

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería



**Aplicación sistemática de algoritmos de aprendizaje
automático y reconocimiento de patrones a señales
bioeléctricas obtenidas de HUMANA y con el sistema Biopac**

Protocolo de trabajo de graduación presentado por Fernando Javier
Sandoval Ruballos, estudiante de Ingeniería Mecatrónica

Guatemala,

2023

Resumen

La presente propuesta de trabajo se basa en la implementación de la inteligencia artificial enfocada en el desarrollo de algoritmos, permitiendo a las computadoras aprender y mejorar a través de la experiencia. Particularmente en el campo de la medicina, los algoritmos de aprendizaje automático permiten la extracción de características de señales bioeléctricas de manera automática, mejorando la eficacia en el diagnóstico médico de afecciones neuronales como la epilepsia, entre otras. El objetivo del proyecto es aplicar algoritmos de aprendizaje automático desarrollados en fases anteriores a una mayor cantidad de señales bioeléctricas, utilizando bases de datos proporcionadas por HUMANA y las obtenidas dentro del departamento con los nuevos sistemas de captura Biopac.

Para la propuesta de desarrollo del trabajo, y con la adición de los sistemas de captura Biopac al departamento, se pretende recolectar señales de manera sistemática y en gran volumen. Con la ayuda de los algoritmos desarrollados en etapas previas, comparar el funcionamiento, eficacia y exactitud de los mismos utilizando una base de datos más robusta, así mismo integrando el uso del software Biopac en las interfaces biomédicas desarrolladas previamente. Con el fin de volver más eficiente los diagnósticos médicos, se busca validar los algoritmos de aprendizaje desarrollados en fases anteriores con especialistas de HUMANA y una base de datos más completa.

Antecedentes

Entre los trastornos neurológicos más comunes a nivel global se encuentra la epilepsia, la cual representa un 6 % de la carga global de enfermedades. Según datos de la OPS en 2021, se registra que alrededor de 60 millones de personas a nivel mundial sufren de epilepsia; cifra que se proyecta duplicar cada 20 años. En Latinoamérica y el Caribe se destaca la mayor prevalencia de la afección neurológica, donde se estima viven 5 millones de personas afectadas por cuadros epilépticos [1].

En Guatemala se encuentra consolidado el Centro de Epilepsia y Neurocirugía Funcional, HUMANA. El centro HUMANA estima que 325,000 personas sufren de epilepsia a nivel nacional, con un 30 % siendo pacientes que presentan un cuadro clínico de difícil control. El diagnóstico de la afección y su categorización no es posible de realizar mediante observación directa, por lo que se implementa el análisis de señales electroencefalográficas, EEG. Las anotaciones de una señal EEG son realizadas por un profesional que resalta los segmentos de interés, sin embargo, este procedimiento puede resultar largo y tedioso (24-72 hrs) [2].

Durante la última década, una gran cantidad de investigaciones se han enfocado en el procesamiento de señales y en la implementación de métodos para la extracción de características de las mismas. Al inicio se creía que las señales EEG eran valores aleatorios; sin embargo, se han descubierto diversos modelos matemáticos en el dominio del tiempo, la frecuencia y tiempo-frecuencia, los cuales son capaces de clasificar y distinguir patrones en las señales [3].

Con el objetivo de crear procesos innovadores que ayuden a personas con este tipo de afecciones, se ha abierto el espacio a la creación de aplicaciones que permiten la manipulación

de sistemas utilizando datos obtenidos de señales bioeléctricas. En la Universidad del Valle de Guatemala se han desarrollado herramientas que permiten explorar diferentes aplicaciones a partir de la información extraída de estas señales. En el año 2020, con el trabajo de graduación presentado por María Fernanda Girón [4], se implementó una herramienta que permite utilizar señales electromiográficas (EMG) para el control de sistemas robóticos. En el trabajo, se implementaron algoritmos de aprendizaje supervisado que permiten la clasificación de movimientos o gestos a partir de características extraídas de señales EMG. El mejor resultado obtenido fue a través de una Red Neuronal Artificial, con rendimiento superior al 90 % clasificando específicamente 4 movimientos. Adicionalmente, se observó que utilizando únicamente características en el dominio del tiempo fue suficiente para obtener rendimientos mayores al 80 %. Rodrigo Ralda con su trabajo de graduación [5], desarrolló la primera iteración de una interfaz que permite la manipulación de sistemas robóticos utilizando señales electroencefalográficas (EEG). En este trabajo se implementó la técnica de aprendizaje supervisado, utilizando Redes Neuronales Artificiales y Máquinas de Vectores de Soporte para encontrar una combinación con menor error al seleccionar características, el mejor rendimiento obtenido fue mayor al 80 %, logrando así manipular un sistema robótico con señales EEG.

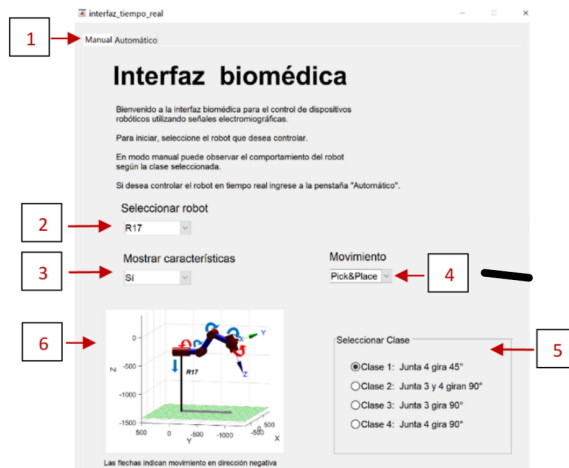


Figura 1: Interfaz de control de sistema robótico con señales EMG [4].

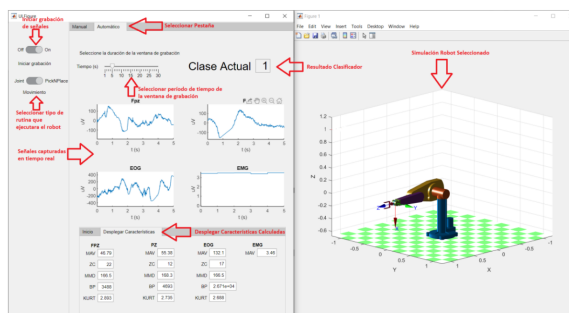


Figura 2: Interfaz de control de sistema robótico con señales EEG [5].

Hasta el año 2020, las interfaces biomédicas no poseían correlación con respecto a las señales utilizadas en funcionamiento, el trabajo posterior presentado por Roberto José Cáceres

García en el año 2021 [6] implementó la unificación de las interfaces, además de presentar control de sistemas robóticos utilizando la combinación de señales biomédicas EMG y EEG. En este trabajo se utilizó entrenamiento con 4 clases combinando las señales, que funcionaron parecido a dos clasificadores distintos. Esto logró un rendimiento superior al 91 %, también se realizaron pruebas grabando 2 clases distintivas para las señales EMG y usando las señales EEG como auxiliares, obteniendo un 87 % de rendimiento, finalmente utilizando características de la señal EMG únicamente se logró un 85 % de rendimiento. Los trabajos anteriormente expuestos nos dan poderosas herramientas y aplicaciones del aprendizaje automático que permiten a la información obtenida de señales bioeléctricas apoyar en la rehabilitación y detección de pacientes con afecciones neurológicas.



Figura 3: Interfaz de control de sistema robótico con señales EMG y EEG [6].

En la última década el área de diagnóstico de epilepsia ha implementado técnicas de aprendizaje automático maximizando la información y utilidades que se pueden obtener de las señales EEG. Entre estas aplicaciones se encuentran el análisis de neuroimágenes, detección automatizada de convulsiones a partir de señales EEG, predicción de mejorías utilizando recetas médicas, predicción de resultados post-quirúrgicos, entre otros [7].

En la Universidad del Valle de Guatemala, el trabajo de graduación presentado por María Jesús Angulo Tijerino en el año 2020 [8], implementó aprendizaje no supervisado como primera iteración en el estudio de señales EEG de pacientes HUMANA con registros epilépticos, al realizar una correcta división y categorización de segmentos, se encontró que es posible detectar crisis epilépticas en señales EEG de forma automática. En este mismo año se generó una base de datos con las muestras extraídas de HUMANA en el trabajo de graduación presentado por María Fernanda Pineda Esmieu [9], además creando una herramienta de software para la interacción con la base de datos y resultados.

En el año 2021, David Alejandro Vela Aguilera en su trabajo de graduación [10] desarrolló una nueva versión de la herramienta de análisis de señales EEG para el estudio de epilepsia, en ella se realizaron ajustes en el clustering dividiendo las muestras en 4 clases, esto obtuvo un promedio de exactitud del 96.7 %, presentando una mejora significativa del proceso de clasificación. Adicionalmente, se realizó una sección que genera anotaciones de

Resultados Clasificador SVM (Máquina de Vectores de Soporte)									
Análisis	Caso	Base de Datos	Rendimiento	VP	FN	FP	VN	Kernel Lineal	Kernel RBF
Directo	1	Ubonn	98%	50%	0%	0%	50%	x	x
			100%	50%	0%	0%	50%	x	x
			100%	50%	0%	0%	50%	x	x
			95%	10%	1%	0%	9%	x	x
		CHB-MIT	96.30%	38%	0%	3%	41%	x	-
			100%	41%	0%	0%	41%	x	-
Wavelets	2	Ubonn	100%	25%	0%	0%	25%	x	-
			100%	25%	0%	0%	25%	x	-
			100%	25%	0%	0%	25%	x	-
			100%	25%	0%	0%	25%	x	-

Figura 4: Resumen de resultados con clasificador SVM [8].

Resultados Clasificador RNA (Red Neuronal Artificial)							
Análisis	Caso	Base de Datos	Rendimiento	VP	FN	FP	VN
Directo	1	Ubonn	99.5%	49.5%	0.0%	0.5%	50.0%
			99.5%	50.0%	0.5%	0.0%	49.5%
			97.1%	47.6%	0.5%	2.4%	49.5%
			100.0%	50.0%	0.0%	0.0%	50.0%
Wavelet	2	Ubonn	99.5%	50.0%	0.5%	0.0%	49.5%
			99.0%	50.0%	1.0%	0.0%	49.0%
			99.5%	50.0%	0.5%	0.0%	49.5%
			99.0%	49.5%	0.5%	0.5%	49.5%
		CHB-MIT					
			89.2%	44.9%	5.4%	5.4%	44.4%
			81.6%	42.6%	10.8%	7.6%	39.0%

Figura 5: Resumen de resultados con clasificador ARN [8].

las señales EEG y las muestra de forma interactiva para facilidad del usuario. Hasta este año la manipulación de las señales fue en tiempo continuo, utilizando mayormente técnicas de aprendizaje supervisado, abriendo espacio a la implementación de otras técnicas de aprendizaje y el análisis en diferentes espectros, no solo el de tiempo continuo.

Características	Tiempo Continuo			Wavelet		
Modelo	RNA	SVM		RNA	SVM	
Kernel	-	Gaussiano	Lineal	-	Gaussiano	Lineal
2 Clases	100%	99.80%	100%	97.70%	98.70%	97.90%
3 Clases	97.90%	98.90%	97.20%	98.20%	98.30%	97.20%
4 Clases	88%	91.30%	88.30%	81.20%	83.30%	77.10%
Promedio	95.30%	96.70%	95.20%	92.40%	93.40%	90.70%
Desv. Estándar	5.23%	3.81%	4.99%	7.90%	7.17%	9.64%

Figura 6: Resumen de resultados con clasificador ARN y SVM [10].

La última iteración en el año 2022, con el trabajo de graduación presentado por Camila Lemus Lone [11], implementó ambos tipos de aprendizaje; supervisado y no supervisado. En este trabajo se utilizaron características obtenidas en dominio de frecuencia, el entrenamiento y evaluación fue utilizando Redes Neuronales Artificiales y Máquinas de Vectores de Soporte para clasificadores binarios de las clases Ictal/Sano y las clases Interictal/Preictal,

este tipo de clasificación obtuvo un rendimiento superior al 99 %. Los trabajos anteriormente expuestos implementaron el estudio de señales extraídas de fuentes públicas en su mayoría, presentando limitaciones al acceso de datos de pacientes HUMANA o de muestras utilizado equipo de la Universidad del Valle de Guatemala, afectando el volumen del estudio. Las limitaciones de anteriores proyectos abren el espacio para explorar la extracción de características utilizando diferentes espectros, aplicar los algoritmos desarrollados en bases de datos más extensas y generar estadísticas más exactas.

Cantidad de características	Características utilizadas (razones)	Exactitud RNA	
		Ubonn - Sano/Ictal	Kaggle - Interictal/Preictal
1	1	91.55%	90.50%
2	1, 2	100%	98.90%
3	1, 2, 3	99.99%	98.80%
4	1, 2, 3, 4	99.99%	98.80%
5	1, 2, 3, 4, 5	99.99%	98.80%
6	1, 2, 3, 4, 5, std	100%	99.40%

Figura 7: Resumen del rendimiento de la RNA para dos clases variando la cantidad de características utilizadas en el dominio de la frecuencia [11].

Extracción de características en dominio de la frecuencia				
Característica	Ubonn - Sano/Ictal		Kaggle - Interictal/Preictal	
	Porcentaje de rendimiento RNA	Tiempo	Porcentaje de rendimiento RNA	Tiempo
razón 1	91.00%	3.10 min	98.60%	4.52 min
razón 2	91.10%	3.08 min	84.20%	4.46 min
razón 3	77.00%	4.60 min	98.50%	6.76 min
razón 4	99.70%	2.88 min	96.10%	4.50 min
razón 5	90.90%	4.60 min	98.30%	6.16 min
std	99.60%	11.00 s	67.40%	9.73 s

Figura 8: Resumen del rendimiento de la RNA para dos clasificadores binarios utilizando características individuales en dominio de la frecuencia [11].

Justificación

En la actualidad, el análisis de señales bioeléctricas por medio de algoritmos de aprendizaje automático desempeña un papel fundamental en el diagnóstico y el tratamiento de diversas condiciones médicas, así como en la investigación científica. No obstante, Guatemala no cuenta con acceso a estas herramientas, habilitando un costo de oportunidad para la aplicación de estas nuevas tecnologías de aprendizaje automático.

El análisis de una señal bioeléctrica se enfoca en la morfología y duración de las etapas presentes en la señal. Sin embargo, los métodos tradicionales de análisis (como el método visual) poseen limitaciones en términos de eficiencia y precisión. En una entrevista realizada en 2020 al Dr. Juan Carlos Lara, director de HUMANA, informó que el método visual de

anotaciones sigue siendo el aplicado en Guatemala, implicando un alto costo económico y temporal. En base a lo anteriormente expuesto, se requiere una automatización más rigurosa y sistemática para aprovechar plenamente el potencial de las señales bioeléctricas y mejorar la calidad del diagnóstico médico.

En este contexto, los algoritmos de aprendizaje automático emergen como una herramienta prometedora. Estos algoritmos tienen la capacidad de reconocer patrones complejos, extraer características relevantes y realizar predicciones precisas a partir de grandes cantidades de datos bioeléctricos. Su aplicación en el análisis de señales bioeléctricas es capaz de conducir a una mejor comprensión de los procesos fisiológicos, al desarrollo de métodos de diagnóstico más precisos y a la identificación de posibles tratamientos personalizados.

Previo a este trabajo, las bases de datos accedidas para el análisis de señales bioeléctricas por medio de aprendizaje automático han sido de fuentes públicas, limitando el alcance del proyecto y eliminando robustez en sus conclusiones. Por ello este trabajo tiene como objetivo expandir el alcance de los resultados obtenidos en fases previas, aplicando de manera sistemática algoritmos de aprendizaje automático desarrollados en etapas anteriores a una base de datos más grande. La elección de HUMANA y el sistema Biopac como fuentes de datos bioeléctricos es fundamentada por varias razones. HUMANA, por su parte, cuenta con el único laboratorio de Vídeo Electro Encefalograma en Guatemala, esta base de datos contiene registros obtenidos por especialistas en Neurociencia, justificando su validez. Por otro lado, el sistema Biopac es conocido por su fiabilidad y precisión en la grabación de señales bioeléctricas, lo que garantiza la calidad de los datos utilizados en este estudio.

En base a lo anterior, el presente trabajo se justifica con la validación de los algoritmos desarrollados en etapas anteriores a muchos más datos, mejorando las interfaces, herramientas desarrolladas, generación de estadísticas de comparación y la introducción de un nuevo método de extracción de señales bioeléctricas en el departamento, utilizando sistemas Biopac. Este trabajo puede contribuir a la mejora de los diagnósticos médicos, permitiendo una detección temprana de enfermedades y una mayor precisión en el seguimiento de la evolución de los pacientes. Además, puede impulsar el desarrollo de dispositivos médicos más eficientes.

Objetivos

Objetivo General

Aplicar los algoritmos de aprendizaje automático desarrollados en fases anteriores a una mayor cantidad de señales bioeléctricas procedentes de HUMANA y señales obtenidas con el sistema Biopac dentro de la UVG, y mejorar las interfaces biomédicas desarrolladas anteriormente.

Objetivos Específicos

- Recolectar señales bioeléctricas de forma sistemática y masiva utilizando el sistema Biopac.

- Aplicar algoritmos de extracción de características y aprendizaje automático, desarrollados en fases anteriores, a las señales recolectadas y señales obtenidas de HUMANA.
- Realizar análisis estadísticos para evaluar el rendimiento de los algoritmos e identificar posibles mejoras a los mismos.
- Mejorar las interfaces biomédicas desarrolladas anteriormente e integrar el uso del sistema Biopac para la captura de las señales bioeléctricas.

Marco teórico

Señales bioeléctricas

Las señales bioeléctricas son señales generadas por potenciales y corrientes eléctricas en organismos vivos. Estas señales son producidas por diversos procesos biológicos, como la actividad neuronal, la contracción muscular y la actividad cardíaca. Las señales bioeléctricas son ampliamente utilizadas en la medicina y la investigación científica para el diagnóstico, la monitorización y el estudio de diversas funciones fisiológicas. Las tres señales principales de estudio son: señales Electromiográficas (EMG), Electroencefalográficas (EEG), y Electrocardiogramas (ECG) [12].

Biopac

Biopac Systems, Inc. es una empresa que se dedica a la fabricación y venta de equipos y software para la adquisición y análisis de datos fisiológicos. Sus sistemas de captura se utilizan en investigaciones científicas, estudios clínicos y aplicaciones educativas en una variedad de campos, como la psicología, la fisiología, la neurociencia y la biomecánica. Los sistemas de captura de Biopac están diseñados para medir y registrar una amplia gama de señales fisiológicas [13], incluyendo:

- **Señales electrofisiológicas:** Estos sistemas pueden capturar señales eléctricas generadas por el cuerpo humano, como el electrocardiograma (ECG), la electroencefalografía (EEG), la electromiografía (EMG) y la electrodermal activity (EDA). Estas mediciones proporcionan información sobre la actividad eléctrica del corazón, el cerebro, los músculos y la respuesta de la piel [13].
- **Señales biomecánicas:** Los sistemas de Biopac también pueden registrar señales biomecánicas, como la actividad muscular, el movimiento y la fuerza. Utilizan sensores y transductores para medir parámetros como la actividad de los músculos, la posición y la aceleración [13].
- **Señales respiratorias:** Los sistemas de Biopac pueden capturar señales relacionadas con la respiración, como el flujo de aire, la presión y la capacidad pulmonar. Estos datos son útiles en el estudio de la función pulmonar y la ventilación [13].

Señales EMG

Las señales electromiográficas se generan a partir de la actividad eléctrica de los músculos. Estas señales son captadas mediante electrodos colocados en la superficie de la piel o mediante agujas de electromiografía insertadas directamente en el músculo. El resultado del EMG proporciona información sobre la actividad muscular, como la contracción, relajación, fuerza y coordinación.

El análisis de las señales EMG se utiliza en diversas áreas, como la medicina, la rehabilitación, la biomecánica y la investigación científica. Se emplea para evaluar el funcionamiento y la salud de los músculos, diagnosticar trastornos neuromusculares, y evaluar la eficacia de terapias y tratamientos [14].

Señales ECG

Las señales electrocardiográficas son generadas por la actividad eléctrica del corazón. Estas señales son captadas mediante electrodos colocados en la superficie del cuerpo, generalmente en el pecho, las extremidades y otras ubicaciones específicas donde se pueda monitorear el comportamiento del corazón. El ECG registra los cambios en el potencial eléctrico que ocurren durante el ciclo cardíaco, lo que permite evaluar la función cardíaca y detectar posibles anomalías o trastornos.

El análisis de las señales ECG se utiliza ampliamente en la medicina para el diagnóstico, la monitorización y el tratamiento de enfermedades cardíacas. Permite evaluar la regularidad del ritmo cardíaco, identificar arritmias, falta de flujo sanguíneo en el corazón, evaluar la función de las cámaras cardíacas, evaluar la respuesta del corazón a diferentes estímulos y medicamentos [15].

Señales EEG

Las señales electroencefalográficas son señales bioeléctricas que registran la actividad eléctrica del cerebro. Estas señales son capturadas mediante electrodos colocados en el cuero cabelludo y proporcionan información sobre la actividad neuronal, los patrones de sueño, la función cognitiva y otras actividades cerebrales [16]. Es de particular interés este tipo de señal para el estudio de la epilepsia, pues permite observar el comportamiento del cerebro al exponerse a estímulos sensoriales. Los patrones más frecuentes en señales EEG para describir la actividad cerebral son denominados ritmos. Los ritmos cerebrales son especialmente afectados por el estado del sueño, edad, y presencia de patologías. Los ritmos cerebrales se pueden dividir en ondas Delta (0.5 - 3 Hz), Theta (4 - 7 Hz), Alpha (8 - 12 Hz), Beta (13 - 30 Hz) y Gamma (30 - 120 Hz) [17].

Procesamiento de señales bioeléctricas

El procesamiento de señales bioeléctricas refiere al análisis y la manipulación de las señales eléctricas generadas por el cuerpo humano. Este procedimiento implica una serie de

técnicas y algoritmos que permiten extraer, analizar y visualizar la información contenida en las señales. Esto incluye la eliminación de ruido y artefactos, la detección y segmentación de eventos de interés, la extracción de características relevantes y la clasificación de patrones. El procesamiento de señales utilizando MATLAB es una tarea comúnmente realizada en el campo de la ingeniería y las ciencias de la computación. La herramienta MATLAB ofrece una amplia gama de herramientas y funciones específicamente diseñadas para el procesamiento de señales, lo que facilita el análisis y la manipulación de datos bioeléctricos [18].

El procedimiento se puede generalizar según el siguiente listado de tareas:

- **Adquisición de datos:** MATLAB proporciona funciones para leer y cargar datos de señales bioeléctricas desde diferentes fuentes, como archivos de formato común (por ejemplo, .mat, .csv) o mediante interfaces de adquisición de datos.
- **Preprocesamiento de señales:** Esta etapa implica la limpieza y preparación de las señales para su análisis. Puede incluir la eliminación de ruido, filtrado, normalización, detección de artefactos y segmentación de eventos de interés [19].
- **Análisis y extracción de características:** MATLAB ofrece una variedad de funciones y herramientas para realizar análisis en el dominio de tiempo, frecuencia y tiempo-frecuencia de las señales. También se pueden utilizar técnicas de transformada de Fourier, transformada Wavelet y métodos de análisis espectral para extraer características relevantes de las señales [19].
- **Visualización de señales:** MATLAB proporciona funciones para visualizar las señales en diferentes formatos, como gráficos de líneas, gráficos de dispersión, espectrogramas y mapas de colores. Estas visualizaciones ayudan a comprender mejor las características y patrones de las señales [19].
- **Modelado y clasificación:** MATLAB ofrece algoritmos y herramientas para construir modelos de aprendizaje automático y realizar tareas de clasificación, como la detección de eventos o la identificación de patrones específicos en las señales bioeléctricas [19].

Filtrado de señales

El filtrado de señales bioeléctricas es una técnica fundamental en el procesamiento de señales, este filtrado tiene como objetivo eliminar el ruido no deseado y resaltar las características relevantes de interés. El filtrado se aplica para mejorar la calidad de las señales bioeléctricas, facilitar su análisis y reducir las interferencias que puedan afectar su interpretación. Existen diferentes tipos de filtros utilizados en el procesamiento de señales bioeléctricas, como filtros pasa banda, filtros rechaza banda, filtros paso alto, filtros paso bajo y filtros de mediana, entre otros. Estos filtros se aplican según las características específicas de la señal y los objetivos del análisis.

La elección del filtro adecuado depende de factores como el tipo de ruido presente en la señal, la frecuencia de la señal de interés y la preservación de las características importantes de la señal original. Además, se pueden utilizar técnicas avanzadas, como el filtrado adaptativo, para adaptar el filtro a las características cambiantes de la señal [20].

Métodos de captura de señales bioeléctricas

Entre los métodos de captura de señales bioeléctricas existen:

- **Electrodos de contacto:** Se utilizan electrodos de metal o material conductor que se colocan en contacto directo con la piel. Estos electrodos capturan las señales bioeléctricas a través de la conducción eléctrica [21].
- **Electrodos secos:** Son electrodos que no requieren de un gel conductor o líquido para realizar la conexión eléctrica con la piel [21].
- **Electrodos no invasivos:** Este método de captura no requiere la inserción de electrodos dentro del cuerpo. Este método es utilizado comúnmente en la electroencefalografía (EEG) de superficie, la electrocardiografía (ECG) de superficie y la electromiografía (EMG) de superficie [22].
- **Electrodos intracorporales:** Son métodos de captura que implican la inserción de electrodos dentro del cuerpo para registrar las señales bioeléctricas directamente desde los órganos o tejidos. Estos métodos se utilizan en aplicaciones específicas, como la electrocorticografía (ECoG) o la estimulación eléctrica directa [22].

Epilepsia

La epilepsia es un trastorno neurológico crónico caracterizado por la presencia recurrente de convulsiones o crisis epilépticas. Se origina debido a una actividad eléctrica anormal y excesiva en el cerebro, lo que provoca una variedad de síntomas que van desde movimientos involuntarios y alteraciones de la conciencia, hasta convulsiones generalizadas. Según datos de la OPS, las afecciones neurológicas representan un 6 % de la carga global de enfermedades, posicionando a la afección entre las más comunes a nivel mundial [1].

Esta afección puede manifestarse de diferentes maneras, algunas convulsiones pueden ser breves y sutiles, como la pérdida de conciencia momentánea o una sensación extraña, otras pueden ser más intensas, con movimientos violentos, pérdida total de conciencia y cambios en la respiración y la función motriz [23].

- **Causas:** Existen múltiples causas, como lesiones cerebrales traumáticas, infecciones cerebrales, trastornos genéticos, trastornos metabólicos, tumores cerebrales o anomalías en el desarrollo del sistema nervioso central. Sin embargo, en muchos casos, la causa exacta no se puede determinar [24].
- **Diagnóstico:** El diagnóstico de la epilepsia implica la evaluación clínica, donde se recopilan los antecedentes médicos y se realizan pruebas, como un electroencefalograma (EEG). En este tipo de diagnóstico se utiliza la observación mediante videocámaras, y el monitoreo de las ondas cerebrales en tiempo real [24].
- **Tratamiento:** El tratamiento de la epilepsia generalmente involucra el uso de medicamentos antiepilépticos para controlar las convulsiones. En algunos casos, cuando los medicamentos no son efectivos, se pueden considerar otras opciones terapéuticas, como la cirugía cerebral o la estimulación del nervio vago [25].

Aprendizaje automático

El aprendizaje automático es un campo de estudio de la inteligencia artificial que se enfoca en el desarrollo de algoritmos y modelos estadísticos que permiten a las computadoras aprender y mejorar automáticamente a través de la experiencia. El aprendizaje automático se utiliza en una amplia gama de aplicaciones, particularmente se utiliza en el diagnóstico médico como herramienta para mejorar la eficiencia y categorización de las señales bioeléctricas. Este proceso de automatización permite la extracción de información que por métodos de visualización es imposible obtener [26]. Algunas de las características del aprendizaje automático son:

- **Extracción automática de características:** El aprendizaje automático permite identificar y extraer automáticamente características relevantes de las señales bioeléctricas, lo que facilita el análisis y la interpretación de los datos. Esto es especialmente útil en señales complejas como la electroencefalografía (EEG) o la electromiografía (EMG), donde las características relevantes pueden ser sutiles y difíciles de detectar para un observador humano [27].
- **Clasificación y diagnóstico:** Los algoritmos de aprendizaje automático pueden aprender a clasificar diferentes patrones en las señales bioeléctricas, lo que resulta crucial para el diagnóstico de enfermedades y trastornos. Por ejemplo, en el caso de la epilepsia, el aprendizaje automático puede ayudar a identificar patrones anormales en las señales EEG que son característicos de los episodios epilépticos [27].
- **Personalización:** El aprendizaje automático también puede ayudar a personalizar los enfoques de tratamiento y cuidado de los pacientes. Al analizar las señales bioeléctricas de un individuo y combinarlas con información clínica adicional, los modelos de aprendizaje automático pueden adaptar las intervenciones terapéuticas para satisfacer las necesidades específicas de cada paciente, mejorando así los resultados y la atención médica en general [27].

Aprendizaje no supervisado

El aprendizaje no supervisado es una de las ramas del aprendizaje automático, en ella el algoritmo de aprendizaje se encarga de descubrir patrones y estructuras ocultas en los datos sin la ayuda de etiquetas o información previa, por ende su nombre. El aprendizaje no supervisado se centra en explorar la estructura de los datos y agruparlos en conjuntos. Este tipo de aprendizaje permite agrupar a los datos en cantidades determinadas por el usuario o de forma natural [28].

Aprendizaje supervisado

El aprendizaje supervisado es otra rama del aprendizaje automático donde se utilizan datos etiquetados para entrenar un modelo. En ella, el modelo aprende a realizar predicciones o tomar decisiones mediante el método de entrenadores y sujetos de prueba, donde se

conoce la respuesta correcta. El objetivo es migrar el conocimiento aprendido de las etiquetas previas, a datos no pertenecientes al grupo entrenado, logrando un rendimiento preciso en la tarea de predicción de los datos no conocidos [29].

Características en el dominio del tiempo

Las características más simples de la señal EEG son características estadísticas, como la media, la mediana, la varianza, la desviación estándar, la asimetría, la curtosis y similares [30].

- **Integración de la señal:** Este valor es obtenido calculando la sumatoria de los valores absolutos de una señal EMG. Puede ser tratada como un estimador de poder [31].
- **Media del valor absoluto:** Estima la media del valor absoluto, sumando el valor absoluto de x_i en un segmento k y dividido por la longitud del segmento [31].
- **Media cuadrática:** Se modela como un proceso aleatorio gaussiano modulado en amplitud cuyo RMS es relacionado con la fuerza constante y la contracción no fatigante [31].
- **Longitud de *Waveform*:** Longitud acumulada en forma de onda sobre un segmento escogido. Los valores resultantes indican una medida de amplitud, frecuencia y duración en un solo valor [31].
- ***Zero Crossings*:** Número de veces que la forma de onda cruza cero, es bueno incluir un umbral para detectar los ceros, ya que por el ruido pueden haber ceros falsos [31].

Características en el dominio de la frecuencia:

Las potencias relativas de ciertas bandas de frecuencia son las características de dominio de frecuencia más utilizadas en todos los campos de análisis de las señales de EEG. Comúnmente, bandas de frecuencia más utilizadas son la Delta, Theta, Alpha, Beta, y Gamma, aunque también existe la banda Sigma (12–14 Hz). Las relaciones entre las bandas de frecuencia (*ratios*) se utilizan ampliamente como características en el análisis de señales de EEG [30].

- **Coefficientes autorregresivos:** Describe cada muestra de la señal EMG como una combinación lineal de muestras anteriores más un término de error de ruido blanco. Los coeficientes AR se utilizan como características en el reconocimiento de patrones EMG [31].
- **Mediana de frecuencia:** Divide la densidad del espectro de potencia en dos partes iguales [31].
- **Densidad espectral de potencia:** Es la base para el cálculo de las características del dominio de la frecuencia, se puede calcular con varios métodos paramétricos y no paramétricos. Los métodos no paramétricos se usan con más frecuencia e incluyen

métodos como la transformada de Fourier (generalmente calculada con el algoritmo de transformada rápida de Fourier, FFT), el método de Welch o el método multicono de Thompson [30].

Metodología

Como primer paso es necesaria la instalación del software para utilizar el sistema de captura Biopac, se debe corroborar el funcionamiento realizando calibración y pruebas de captura para los diferentes tipos de señales bioeléctricas. El siguiente paso involucra un proceso sistemático para la obtención de las señales bioeléctricas por medio del sistema de captura Biopac, definiendo gestos de análisis (etiquetas), y una estructura común para su almacenamiento en la base de datos. Los datos proporcionados por HUMANA y los obtenidos con los sistemas Biopac serían sometidos a un algoritmo de conversión a variables tipo struct, haciendo posible la interpretación dentro de MATLAB y por ende, la interfaz biomédica.

Posteriormente, es necesaria una investigación a fondo de los algoritmos de aprendizaje automático utilizados en proyectos anteriores, enfatizando el desarrollo en los algoritmos de aprendizaje supervisado. La etapa siguiente se enfocaría en la reconstrucción de las aplicaciones previamente desarrolladas en la UVG, validando su funcionamiento con los métodos y datos extraídos en esta nueva iteración. Comprender las estructuras de estas aplicaciones es un paso esencial para cumplir el objetivo general en este trabajo: aplicar los algoritmos de aprendizaje automático desarrollados en fases anteriores a una mayor cantidad de señales bioeléctricas procedentes de HUMANA, incluyendo señales obtenidas con el sistema Biopac, y mejorar las interfaces biomédicas desarrolladas anteriormente.

La etapa de desarrollo del proyecto culminaría con la implementación del sistema Biopac para la captura de señales directamente desde las interfaces biomédicas. Dando inicio a la etapa de extracción de resultados, dónde se aplicarían de manera sistemática los algoritmos desarrollados en fases anteriores, y utilizando estadística, se estaría comparando el funcionamiento, eficacia y exactitud de los algoritmos, y la calidad de los nuevos métodos de extracción de datos. Con el objetivo de ayudar a los especialistas a encontrar métodos automatizados y precisos de crear anotaciones, se buscaría corroborar el funcionamiento consultando con especialistas de HUMANA, obteniendo retroalimentación respectiva.

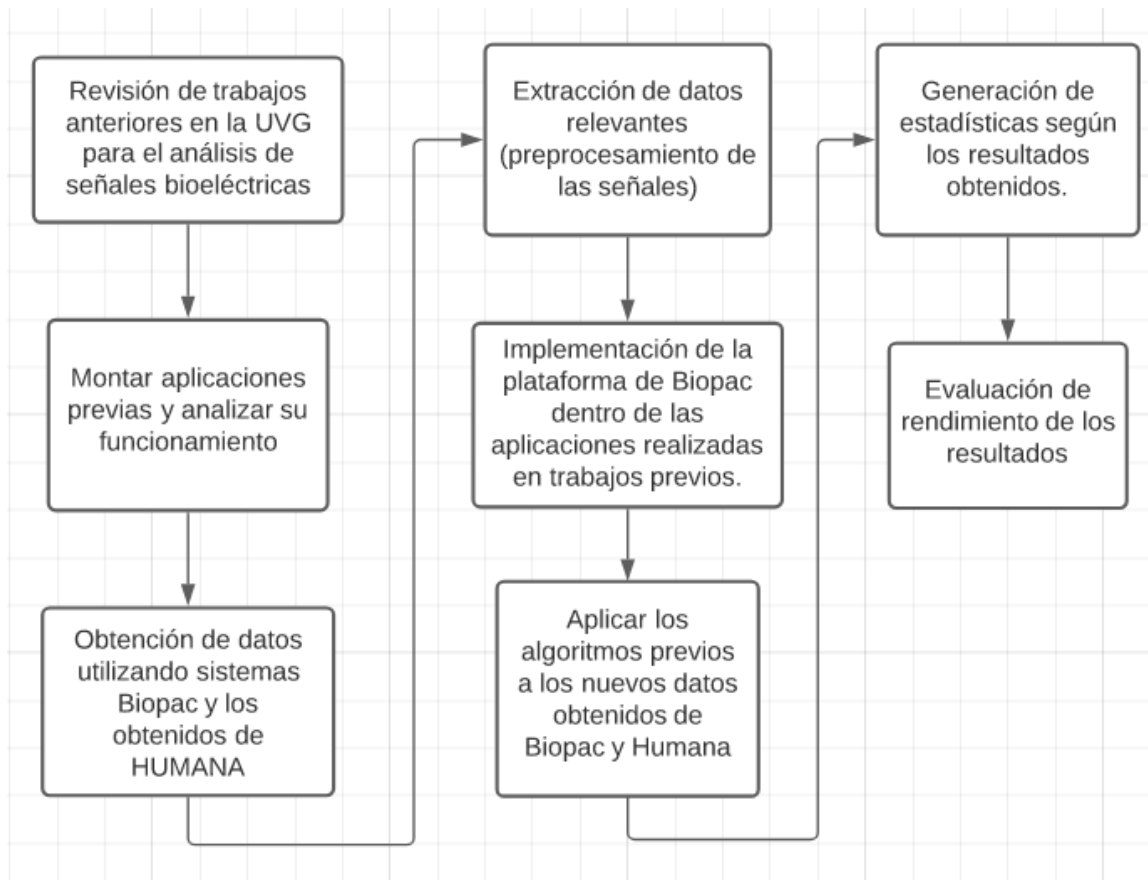


Figura 9: Cronograma de actividades.

Cronograma de actividades

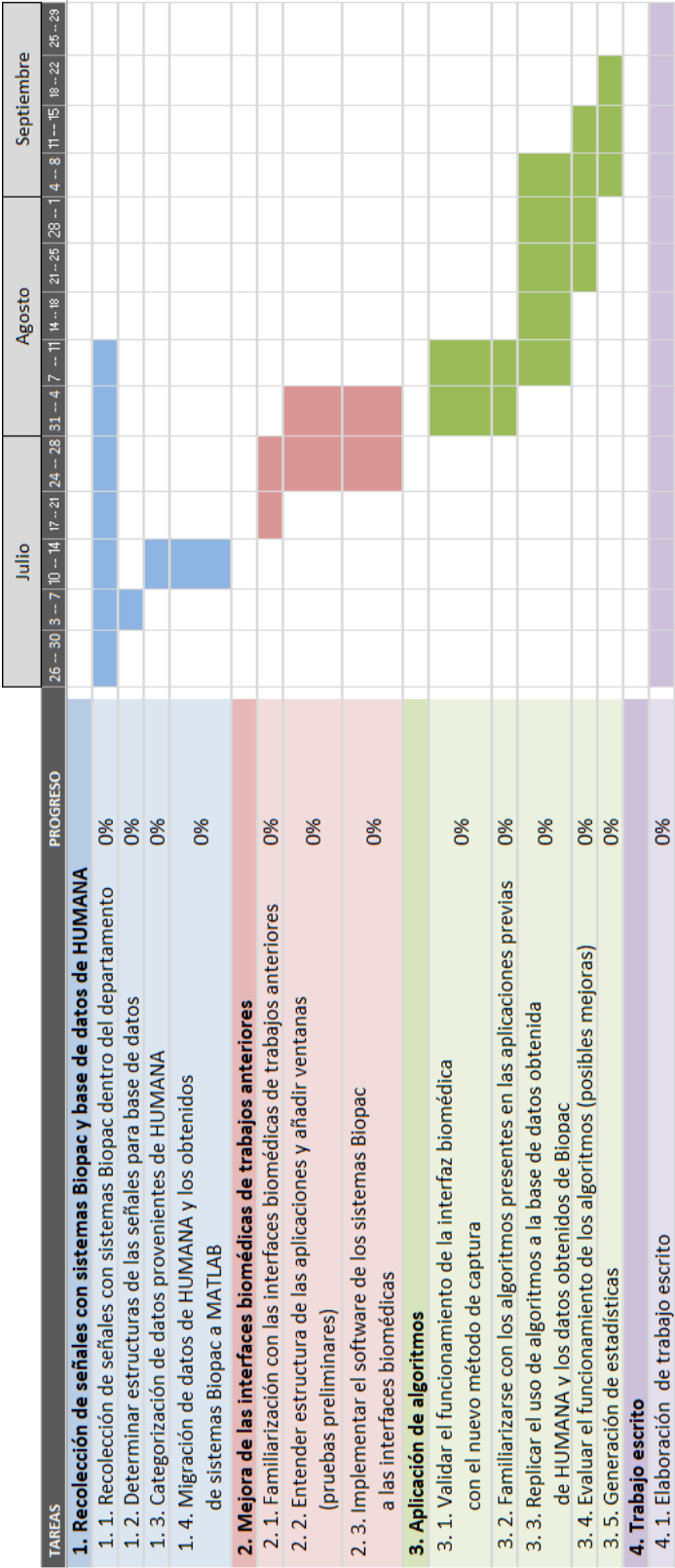


Figura 10: Cronograma de actividades.

Índice preliminar

1. Prefacio
2. Lista de figuras
3. Lista de cuadros
4. Resumen
5. Abstract
6. Introducción
7. Antecedentes
8. Justificación
9. Objetivos
 - a) Objetivo general
 - b) Objetivos específicos
10. Alcance
11. Marco teórico
 - a) Señales Bioeléctricas
 - b) Señales EMG
 - c) Señales ECG
 - d) Epilepsia
 - 1) Causas
 - 2) Diagnósticos
 - 3) Tratamiento
 - e) Señales EEG
 - f) Aprendizaje Automático
 - 1) Extracción automática de características
 - 2) Clasificación y diagnóstico
 - 3) Personalización
 - g) Aprendizaje no supervisado
 - h) Aprendizaje supervisado
12. Procesamiento de Datos
 - a) Procesamiento de datos obtenidos de sistemas Biopac
 - b) Procesamiento de datos obtenidos de HUMANA
 - c) Manipulación del conjunto de datos de forma automática
 - 1) Creación de .mats y structs

13. Experimentos y Resultados
 - a) Implementación de Biopac como herramienta en las interfaces biomédicas
 - 1) Pruebas en tiempo continuo
 - b) Aplicación de algoritmos
 - 1) Aprendizaje supervisado
 - a' Resultados con datos provenientes de HUMANA
 - b' Resultados con datos provenientes de Biopac
 - 2) Aprendizaje no supervisado
 - a' Resultados con datos provenientes de HUMANA
 - b' Resultados con datos provenientes de Biopac
 - c) Generación de estadísticas
 - d) Evaluación de estadísticas con HUMANA
14. Conclusiones
15. Recomendaciones
16. Bibliografía
17. Anexos
18. Glosario

Referencias

- [1] Organización Panamericana de la Salud, *Informe sobre la epilepsia en América y El Caribe*, <https://www3.paho.org/>, Accesado: 23-04-2023, 2021.
- [2] HUMANA, *Centro de Epilepsia y Neurología Funcional*, <https://humanagt.org/epilepsia/>, Accesado: 23-04-2023, 2023.
- [3] R. Sepulveda, O. Montiel y G. Díaz, "Clasificación de Señales Encefalográficas Mediante Redes Neuronales Artificiales," *Redalyc*, vol. 19, n.º 1, págs. 69-88, 2015.
- [4] M. F. Girón, "Interfaz Biomédica para el Control de Sistemas Robóticos Utilizando Señales EMG," Tesis de licenciatura, Universidad Del Valle de Guatemala, 2021.
- [5] R. Ralda, "Interfaz Biomédica para el Control de Sistemas Robóticos Utilizando Señales EEG," Tesis de licenciatura, Universidad Del Valle de Guatemala, 2020.
- [6] R. J. Cáceres, "Interfaz para el Control de Sistemas Robóticos Utilizando Combinaciones de Señales Biomédicas," Tesis de licenciatura, Universidad Del Valle de Guatemala, 2022.
- [7] B. Abbasi y D. Goldenholz, "Machine learning applications in epilepsy," *Epilepsia*, vol. 60, n.º 10, págs. 1528-1167, 2019.
- [8] M. J. Angulo, "Análisis y Reconocimiento de Patrones de Señales Biomédicas de Pacientes con Epilepsia," Tesis de licenciatura, Universidad Del Valle de Guatemala, 2020.

- [9] M. F. Pineda, "Diseño e Implementación de una Base de Datos de Señales Biomédicas de Pacientes con Epilepsia," Tesis de licenciatura, Universidad Del Valle de Guatemala, 2020.
- [10] D. A. Vela, "Automatización del Proceso de Anotación de Señales EEG de Pacientes con Epilepsia por Medio de Técnicas de Aprendizaje Automático," Tesis de licenciatura, Universidad Del Valle de Guatemala, 2021.
- [11] C. Lemus, "Análisis y Anotación de Señales Bioeléctricas de Pacientes con Epilepsia Utilizando Técnicas de Aprendizaje Automático Supervisado y No Supervisado," Tesis de licenciatura, Universidad Del Valle de Guatemala, 2022.
- [12] G. Pfurtscheller, G. Bauernfeind y T. Solis-Escalante, "Brain-Computer Interfaces and Signal Processing," *IEEE Signal Processing Magazine*, vol. 36, n.º 2, págs. 16-23, 2019. DOI: 10.1109/MSP.2018.2876352.
- [13] Biopac Systems Inc., *Biopac Systems, Inc.* <https://www.biopac.com>, 2021.
- [14] R. Merletti y P. Parker, *Electromyography: Physiology, Engineering, and Non-Invasive Applications*. Hoboken, NJ: Wiley-IEEE Press, 2004.
- [15] A. Goldberger y Z. Goldberger, *Clinical Electrocardiography: A Simplified Approach*. Philadelphia, PA: Elsevier, 2017.
- [16] E. Niedermeyer y F. Lopes da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. New York, NY: Oxford University Press, 2018.
- [17] J. Smith y S. Johnson, "Analysis of EEG Signals for Brain Research," *Journal of Neuroscience*, vol. 25, n.º 3, págs. 45-60, 2018.
- [18] A. Oppenheim, R. Schafer y J. Buck, *Discrete-Time Signal Processing*. Upper Saddle River, NJ: Pearson, 2010.
- [19] S. Monk, *MATLAB: A Practical Introduction to Programming and Problem Solving*. Amsterdam, Netherlands: Elsevier, 2016.
- [20] K. Gopalakrishnan, K. Sivakumar y A. Khosla, "Review on Signal Filtering Techniques for Biomedical Signal Processing," *Biocybernetics and Biomedical Engineering*, vol. 39, n.º 1, págs. 23-39, 2019. DOI: 10.1016/j.bbe.2018.11.005.
- [21] J. Webster, *Medical Instrumentation: Application and Design*. Hoboken, NJ: Wiley, 2009.
- [22] R. Plonsey y R. Barr, *Bioelectricity: A Quantitative Approach*. New York, NY: Springer, 2000.
- [23] J. Engel Jr., *Epilepsy: A Comprehensive Textbook*, 2nd Edition. Philadelphia, PA: Lippincott Williams Wilkins, 2013.
- [24] R. S. Fisher, C. Acevedo y A. Arzimanoglou, "ILAE Official Report: A practical clinical definition of epilepsy," *Epilepsia*, vol. 55, n.º 4, págs. 475-482, 2014.
- [25] D. R. Fish, S. C. Schachter, W. E. Dodson y E. L. So, *Epilepsy: A Guide to Balancing Your Life*, 2nd Edition. New York, NY: Demos Medical Publishing, 2008.
- [26] Z. Obermeyer y E. Emanuel, "Predicting the Future - Big Data, Machine Learning, and Clinical Medicine," *The New England Journal of Medicine*, vol. 375, n.º 13, págs. 1216-1219, 2016. DOI: 10.1056/NEJMp1606181.

- [27] V. François-Bongarçon, J. Lecailtel y G. Coatrieux, “Machine learning in bioelectrical signal analysis: A survey,” *Biomedical Signal Processing and Control*, vol. 42, págs. 73-88, 2018.
- [28] C. M. Bishop, *Pattern Recognition and Machine Learning*, 1st Edition. New York: Springer, 2006.
- [29] S. J. Russell y P. Norvig, “Artificial Intelligence: A Modern Approach,” *Pearson*, 2010.
- [30] I. Stancin, M. Cifrek y A. Jovic, “A Review of EEG Signal Features and their Application in Driver Drowsiness Detection Systems,” *Sensors*, vol. 21, n.º 11, 2021. DOI: 10.3390/s21113971.
- [31] U. Acharya, S. Oh, Y. Hagiwara, J. Tan y H. Adeli, “Deep Learning Applications in Healthcare Diagnosis,” *Computational and Structural Biotechnology Journal*, vol. 16, págs. 1-13, 2018. DOI: 10.1016/j.csbj.2018.01.003.