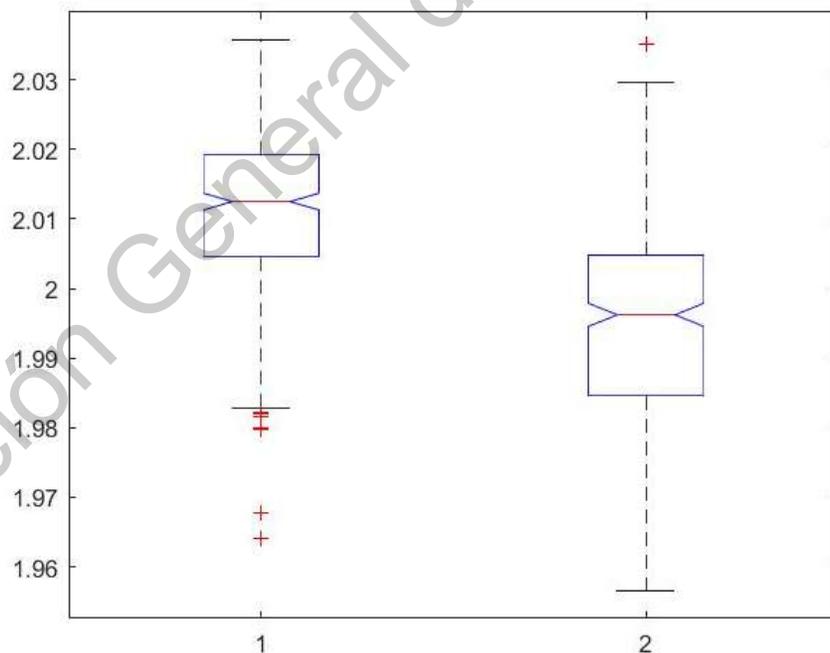


Figura 3.6. Diagrama de cajas y tabla de resultados obtenida tras el análisis estadístico no paramétrico Kruskal-Wallis. El primer resultado de interés es el valor de p (prob > chi-cuadrada) resaltado en un recuadro rojo. En el eje horizontal de la figura se pueden identificar las clases con el número 1 y 2, donde el 1 corresponde a control y 2 a alcoholismo. El segundo resultado de interés son las medias de las clases, representada por la línea roja a la mitad de ambas cajas. Entre mayor sea la diferencia entre medias, mayor será la distancia entre ellas.

3.6. ETAPA V: Clasificación

Una vez seleccionadas todas las características posibles, se entrenaron los modelos de aprendizaje automático y se validó su desempeño. Para llevarlo a cabo, se construyó una base de datos con las características seleccionadas del análisis estadístico, donde las filas correspondían a las pruebas (observaciones), las columnas a cada una de las características y al final se añadieron dos columnas extra con una etiqueta de identificación para cada clase de cada observación, representada con 0 para sujetos control y 1 para sujetos alcohólicos (Figura 3.7). Como ya se ha mencionado, los modelos seleccionados para realizar la tarea de clasificación son: perceptrón multicapa y máquinas de vectores de soporte. Se utilizó el 80% de los datos para el entrenamiento y el 20% para la prueba del desempeño. Se designaron tantos sujetos control como alcohólicos a ambos conjuntos de datos para mantener un equilibrio del 50% - 50%. La plataforma y lenguaje de programación utilizados para entrenar los modelos fue Matlab R2020a.

Observaciones	x_1	x_2	x_3	x_4	x_5	x_6	C_1	C_2
1	0	1
2	0	1
...	1	0
n	0	1

Figura 3.7. Acomodo de los datos para entrenar los modelos de aprendizaje de máquinas. En el caso de la SVM no es necesario añadir la última columna correspondiente a la segunda clase.

3.6.1. Perceptrón Multicapa

Se utilizó el Deep Learning Toolbox de Matlab para entrenar el perceptrón multicapa, específicamente con la función “*patternnet*” que permite generar una red reconocedora de patrones. Para afinar el proceso de entrenamiento se seleccionó el algoritmo de optimización de mínimos cuadrados y la función de entrenamiento Levenberg-Marquardt, también conocida como el algoritmo de mínimos cuadrados amortiguados.

Tanto el número de neuronas por capa oculta como el número de capas ocultas se seleccionó de manera heurística en el rango de 5 a 20 neuronas y de 1 a 3 capas ocultas, dando los mejores resultados 3 capas ocultas y 10 neuronas por capa oculta. En particular, con las redes neuronales, es necesario añadir una columna más a la base de datos de entrenamiento, la cual contiene la clase contraria a la que pertenece cada observación, esto es necesario ya que el número de neuronas en la capa de salida de una red neuronal es igual al número de clases en las que se desea clasificar cada una.

3.6.2. Máquina de Vectores de Soporte

Se entrenó el modelo de Máquina de Vectores de Soporte con la librería de acceso libre 'LIBSVM' que se puede implementar en varios lenguajes. En el caso de Matlab, se requiere un compilador C++ para poder crear las librerías y utilizarlas en la plataforma. Dichas librerías permiten mayor personalización a la hora de realizar el afinado del entrenamiento.

Tanto los kernel como parámetros de ajuste fueron seleccionados por desempeño. Los kernel probados que cumplen con el teorema de Mercer fueron: kernel polinomial, kernel gaussiano y kernel RBF; siendo el último el mejor en la práctica. Para elegir los mejores parámetros de ajuste (C, γ) se siguieron las recomendaciones descritas en (Hsu et al., 2003), lo cual conllevó a desarrollar una búsqueda de rejilla manual (*Manual Grid Search*, en inglés) para seleccionar el par de parámetros que tuviera el mejor porcentaje de precisión en la validación cruzada de ν pliegues.

En la validación cruzada de ν pliegues consiste en dividir el conjunto de entrenamiento en ν subconjuntos de igual tamaño. Secuencialmente se prueba un subconjunto utilizando el clasificador entrenado en los subconjuntos $\nu - 1$ restantes. Por lo tanto, cada instancia de todo el conjunto de entrenamiento se predice una vez, por lo que la precisión de la validación cruzada es el porcentaje de datos que se clasifican correctamente. Esto permite reducir el sesgo y varianza a la hora de implementar el modelo, lo que se traduce a obtener un modelo generalizado a los

datos con esa naturaleza más que a un modelo especializado a la base de datos utilizada, lo que previene el sobre entrenamiento.

Se crearon dos secuencias de números con crecimiento exponencial 2^{+2} , desde 2^{-7} hasta 2^{15} para el valor de C y otro desde 2^{-15} hasta 2^{11} para el valor de γ . Se evaluó el desempeño de cada posible combinación de par de valores y se seleccionaron aquellos con precisión más alta. Posteriormente se evaluó el modelo entrenado con tales parámetros y se seleccionó el mejor modelo.

3.7. Evaluación diagnóstica de los modelos

Para diagnosticar el desempeño de los modelos entrenados se utilizaron métricas de exactitud, precisión, sensibilidad y especificidad. En seguida, se presenta su definición y fórmulas para estimarlas que hacen referencia a la matriz de confusión (Figura 3.8):

- **Exactitud:** Se refiere a la dispersión del conjunto de valores obtenidos a partir de mediciones repetidas de una magnitud. Cuando menor es la dispersión mayor la precisión. Se representa por la proporción entre el número de predicciones correctas (positivas como negativas) y el total de predicciones. En otras palabras, es un parámetro que indica el porcentaje de predicciones correctas. Hay que tener en cuenta que carece de utilidad cuando el conjunto de datos no está equilibrado. Su cálculo se puede efectuar de la siguiente manera (Recuerdo, 2018):

$$Exactitud = \frac{a + d}{a + b + c + d}$$

- **Precisión:** está relacionada con el sesgo de una estimación y también se le conoce como 'Verdadero Positivo'. Se representa por la proporción entre los positivos reales predichos por el algoritmo y todos los casos positivos. De otra manera, se refiere a lo próximo que está el resultado de una medición del valor verdadero. Se estima a través de la siguiente ecuación (Recuerdo, 2018):

$$Precisión = \frac{d}{b+d} \tag{3.1}$$

- **Sensibilidad:** Corresponde a la proporción de sujetos evaluados correctamente con la condición en cuestión. Es decir, es la proporción de verdaderos positivos correctamente identificados por el test del total de individuos enfermos según el estándar de referencia (Bravo-Grau et al., 2015). Su cálculo es el realiza de la siguiente manera (Recuerdo, 2018):

$$\text{Sensibilidad} = \frac{d}{c+d} \quad (3.2)$$

- **Especificidad:** Es la proporción de sujetos correctamente diagnosticados con ausencia de la condición en cuestión por la prueba diagnóstica. Lo que es igual a decir que es la proporción de verdaderos negativos identificados correctamente (Bravo-Grau et al., 2015). El cálculo se realiza a través de la siguiente ecuación (Recuerdo, 2018):

$$\text{Especificidad} = \frac{a}{a+b} \quad (3.2)$$

A)

Matriz de Confusión		Valores estimados por el modelo	
		Negativo	Positivo
Valores Reales	Negativo	a: negativos verdaderos	b: falsos positivos
	Positivo	c: falsos negativos	d: positivos verdaderos

B)

Exactitud: Porcentaje de predicciones correctas.

$$\frac{a + d}{a + b + c + d}$$

Precisión: verdaderos positivos

$$\frac{d}{b + d}$$

Sensibilidad: Porcentaje casos positivos detectados.

$$\frac{d}{c + d}$$

Especificidad: Porcentaje de casos negativos detectados.

$$\frac{a}{a + b}$$

Figura 3.8. A) Matriz de confusión, donde las filas se nombran según las clases reales y las columnas según las clases previstas. La 'a' es el número de negativos reales; 'b' es el número de falsos positivos, 'c' es el número de falsos negativos y 'd' es el número de positivos reales. B) Resumen de las métricas para evaluar los modelos, donde las letra de las fórmulas provienen de la matriz de confusión.

Dirección General de Bibliotecas UAQ

CAPÍTULO 4. EXPERIMENTACIÓN Y RESULTADOS

Finalizada la descripción y justificación de la metodología implementada se presentan en este capítulo los resultados obtenidos después de llevar a la práctica cada uno de los pasos descritos en el capítulo anterior. Los resultados se detallarán de manera ordenada siguiendo el procedimiento propuesto.

4.1. Resultados Obtenidos

Como ya se ha mencionado en la sección 3.2, tras realizar la selección de pacientes se obtuvieron 30 sujetos; de los cuales 12 fueron control y 18 alcohólicos. Tomándose en cuenta los 64 electrodos como señales, todas derivadas del Sistema 10-20 que integran la información de cada una de las pruebas y, contemplando 12 sujetos de ambas clases de pacientes, se compilaron en total 46080 señales ($24 \text{ pacientes} \times 30 \text{ pruebas} \times 64 \text{ electrodos}$).

Una vez reunidas las señales, se sometieron a preprocesamiento con la transformada wavelet discreta. En la Figura 4.1 se pueden apreciar gráficamente la perspectiva de la DWT como una herramienta de descomposición de una señales en sus deferentes sub-bandas de frecuencia, frecuencias que fueron establecidas en la sección 3.3 del capítulo anterior. Para realizar dichos gráficos se utilizó la señal proveniente del electrodo 1 de la prueba 5 tanto del sujeto alcohólico como control número 9. Como se planteó, 12 descomposiciones se obtienen de una señal y cada una de ellas corresponde a las mostradas en la Figura 3.4 de la sección anterior. Es de esta manera en la cual cada señal se trató antes de medir su variación geométrica a lo largo del tiempo. Observando el comportamiento de las sub-bandas de señal, se comprueba que las altas frecuencias se encuentran en las primeras descomposiciones y van disminuyendo conforme se baja de nivel de descomposición.

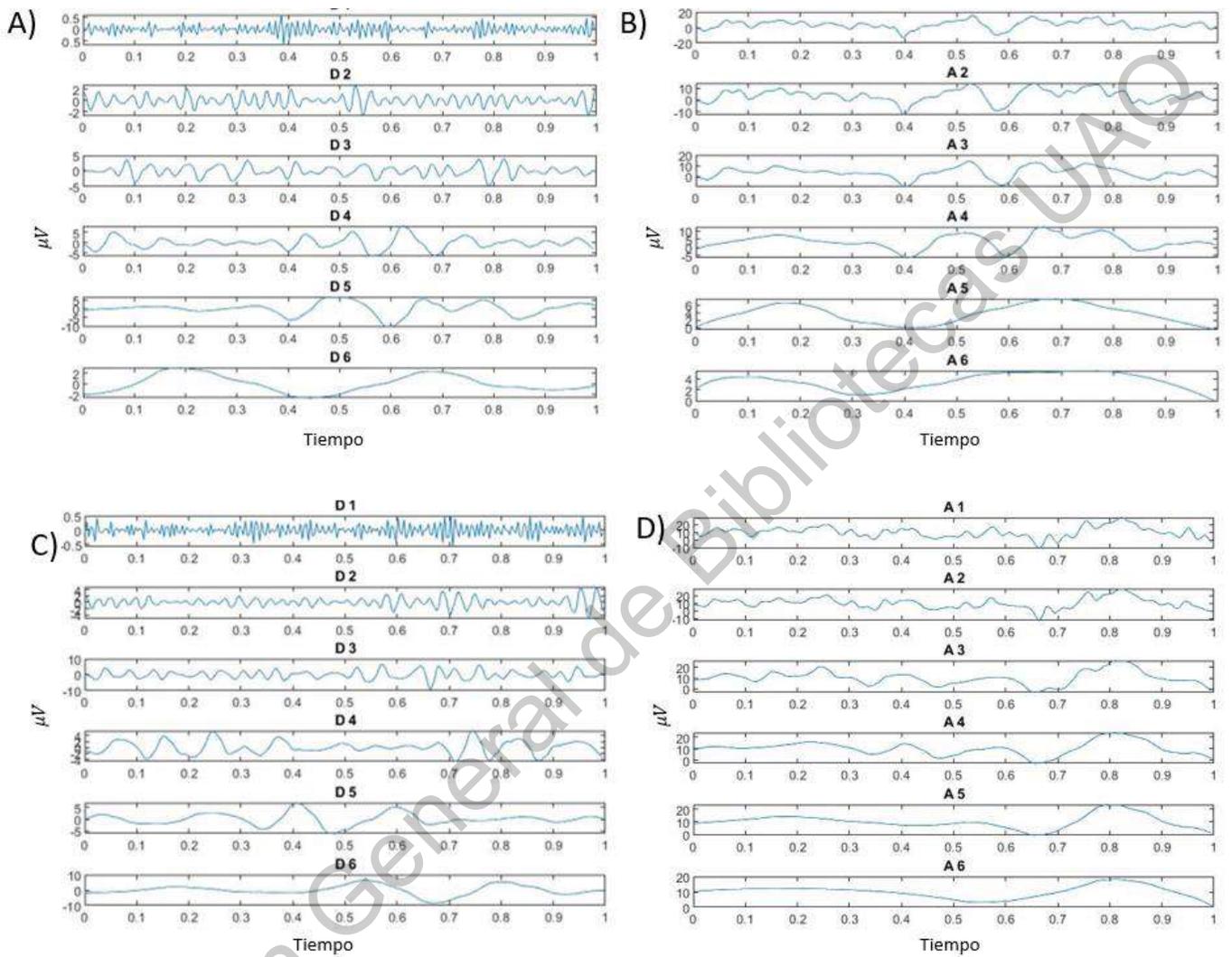


Figura 4.1. Representación gráfica de la descomposición de las señales de EEG. En esta imagen se muestran las descomposiciones obtenidas tras utilizar la DWT en la señal del electrodo 1 de la prueba 5 del paciente alcohólico y control número 9. Las letras 'A' indican aproximación y las letras 'D' indican detalle. Los gráficos del inciso A) y B) corresponden descomposiciones derivadas del EEG del paciente considerado alcohólico; mientras que los incisos C) y D) corresponden al paciente control.

Tabla 4.1. Valores de Fractalidad seleccionados tras la prueba estadística no paramétrica ANOVA de Kursk-Wallis.

Prueba	HFD	HFD	HFD	HFD	PET	PET
	D1	D1	A1	A2	D2	D1
	Electrodo 55	Electrodo 56	Electrodo 38	Electrodo 38	Electrodo 54	Electrodo 19
1	2.0137	2.0275	1.7241	1.6428	1.0259	1.0439
2	2.0192	2.0256	1.6614	1.5781	1.0253	1.0461
3	2.0213	2.0157	1.4876	1.4604	1.0251	1.0470
4	2.0233	2.0121	1.7571	1.7356	1.0261	1.0477
5	2.0192	2.0173	1.6970	1.6549	1.0251	1.0456
6	2.0212	2.0136	1.7488	1.7097	1.0269	1.0468
7	2.0227	2.0275	1.7143	1.6891	1.0274	1.0468
8	2.0177	2.0171	1.7361	1.7209	1.0269	1.0489
9	2.0269	2.0247	1.7167	1.6961	1.0282	1.0470
10	2.0277	2.0198	1.6824	1.6373	1.0269	1.0482
11	2.0172	2.0183	1.6997	1.6693	1.0259	1.0475
12	2.0229	2.0379	1.7301	1.6826	1.0253	1.0470
13	2.0307	2.0269	1.5634	1.4943	1.0277	1.0453
14	2.0270	2.0306	1.7426	1.7260	1.0266	1.0480
15	2.0173	2.0217	1.6424	1.6139	1.0274	1.0468
16	2.0225	2.0196	1.6104	1.5530	1.0289	1.0461
17	2.0187	2.0144	1.5728	1.5554	1.0259	1.0453
18	2.0193	2.0107	1.6681	1.6345	1.0274	1.0463
19	2.0247	2.0191	1.5459	1.5349	1.0259	1.0477
20	2.0179	2.0213	1.5803	1.5446	1.0261	1.0451
21	2.0212	2.0205	1.5632	1.5434	1.0282	1.0461
22	2.0035	2.0106	1.7094	1.6857	1.0261	1.0453
23	2.0187	2.0134	1.7730	1.7043	1.0269	1.0453
24	2.0140	2.0110	1.6986	1.6487	1.0253	1.0470
25	2.0108	2.0181	1.6076	1.5562	1.0277	1.0473
26	2.0068	2.0276	1.4876	1.4546	1.0271	1.0477
27	2.0060	2.0059	1.6814	1.6426	1.0292	1.0468
28	2.0147	2.0237	1.7222	1.7130	1.0256	1.0473
29	2.0244	2.0159	1.6021	1.5695	1.0253	1.0470
30	2.0185	2.0256	1.7304	1.6993	1.0271	1.0456

Con la descomposición de cada señal se consiguieron en total 552960 sub-bandas ($46080 \text{ señales} \times 12 \text{ descomposiciones}$). A cada una de ellas se le calculó la dimensión fractal con los 5 diferentes algoritmos establecidos, es decir 2,764,800 valores (552960×5). Para su mejor visualización y representación se exhiben en la Tabla 4.1 los primeros 30 valores de fractalidad de aquellas sub-bandas y algoritmos que muestran mayor rango de variación, de acuerdo con su respectiva prueba (estímulo S1 y S1-S2). Los 30 valores corresponden a un solo paciente. Como es leído en la metodología, esto se deduce tras la prueba estadística no paramétrica ANOVA de Kruskal-Wallis.

Volviendo a la Tabla 4.1, se muestra inicialmente como encabezado de cada columna el algoritmo de dimensión fractal utilizado, en seguida de qué sub-banda y electrodo muestran los datos y, por último, en el resto de las filas del lado izquierdo, cada uno de los valores obtenidos en cada una de las pruebas con el respectivo electrodo, sub-banda y algoritmo. Esta estructura también ayudará al lector a comprender el motivo del acomodo de los datos y la perspectiva que se tiene de ellos, que es buscar aquel electrodo, descomposición y algoritmo de fractalidad que entre las distintas pruebas proporcione mayor significancia para la diferenciación de clases de sujetos; conformando así las características para entrenar los modelos de machine learning.

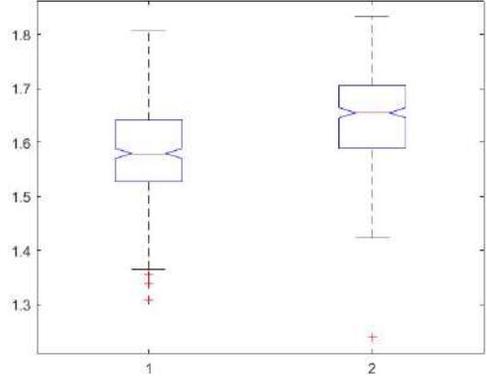
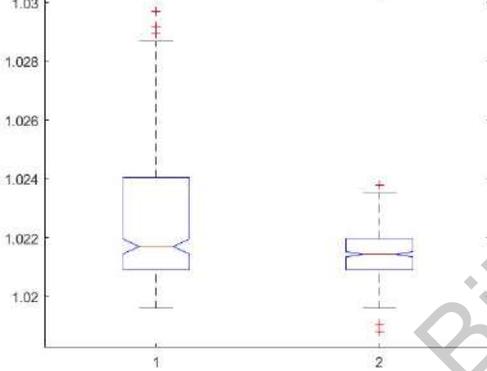
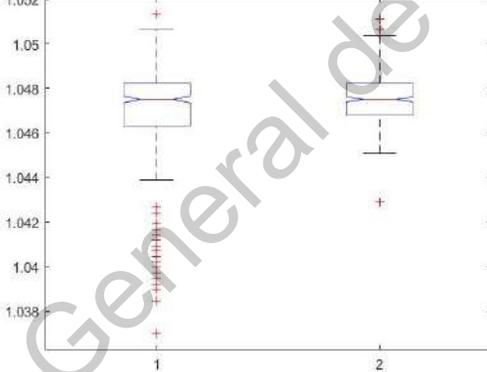
Continuando con la metodología, los datos se acomodaron como se describió en la sección 3.4 para continuar con lo descrito en la sección 3.5. Las gráficas de cajas y bigotes y valores de p relevantes para el trabajo en cuestión se presentan en las Tablas 4.2 y 4.3.

Tabla 4.2. p-values de cada uno de los algoritmos de fractalidad con mayor importancia para el método de clasificación, considerando electrodo y descomposición.

	HFD	HFD	HFD	HFD	PET	PET
	<i>D1</i>	<i>D1</i>	<i>A1</i>	<i>A2</i>	<i>D2</i>	<i>D1</i>
	<i>Electrodo 55</i>	<i>Electrodo 56</i>	<i>Electrodo 38</i>	<i>Electrodo 38</i>	<i>Electrodo 54</i>	<i>Electrodo 19</i>
<i>p-value</i>	$1.78538e^{-49}$	$2.40281e^{-39}$	$9.71443e^{-24}$	$1.16954e^{-21}$	$4.67965e^{-08}$	0.0717

Tabla 4.3. Principales diagramas de cajas y bigotes correspondientes a los mejores p-values.

Número	Diagrama de Cajas y Bigotes	Característica
1		HFD Electrodo 55 D1
2		HFD Electrodo 56 D1
3		HFD Electrodo 38 A1

4		<p>HFD Electrodo 38 D2</p>
5		<p>PET Electrodo 54 D2</p>
6		<p>PET Electrodo 19 D1</p>

Todas las anteriores se consideran pertinentes para ser características de entrenamiento para los modelos de aprendizaje, siendo las observaciones de la base de datos cada una de las pruebas por paciente. Cabe resaltar que la característica número 6 se consideró pertinente para el entrenamiento de los modelos a pesar de presentar una cantidad considerable de valores atípicos muy distantes del valor central. Esta zona que no comparte información con la clase número dos proporciona una nueva fuente información no compartida que ninguna otra característica aportó, ayudando a mejorar la eficiencia de clasificación de los

modelos. Considerando 30 pruebas por cada uno de los sujetos de cada clase suma una base de datos de 6 características con 720 observaciones, balanceado 50-50%.

Con la base de datos generada se puso en práctica el paso 3.6 la metodología descrita. En las Tablas 4.4 y 4.5 se muestran las características principales de cada modelo de entrenamiento y su respectiva matriz de confusión que demuestra su desempeño.

Tabla 4.4. Desempeño del Perceptrón Multicapa.

Detalles	Matriz de Confusión en Validación	Desempeño del Modelo						
<ul style="list-style-type: none"> • Función de entrenamiento: <i>Levenberg-Marquardt.</i> • Número de Capas: 3 <i>Capas</i> • Número de Características: 6 • Función de Transferencia: <i>Tangente Hiperbólica Sigmoides.</i> • Número de neuronas por capa: 10 • Entrenamiento: 80% <i>balanceado 50-50%</i> • Validación: 20% <i>balanceado 50-50%</i> 	<p style="text-align: center;">Valores Estimados</p> <table border="1"> <tr> <td rowspan="2" style="vertical-align: middle;">Valores Reales</td> <td style="text-align: center;">1</td> <td style="text-align: center;">2</td> </tr> <tr> <td style="text-align: center;">1</td> <td style="text-align: center;">2</td> </tr> </table>	Valores Reales	1	2	1	2	Especificidad	97.22%
			Valores Reales	1	2			
		1		2				
		Sensibilidad	69.44%					
		Precisión	96.15%					
Exactitud	83.33%							

Tabla 4.5. Desempeño de la Máquina de Vectores de Soporte

Detalles	Matriz de Confusión en Validación	Desempeño del Modelo												
<ul style="list-style-type: none"> • Optimización: <i>Grid search manual</i> • Kernel: <i>RBF</i> • Validación Cruzada: <i>5 particiones</i> • Número de características: <i>6</i> • Iteraciones: <i>256</i> • C-Value: <i>0.5</i> • Gama: <i>1024</i> • Entrenamiento: <i>80%</i> • Validación: <i>balanceado 50-50 %</i> 	<p style="text-align: center;">Valores Estimados</p> <table border="1"> <tr> <td rowspan="2" style="vertical-align: middle;">Valores Reales</td> <td style="text-align: center;">1</td> <td style="text-align: center;">68 47.22%</td> <td style="text-align: center;">1 0.69%</td> </tr> <tr> <td style="text-align: center;">2</td> <td style="text-align: center;">4 2.78%</td> <td style="text-align: center;">71 49.31%</td> </tr> <tr> <td></td> <td></td> <td style="text-align: center;">1</td> <td style="text-align: center;">2</td> </tr> </table>	Valores Reales	1	68 47.22%	1 0.69%	2	4 2.78%	71 49.31%			1	2	Especificidad	98.61%
			Valores Reales	1	68 47.22%	1 0.69%								
		2		4 2.78%	71 49.31%									
				1	2									
		Sensibilidad	94.44%											
Precisión	98.55%													
Exactitud	96.52%													

Por un lado, de acuerdo con la evaluación presentada en la Tabla 4.4, el perceptrón multicapa tuvo un porcentaje general de predicciones correctas aceptable, pues su exactitud es de 83.33%. El modelo es capaz de identificar con éxito pacientes control, pues tiene una especificidad de 97.22%; lamentablemente, su desempeño es deplorable para identificar correctamente la condición de alcoholismo en la base de datos, esto se deduce de su porcentaje de sensibilidad de 69.44%, pero la proporción de positivos reales correctamente identificados es alta con un porcentaje de 97.22%

Por otro lado, la máquina de vectores de soporte conforme a la Tabla 4.5, tuvo una exactitud considerablemente mayor que el perceptrón multicapa, con valor de 96.52%. El modelo de aprendizaje automático acorde a los resultados logra diferenciar las condiciones de ambos pacientes con gran desempeño, pues identifica alcohólicos con sensibilidad de 94.44% y pacientes control con especificidad de 98.61%.; incluso la proporción identificada correctamente de positivos al alcoholismo es superior con un valor de 98.55%.

Dados los resultados anteriores se puede concluir que la máquina de vectores de soporte es mejor algoritmo para diferenciar sujetos con predisposición a alcoholismo de sujetos control para la base de datos y características usadas, con un desempeño general superior al perceptrón multicapa.

En el siguiente apartado se compararán los resultados de este trabajo de tesis con el estado del arte de clasificación de señales de EEG que utilizan la misma base de datos.

4.2. Discusión de Resultados

Después de llevar a la práctica la metodología propuesta y revisar los resultados obtenidos se consiguió una exactitud de 96.52% con la máquina de vectores de soporte, sobreponiéndose al perceptrón multicapa en capacidad para diferenciar señales de EEG entre pacientes con predisposición alcohólica y control con la naturaleza de características implementadas. Dicha metodología es novedosa según la situación actual del conocimiento que se tiene, pues no ha sido propuesto método similar que utilice DWT y dimensión fractal con la base de datos utilizada. En seguida se presenta en la Tabla 4.6 un breve resumen producción científica más reciente que implementan metodologías para la clasificación de señales de EEG aplicada a la misma base de datos (<https://archive.ics.uci.edu/ml/datasets/EEG+Database>).

Tabla 4.6. Tabla comparativa de Metodologías

Autores	Metodología	Clasificadores	Características	Resultados
(Thilagaraj & Pallikonda Rajasekaran, 2019)	Adquisición de señales de EEG (base de datos). Utilizando 122 pacientes (sin especificar número de señales de EEG). Señales filtradas para remover artefactos. Descomposición por medio de EMD. Extracción de características y reducción por medio de Análisis de Componentes Independientes (ICA).	K vecinos más próximos ponderado con validación cruzada de 10 particiones	Coeficientes de IMFs asociados a los espectros de potencia, dominio de frecuencia y naturaleza multifacética.	Exactitud: 98.91% Sensibilidad: 99.02% Especificidad: 99.24%
(Rodrigues et al., 2019)	Se utilizó la transformada wavelet con cinco niveles de descomposición, empleando diferentes familias de	• Máquina de Vectores de Soporte	• Valor mínimo • Valor máximo	SVM db4 (75-25): • Exactitud: 97.93%

	<p>ondeletas (Biorotogonal, Coiflet, Daubechies y Symlets) para extraer características estadísticas de las descomposiciones.</p> <p>Experimentaron con la base de datos de 10 sujetos control y 10 sujetos alcohólicos, en total 600 señales.</p> <p>Compararon diferentes algoritmos de aprendizaje de máquinas.</p> <p>Se emplearon todas las sub-bandas resultantes de la descomposición y los porcentajes del total de la base de datos para el entrenamiento-prueba fueron: 25-75%, 50-50% y 75-25%.</p>	<ul style="list-style-type: none"> • Perceptrón Multicapa • Bosque de camino óptimo • Nave Bayes • K vecinos más próximos 	<ul style="list-style-type: none"> • Media • Desviación Estándar • Potencia • Media absoluta • Razón de media absoluta 	<ul style="list-style-type: none"> • Sensibilidad: 97.93 • Especificidad: 97.93 • MLP db4 (75-25): • Exactitud: 88% • Sensibilidad: 88% • Especificidad: 88%
(Siuly et al., 2019)	<p>En su metodología utilizan una técnica llamada 'muestreo basado en la asignación óptima' (OAS) para descubrir los puntos de datos más representativos de cada ventana de tiempo en los EEG, se usa la base de datos pequeña.</p> <p>Se extrajeron características estadísticas y se seleccionaron por medio del U-test de Mann-Whitney.</p> <p>Con las características extraídas se entrenaron modelos de machine learning.</p>	<ul style="list-style-type: none"> • Máquina de Vectores de Soporte • Tabla de decisión • Regresión Logística • K vecinos más próximos 	<ul style="list-style-type: none"> • Valor mínimo • Valor máximo • Media • Desviación Estándar • Mediana • Moda • Primer cuartil • Tercer cuartil • Rango intercuartil (IQR) • Curtosis • Asimetría 	<p>Con 11 Características</p> <p>SVM (Kernel Polinomial)</p> <ul style="list-style-type: none"> • Exactitud: 90.83% • Precisión: 90.90% <p>SVM (Kernel Lineal)</p> <ul style="list-style-type: none"> • Exactitud: 90.0% • Precisión: 90.0% <p>SVM (Kernel RBF)</p> <ul style="list-style-type: none"> • Exactitud: 70.83% • Precisión: 68.40%
(Anuragi & Sisodia, 2020)	<p>Los autores utilizan la transformada wavelet empírica empleando la transformada de Hilbert-Huang para extraer IMFs en el dominio tiempo-frecuencia.</p> <p>10 IMF fueron extraídas y a cada una de ellas se le extrajeron características estadísticas.</p> <p>Por medio de un estudio estadístico de significancia (<i>t-test</i>) seleccionaron 4 características.</p> <p>Con las características seleccionadas entrenaron modelos de machine learning para clasificar las señales utilizando validación Cruzada dejando uno afuera.</p> <p>Utilizan la base de datos pequeña.</p>	<ul style="list-style-type: none"> • Máquina de Vectores de Soporte mínimos cuadrados (LS-SVM) • Máquina de Vectores de Soporte • Clasificador ingenuo de Bayes • K vecinos más próximos 	<ul style="list-style-type: none"> • Media • Desviación Estándar • Varianza • Asimetría 	<p>LS-SVM RBF:</p> <ul style="list-style-type: none"> • Exactitud: 96.67% • Sensibilidad: 95.90% • Especificidad: 97.46% • Precisión: 97.50% <p>SVM Gaussiano:</p> <ul style="list-style-type: none"> • Exactitud: 90.46% • Sensibilidad: 87.60% • Especificidad: 93.75% • Precisión: 94.17%

<p>(Farsi et al., 2021)</p>	<p>En este trabajo se proponen dos algoritmos:</p> <p>Algoritmo 1: basado en redes neuronales profundas extrae los componentes principales los datos crudos y los componentes extraídos entrenan y validan una red neuronal</p> <p>Algoritmo 2: los datos crudos son directamente usadas en un modelo de aprendizaje profundo LSTM (Long Short-Term Memory, en inglés) que extrae características importantes en cada capa y las aplica como entrada del modelo de clasificación.</p>	<ul style="list-style-type: none"> • PCA + MLP • LSTM 	<ul style="list-style-type: none"> • Datos crudos tratados con PCA • Datos crudos 	<p>PCA + MLP:</p> <ul style="list-style-type: none"> • Exactitud: 86.00% • Sensibilidad: 85.00% • Especificidad: 84.00% <p>SVM Gaussiano:</p> <ul style="list-style-type: none"> • Exactitud: 93.00% • Sensibilidad: 95.00% • Especificidad: 92.00%
<p>(Agarwal et al., 2021)</p>	<p>En este trabajo se introduce el análisis deslizante de espectro singular (S-SSA) para descomponer y eliminar el ruido de señales de EEG.</p> <p>Como método de extracción de características se utiliza ICA. Los componentes seleccionados se utilizan para entrenar y validar modelos de machine learning.</p> <p>Se tomaron 10,000 muestras de alcohol y normal, para después dividirse en 50 canales con longitud de ventana (L=70).</p>	<ul style="list-style-type: none"> • Máquina de Vectores de Soporte • Perceptrón Multicapa • GBoost • AdaBoost • CGBoost • K vecinos más próximos 	<p>Componentes independientes (IC) extraídos tras utilizar S-SSA en las señales.</p>	<p>SVM con 5 características:</p> <ul style="list-style-type: none"> • Exactitud: 67.7% • Precisión: 72% <p>MLP con 5 características</p> <ul style="list-style-type: none"> • Exactitud: 76.53% • Sensibilidad: 79%

Después de revisar los resultados de diferentes trabajos de investigación con una diversidad de formidables metodologías implementadas, Tabla 4.6, se pueden resaltar los siguientes puntos de este trabajo de tesis con respecto al estado del arte:

- De acuerdo con la situación actual del conocimiento, ninguna metodología mide el cambio geométrico de la señal con dimensión fractal, ni utiliza alguno de estos algoritmos junto con DWT, por lo que el presente trabajo se traduce en un novedoso método para la identificación de alcoholismo por medio de EEG. Cabe recordar que la caracterización estadística de EEGs puede no ser la manera más adecuada de interpretar y caracterizar señales fisiológicas

con una naturaleza compleja y caótica, como se ha mencionado previamente.

- Conforme a lo descrito en los artículos científicos que describen claramente el manejo de la base de datos, el número de sujetos utilizado para llevar a cabo las fases de experimentación y prueba es menor a los 24 sujetos estudiados; ya sea por utilizar la base de datos pequeña (Anuragi & Sisodia, 2020; Siuly et al., 2019) o menor cantidad de datos (Agarwal et al., 2021; Farsi et al., 2021; Rodrigues et al., 2019; Thilagaraj & Pallikonda Rajasekaran, 2019). Así, el método presentado generaliza a mayor cantidad de señales de EEG.
- Sin lugar a dudas, hay trabajos que presentan resultados superiores en escasas cifras al desempeño de la SVM expuesta (Anuragi & Sisodia, 2020; Rodrigues et al., 2019), sin embargo contemplan menor número de sujetos al utilizar la bases de datos más pequeñas. En cuanto al resto de trabajos la SVM descrita sobre sale en su desempeño.
- En esta tesis también se compara el desempeño de dos métodos de machine learning y concuerda con el estado del arte en que las SVM tienen mayor eficacia para diferenciar entre EEG de sujetos control y predisposición alcohólica.
- En el presente trabajo además de contar con mayor número de sujetos, también se considera que las señales que se analicen por dimensión fractal hayan sido generadas bajo las mismas condiciones, es decir, se evalúan señales que fueron adquiridas con los mismos estímulos.

Dirección General de Bibliotecas UAQ

CAPÍTULO 5. CONCLUSIONES Y PROSPECTIVAS

5.1. Conclusiones

Incuestionablemente, el alcoholismo es un problema en la actualidad no solo en México (Infocop, 2018), sino que también en el resto del mundo (World Health Organization, 2018) y que ha prevalecido a lo largo de los años. Siendo muy severas las consecuencias derivadas del alcohol, como ser la principal causa de accidentes de tránsito en nuestro país (Ahumada-Cortez et al., 2017), los métodos vigentes para la detección del alcoholismo son ineficientes y desactualizados (Fleming, 1997).

Es por ello por lo que se presentó una novedosa metodología para la detección objetiva de predisposición al alcoholismo a través de señales de EEG. Utilizando la transformada wavelet discreta para descomponer estas señales en sus diferentes sub-bandas de frecuencia se posibilita la medición del cambio geométrico de la señal con naturaleza caótica en sus diversas descomposiciones por medio de dimensión fractal, generando datos que pudieran ser de gran significancia para el fenómeno estudiado y, una vez seleccionados por medio de un método estadístico no paramétrico, con ellos entrenar modelos de machine learning para la detección de alcoholismo en señales de la misma naturaleza.

Después de llevar a la práctica la metodología descrita con la base de datos detallada (Begleiter, 1999), se obtuvieron resultados satisfactorios para la detección de alcoholismo, logrando evaluar correctamente la condición en 94.44% y su ausencia en 98.61%, consiguiendo una eficacia general para la correcta distinción entre grupos de estudio del 96.52%, lo anterior con la máquina de vectores de soporte. Dichos resultados colocan al presente trabajo como un método formidable en el actual estado del arte para la detección de predisposición al alcoholismo por medio de procesamiento de señales y máquinas de vectores de soporte. Así, se sientan las bases para que en un futuro el diagnóstico de alcoholismo a nivel nacional se realice objetivamente y de esta manera poder disminuir las consecuencias de esta enfermedad.

5.2. Trabajos a Futuro

A pesar de presentar un trabajo con excelentes resultados, se encontraron áreas de oportunidad para pulir el método:

- Se propone mejorar la afinación del MLP, ya sea por selección de parámetros, número de capas ocultas y de neuronas o por medio otras funciones de activación y de entrenamiento.
- Se descartó una gran cantidad de información y es por ello por lo que se propone hallar la manera de integrar dichos sujetos filtrados a la base de datos final; señales generadas bajo las mismas condiciones.
- La DWT pierde información en las altas frecuencias, de ahí que se sugiera utilizar la WPD para explorar otras sub-bandas de frecuencia y con suerte, descubrir nuevas descomposiciones que sean de gran significancia para identificar la condición estudiada.

Lo que anteriormente se expuso sienta las bases para el diagnóstico objetivo del alcoholismo, por lo que se abren las puertas a futuras investigaciones y desarrollos:

- Aun cuando hay una extensa variedad de modelos de machine learning, únicamente se evalúan dos de ellos. Por ese motivo, se podría investigar otros modelos que se utilizan en el estado del arte que se ajusten adecuadamente con la metodología y explorar su desempeño.
- El tiempo de respuesta fue ignorado y es de suma importancia para llevarlo a un dispositivo real; por ello evaluar el tiempo de diagnóstico es incentivo para nuevos trabajos.
- Se propone generar una nueva base de datos donde se controle la experimentación y se documente adecuadamente el modo de adquisición de las señales, para así evitar la falta de esta información que se menciona en el apartado 3.2 de la metodología.
- Por último, junto con los dos puntos anteriores, no solo desarrollar un dispositivo de adquisición de señales de EEG para el diagnóstico de alcoholismo, sino también, probar la metodología propuesta en el hardware diseñado con la intención de generalizar a una mayor cantidad de sujetos.

BIBLIOGRAFÍA

- Agarwal, S., Zubair, M., & Member, S. (2021). Classification of Alcoholic and Non-Alcoholic EEG Signals Based on Sliding-SSA and Independent Component Analysis Classification of Alcoholic and Non-Alcoholic EEG Signals Based on Sliding-SSA and Independent Component Analysis. *TechRxiv*, 0–8. <https://doi.org/10.36227/techrxiv.14837226.v1>
- Ahumada-Cortez, J. G., Enrique Gámez-Medina, M., & Valdez-Montero, C. (2017). EL CONSUMO DE ALCOHOL COMO PROBLEMA DE SALUD PÚBLICA ALCOHOL CONSUMPTION AS A PUBLIC HEALTH PROBLEM. *Ra Ximhai*, 13(2), 13–24. <https://www.redalyc.org/articulo.oa?id=46154510001>
- Al-Qazzaz, N. K., Mohd Ali, S. H. Bin, Ahmad, S. A., Islam, M. S., & Escudero, J. (2015). Selection of mother wavelet functions for multi-channel EEG signal analysis during a working memory task. *Sensors (Switzerland)*, 15(11), 29015–29035. <https://doi.org/10.3390/s151129015>
- Amat, J. (2017, April). *Máquinas de Vector Soporte (Support Vector Machines, SVMs)*. https://www.cienciadedatos.net/documentos/34_maquinas_de_vector_soporte_support_vector_machines
- Anuragi, A., & Sisodia, D. S. S. (2020). Empirical wavelet transform based automated alcoholism detecting using EEG signal features. *Biomedical Signal Processing and Control*, 57, 101777. <https://doi.org/10.1016/j.bspc.2019.101777>
- Arias Duque, R. (2005). Reacciones fisiológicas y neuroquímicas del alcoholismo. *Diversitas*, 1(2), 138. <https://doi.org/10.15332/s1794-9998.2005.0002.02>
- Awasthi, S. (2020, December 17). *Seven Most Popular SVM Kernels*. <https://dataaspirant.com/svm-kernels/>
- Bae, Y., Yoo, B. W., Lee, J. C., & Kim, H. C. (2017). Automated network analysis to measure brain effective connectivity estimated from EEG data of patients with alcoholism. *Physiological Measurement*, 38(5), 759–773. <https://doi.org/10.1088/1361-6579/aa6b4c>
- Bajaj, N. (2021). Wavelets for EEG Analysis. In *Wavelet Theory*. IntechOpen. <https://doi.org/10.5772/intechopen.94398>
- Bavkar, S., Iyer, B., & Deosarkar, S. (2019). Detection of alcoholism: An EEG hybrid features and ensemble subspace K-NN based approach. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 11319 LNCS*. Springer International Publishing. https://doi.org/10.1007/978-3-030-05366-6_13
- Begleiter, H. (1999). *UCI Machine Learning Repository: EEG Database Data Set*. <https://archive.ics.uci.edu/ml/datasets/EEG+Database>

- Begleiter, H., & Porjesz, B. (1999). What Is Inherited in the Predisposition Toward Alcoholism? A Proposed Model. *Alcoholism: Clinical and Experimental Research*, 23(7), 1125–1135. [https://doi.org/https://doi.org/10.1111/j.1530-0277.1999.tb04269.x](https://doi.org/10.1111/j.1530-0277.1999.tb04269.x)
- Blum, K., Noble, E. P., Sheridan, P. J., Montgomery, A., Ritchie, T., Ozkaragoz, T., Fitch, R. J., Wood, R., Finley, O., & Sadlack, F. (1993). Genetic predisposition in alcoholism: Association of the D2 dopamine receptor TaqI B1 RFLP with severe alcoholics. *Alcohol*, 10(1), 59–67. [https://doi.org/10.1016/0741-8329\(93\)90054-R](https://doi.org/10.1016/0741-8329(93)90054-R)
- Bravo-Grau, S., Cruz, P., & Q, J. (2015). Estudios de exactitud diagnóstica: Herramientas para su Interpretación Diagnostic accuracy studies: Tools for its Interpretation. *Revista Chilena de Radiología. Año*, 21(4), 158–164.
- Brunton, S., & Kutz, N. (2017). *Data Driven Science & Engineering Machine Learning, Dynamical Systems, and Control*. <http://databookuw.com/databook.pdf>
- Burns, T., & Rajan, R. (2015). Combining complexity measures of EEG data: Multiplying measures reveal previously hidden information. *F1000Research*, 4(0), 1–7. <https://doi.org/10.12688/f1000research.6590.1>
- Caraballo, S., Del Pilar, Y., Alfonso, C., & Yuri, A. (2017). *ESTUDIO DE REVISIÓN SOBRE LOS FACTORES GENÉTICOS Y EPIGENÉTICA DEL ALCOHOLISMO* [Universidad Distrital Francisco José de Caldas]. <https://repository.udistrital.edu.co/handle/11349/6769>
- Castillero Mimenza, O. (n.d.). *Neuronas piramidales: funciones y localización en el cerebro*. Retrieved May 21, 2021, from <https://psicologiyamente.com/neurociencias/neuronas-piramidales>
- CEVECE. (2012). *Accidentes y alcoholismo*. 1–9.
- Corredor, C. (2017). Metabolismo del Alcohol. *Colombia Médica*, 2(1), 29–33. <https://doi.org/doi:10.2510/colomb.med.v2i1.3219>
- Costanzo, L. S. (2011). *Fisiología* (4th ed.). Elsevier. <https://books.google.com.mx/books?id=7iOR5x3tKg0C>
- Costas, J. (2015). Un abordaje genómico del alcoholismo. *Adicciones*, 27(2), 87. <https://doi.org/10.20882/adicciones.693>
- Di Ieva, A. (2016). *The Fractal Geometry of the Brain*. <https://doi.org/10.1007/978-1-4939-3995-4>
- Díaz-soto, M. T., & Calderín-M, iranda, J. M. (2020). Síndrome de abstinencia alcohólica: Resultado del estrés oxidativo y desequilibrio neuronal. Estado del arte. *Revista Biomédica*.
- Eke, A., Herman, P., Kocsis, L., & Kozak, L. R. (2002). Fractal characterization of complexity in temporal physiological signals. *Physiological Measurement*, 23(1).

<https://doi.org/10.1088/0967-3334/23/1/201>

- Escobar, J. J. M., Matamoros, O. M., Padilla, R. T., Hernández, L. C., Durán, J. P. F. P., Martínez, A. K. P., Reyes, I. L., & Espinosa, H. Q. (2020). Biomedical signal acquisition using sensors under the paradigm of parallel computing. *Sensors (Switzerland)*, *20*(23), 1–36. <https://doi.org/10.3390/s20236991>
- Farsi, L., Siuly, S., Kabir, E., & Wang, H. (2021). Classification of Alcoholic EEG Signals Using a Deep Learning Method. *IEEE Sensors Journal*, *21*(3), 3552–3560. <https://doi.org/10.1109/JSEN.2020.3026830>
- Faust, O., Acharya, R. U., Allen, A. R., & Lin, C. M. (2008). Analysis of EEG signals during epileptic and alcoholic states using AR modeling techniques. *ITBM-RBM*, *29*(1), 44–52. <https://doi.org/10.1016/j.rbmret.2007.11.003>
- Faust, Oliver, Hagiwara, Y., Hong, T. J., Lih, O. S., & Acharya, U. R. (2018). Deep learning for healthcare applications based on physiological signals: A review. *Computer Methods and Programs in Biomedicine*, *161*, 1–13. <https://doi.org/10.1016/j.cmpb.2018.04.005>
- Faust, Oliver, Yu, W., & Kadri, N. A. (2013). Computer-based identification of normal and alcoholic EEG signals using wavelet packets and energy measures. *Journal of Mechanics in Medicine and Biology*, *13*(3). <https://doi.org/10.1142/S0219519413500334>
- Fenotipo | NHGRI. (n.d.). Retrieved May 20, 2021, from <https://www.genome.gov/es/genetics-glossary/Fenotipo>
- Fleming, M. F. (1997). Strategies to increase alcohol screening in health care settings. *Alcohol Research and Health*, *21*(4), 340–347. [/pmc/articles/PMC6827677/](https://pubmed.ncbi.nlm.nih.gov/6827677/)
- Fuente, J. R. D. la, & Kershenobich, D. (1987). Detección oportuna del paciente alcohólico y de sus alteraciones hepáticas. *Salud Mental*, *10*(4), 76–80. http://www.revistasaludmental.mx/index.php/salud_mental/article/view/320
- G., M. M., C., G., & N., M. C. (2016). Alcohol, cirrosis y predisposición genética. *Revista Colombiana de Gastroenterología*, *31*, 27–35. <https://www.redalyc.org/articulo.oa?id=337745613005>
- Gandhi, T., Panigrahi, B. K., & Anand, S. (2011). A comparative study of wavelet families for EEG signal classification. *Neurocomputing*, *74*(17), 3051–3057. <https://doi.org/10.1016/j.neucom.2011.04.029>
- Gao, R. X., & Yan, R. (2011). Wavelets: Theory and applications for manufacturing. In *Wavelets: Theory and Applications for Manufacturing*. Springer US. <https://doi.org/10.1007/978-1-4419-1545-0>
- García González, T., Jiménez González, A., Ortiz Pedroza, M. del R., & Peña Castillo, M. Á. (1998). *Potenciales bioeléctricos: origen y registro* (1st ed.). Universidad Autónoma Metropolitana.

- Gaviria, C., Marcela, M., Arango, C., Navas, N., & Cristina, M. (2016). Alcohol , cirrosis y predisposición genética. *Revista Colombiana de Gastroenterología*, 31(1), 27–35.
- Guerri, C. (2000). Cómo actúa el alcohol en nuestro cerebro. *Trastornos Adictivos*, 2(1), 14–25. <http://www.elsevier.es/es-revista-trastornos-adictivos-182-articulo-como-actua-el-alcohol-nuestro-10016452>
- Guyton, A. C., & Hall, J. E. (2016). *Guyton y Hall, tratado de fisiología médica* (13^a ed.). Elsevier.
- Handy, T. C. (2005). *Event-related Potentials: A Methods Handbook*. MIT Press. <https://books.google.com.mx/books?id=OQyZefgEzRUC>
- Haykin, S. (2008). Neural Networks and Learning Machines. In *Pearson Prentice Hall New Jersey USA 936 pLinks* (Third, Vol. 3). <https://doi.org/978-0131471399>
- Haykin, S. O. (2011). *Neural Networks and Learning Machines*. Pearson Education. <https://books.google.com.mx/books?id=faouAAAAQBAJ>
- Hesselbrock, V. (1995). The genetic Epidemiology of alcoholism, in Begleiter H, Kissin B (eds): Alcohol and Alcoholism. *The Genetics of Alcoholism*, 1, 17–39.
- Higuchi, S., Matsushita, S., Masaki, T., Yokoyama, A., Kimura, M., Go, S., & Mochizuki, H. (2004). Influence of genetic variations of ethanol-metabolizing enzymes on phenotypes of alcohol-related disorders. *Annals of the New York Academy of Sciences*, 1025(1), 472–480. <https://doi.org/10.1196/annals.1316.058>
- Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena*, 31(2), 277–283. [https://doi.org/10.1016/0167-2789\(88\)90081-4](https://doi.org/10.1016/0167-2789(88)90081-4)
- Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). *A Practical Guide to Support Vector Classification*. <http://www.csie.ntu.edu.tw/~cjlin>
- IMSS. (2015, July 15). *Alcoholismo*. <http://www.imss.gob.mx/salud-en-linea/alcoholismo>
- Infocop. (2018). Nuevo informe sobre la situación mundial del alcohol y la salud 2018. *Infocop Online*, 6–8. http://www.infocop.es/view_article.asp?id=7690
- Joaquín Amat, R. (2016). ANOVA análisis de varianza para comparar múltiples medias. 1–49. https://www.cienciadedatos.net/documentos/19_anova
- Katz, M. J. (1988). Fractals and the analysis of waveforms. *Computers in Biology and Medicine*, 18(3), 145–156. [https://doi.org/10.1016/0010-4825\(88\)90041-8](https://doi.org/10.1016/0010-4825(88)90041-8)
- Kershenobich, D. (2010). Alcohol y alcoholismo: definiciones actuales, mecanismos de daño y tratamiento clínico. *Revista de Gastroenterología de México*, 75(0375), 177–178.

- Lindsay, R. W., Percival, D. B., & Andrew Rothrock, D. (1996). The discrete wavelet transform and the scale analysis of the surface properties of sea ice. *IEEE Transactions on Geoscience and Remote Sensing*, 34(3), 771–787. <https://doi.org/0.1109/36.499782>
- Mandal, A. (2019, April 9). *las funciones de la dopamina*. [https://www.news-medical.net/health/Dopamine-Functions-\(Spanish\).aspx](https://www.news-medical.net/health/Dopamine-Functions-(Spanish).aspx)
- Méndez Plazas, C. B. (2009). *Tipificación de polimorfismos de los genes alcohol deshidrogenasa, aldehído deshidrogenasa y citocromo P4502E1 en una población de la ciudad de Bogotá y la evaluación de su posible relación con el consumo problemático del alcohol* [Universidad Nacional de Colombia]. <https://repositorio.unal.edu.co/handle/unal/63018>
- Minitab. (n.d.). *Explicación de los métodos no paramétricos - Minitab*. Retrieved May 31, 2021, from <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/statistics/nonparametrics/supporting-topics/understanding-nonparametric-methods/>
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill Science.
- Mohanchandra, K., Saha, S., & Azar, A. T. (2016). Evidence of chaos in EEG signals: An application to BCI. In *Studies in Fuzziness and Soft Computing* (Vol. 337, pp. 609–625). Springer Verlag. https://doi.org/10.1007/978-3-319-30340-6_25
- Moreno-Gomez, A., MacHorro-Lopez, J. M., Amezquita-Sanchez, J. P., Perez-Ramirez, C. A., Valtierra-Rodriguez, M., & Dominguez-Gonzalez, A. (2020). Fractal Dimension Analysis for Assessing the Health Condition of A Truss Structure Using Vibration Signals. *Fractals*, 28(7). <https://doi.org/10.1142/S0218348X20501273>
- Neuron, A., Hammer, B., Iliadis, L., Hutchison, D., Caterini, A. L., Chang, D. E., Iba, H., Aggarwal, C. C., Zheng, A., Jin, J., Cornally, P., Butler, M., Murphy, M., Rath, A., Canty, G., & Wenzl, J. (2018). and Machine Learning – ICANN 2018. In *Evidence Based Midwifery*.
- Nunez, P. L., Nunez, E. P. B. E. P. L., Srinivasan, R., Press, O. U., & Srinivasan, A. P. C. S. R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press. https://books.google.com.mx/books?id=fUv54as56%5C_8C
- Nunez, P., & Srinivasan, R. (2006). *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press. https://books.google.com/books?hl=es&lr=&id=fUv54as56_8C&oi=fnd&pg=PR11&dq=electrical+field+theory+eeg+sources&ots=nZQc4TIJIT&sig=qczb6cPh3uc74UwfHSyAl_6Ngyo
- Ortiz González, J. P., & Reinoso Avecillas, M. J. (2010). *Diseño y construcción de un prototipo de electroencefalograma para adquisición de señales cerebrales* [UNIVERSIDAD POLITÉCNICA SALESIANA]. <https://dspace.ups.edu.ec/bitstream/123456789/417/15/UPS-CT001873.pdf>

- Osorio, L. A. (2007). *Acondicionamiento de Señales Bioeléctricas*.
- Pandey, A. K., Rathore, P. S., & Balamurugan, S. (2019). *A Practical Approach for Machine Learning and Deep Learning Algorithms: Tools and Techniques Using MATLAB and Python*. BPB Publications. <https://books.google.com.mx/books?id=1z6wDwAAQBAJ>
- Petrosian, A. (1995). *CBMS '95: Proceedings of the Eighth Annual IEEE Symposium on Computer-Based Medical Systems*.
- Planas-Ballvé, A., Grau-López, L., Morillas, R. M., & Planas, R. (2017). Neurological manifestations of excessive alcohol consumption. *Gastroenterología y Hepatología*, 40(10), 709–717. <https://doi.org/10.1016/j.gastrohep.2017.05.011>
- Polimorfismos de nucleótido único (SNPs) | NHGRI. (n.d.). Retrieved May 20, 2021, from <https://www.genome.gov/es/genetics-glossary/Polimorfismos-de-nucleotido-unico>
- Quiñones, J. G. (n.d.). *Amplificador Operacional Real*. Retrieved May 25, 2021, from https://www.mty.itesm.mx/etie/deptos/ie/profesores/jgomez/eap/opamp_real.pdf
- Quispe Andía, A., Calla Vasquez, K., Yangali Vicente, J., Rodríguez López, J., & Pumacayo Palomino, I. (2019). *Estadística no paramétrica aplicada a la investigación científica con software SPSS, MINITAB Y EXCEL* (1st ed.). EIDEC. <https://www.editorialeidec.com/product/estadistica-no-parametrica-aplicada-a-la-investigacion-cientifica-con-software-spss-minitab-y-excel/>
- Rachman, N. T., Tjandrasa, H., & Fatichah, C. (2017). Alcoholism classification based on EEG data using Independent Component Analysis (ICA), Wavelet denoising and Probabilistic Neural Network (PNN). *Proceeding - 2016 International Seminar on Intelligent Technology and Its Application, ISITIA 2016: Recent Trends in Intelligent Computational Technologies for Sustainable Energy*, 17–20. <https://doi.org/10.1109/ISITIA.2016.7828626>
- Recuerdo, P. (2018, January 23). *Machine Learning a tu alcance: La matriz de confusión - Think Big Empresas*. <https://empresas.blogthinkbig.com/ml-a-tu-alcance-matriz-confusion/>
- Rey Buitrago, M. (2015). Genética Molecular Del Alcoholismo. *Revista de La Facultad de Medicina*, 63(3), 483–494. <https://doi.org/10.15446/revfacmed.v63n3.49352>
- Rioul, O., & Vetterli, M. (1991). Wavelets and Signal Processing. *IEEE Signal Processing Magazine*, 8(4), 14–38. <https://doi.org/10.1109/79.91217>
- Rodrigues, J. das C., Filho, P. P. R., Peixoto, E., N, A. K., & de Albuquerque, V. H. C. (2019). Classification of EEG signals to detect alcoholism using machine learning techniques. *Pattern Recognition Letters*, 125, 140–149. <https://doi.org/10.1016/j.patrec.2019.04.019>
- Sáiz, P. A., G-Portilla, M. P., Paredes, M. B., Bascarán, M. T., & Bobes, J. (2002).

- Instrumentos de evaluación en alcoholismo. *Adicciones*, 14(SUPPL. 1), 387–403. [https://doi.org/10.1016/s1575-0973\(03\)70101-9](https://doi.org/10.1016/s1575-0973(03)70101-9)
- Salis, C. I., Malissovass, A. E., Bizopoulos, P. A., Tzallas, A. T., Angelidis, P. A., & Tsalikakis, D. G. (2013). Denoising simulated EEG signals: A comparative study of EMD, wavelet transform and Kalman filter. *13th IEEE International Conference on Bioinformatics and BioEngineering, IEEE BIBE 2013*. <https://doi.org/10.1109/BIBE.2013.6701613>
- SEFO. (2019). *Pruebas no paramétricas*. <https://www.scientific-european-federation-osteopaths.org/wp-content/uploads/2019/01/ALGUNAS-PRUEBAS-NO-PARAMÉTRICAS.pdf>
- Sevcik, C. (2010). A procedure to Estimate the Fractal Dimension of Waveforms. *Complexity International*, 5. <http://arxiv.org/abs/1003.5266>
- Sharma, M., & Pachori, R. B. (2017). A novel approach to detect epileptic seizures using a combination of tunable-Q wavelet transform and fractal dimension. *Journal of Mechanics in Medicine and Biology*, 17(7), 1–20. <https://doi.org/10.1142/S0219519417400036>
- Silverthorn, D. U., Johnson, B. R., Ober, W. C., Garrison, C. W., & Silverthorn, A. C. (2013). *Human Physiology: An Integrated Approach* (6th ed.). Pearson Education. <https://books.google.com.mx/books?id=UejvkQEACAAJ>
- Siuly, S., Bajaj, V., Sengur, A., & Zhang, Y. (2019). An Advanced Analysis System for Identifying Alcoholic Brain State Through EEG Signals. *International Journal of Automation and Computing*, 16(6), 737–747. <https://doi.org/10.1007/s11633-019-1178-7>
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 174–215. <https://doi.org/10.1037/0278-7393.6.2.174>
- Stavrinidis, G., Michelakis, K., Kontomitrou, V., Giannakakis, G., Sevrisianos, M., Sevrisianos, G., Chaniotakis, N., Alifragis, Y., & Konstantinidis, G. (2016). SU-8 microneedles based dry electrodes for Electroencephalogram. *Microelectronic Engineering*, 159, 114–120. <https://doi.org/10.1016/j.mee.2016.02.062>
- Sun, Y., Ye, N., & Xu, X. (2006). EEG analysis of alcoholics and controls based on feature extraction. *International Conference on Signal Processing Proceedings, ICSP*, 1. <https://doi.org/10.1109/ICOSP.2006.344501>
- Taran, S., & Bajaj, V. (2018). Rhythm-based identification of alcohol EEG signals. *IET Science, Measurement & Technology*, 12(3), 343–349. <https://doi.org/10.1049/iet-smt.2017.0232>
- Teschke, R. (2018). Alcoholic liver disease: Alcohol metabolism, cascade of molecular mechanisms, cellular targets, and clinical aspects. *Biomedicines*,

6(4). <https://doi.org/10.3390/biomedicines6040106>

Thilagaraj, M., & Pallikonda Rajasekaran, M. (2019). An empirical mode decomposition (EMD)-based scheme for alcoholism identification. *Pattern Recognition Letters*, 125, 133–139. <https://doi.org/10.1016/j.patrec.2019.03.010>

UNAM. (2019). *MANUAL SISTEMA 10/20 INTERNACIONAL ANEXO*.

Villatoro-Velázquez, J. A., Reséndiz, E. E., Mujica, S. A., Bretón-Cirret, M., Cañas-Martínez, V., Soto-Hernández, I., Fregosso-Ito, D., Fleiz-Bautista, C., Medina-Mora, M., Gutiérrez, J., Franco-Nuñez, A., Romero, M., & Mendoza, L. (2017). Encuesta Nacional de Consumo de Drogas, Alcohol y Tabaco. *Secretaria de Salud*, 1, 190. <https://www.gob.mx/salud%7Cconadic/acciones-y-programas/encuesta-nacional-de-consumo-de-drogas-alcohol-y-tabaco-encodat-2016-2017-136758>

Wabnitz, H., Sander, T., & Klaus-robert, M. (2017). M3BA : A Mobile , Modular , Multimodal Biosignal Acquisition Architecture for Miniaturized. *IEEE Transactions on Biomedical Engineering*, 64(6), 1199–1210.

Wang, B. (2005). *Detection of structural damage using fractal dimension technique*. 24, 87-88+110.

World Health Organization. (2018). Informe sobre la situación mundial del alcohol y la salud. *Organización Panamericana de La Salud*, 28. http://iris.paho.org/xmlui/bitstream/handle/123456789/51352/OPSNMH19012_spa.pdf?sequence=1&isAllowed=y

Zhang, X. Da. (2020). A matrix algebra approach to artificial intelligence. In *A Matrix Algebra Approach to Artificial Intelligence*. Springer Singapore. <https://doi.org/10.1007/978-981-15-2770-8>

Zhou, J., Schalkoff, R. J., Dean, B. C., & Halford, J. J. (2012). Morphology-based wavelet features and multiple mother wavelet strategy for spike classification in EEG signals. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 3959–3962. <https://doi.org/10.1109/EMBC.2012.6346833>