

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería



**Automatización del Proceso de Anotación de Señales EEG de
Pacientes con Epilepsia por Medio de Técnicas de
Aprendizaje Automático**

Protocolo de trabajo de graduación presentado por David Alejandro
Vela Aguilera, estudiante de Ingeniería Mecatrónica

Guatemala,

2021

Resumen

El trabajo de graduación propuesto busca aplicar algoritmos de aprendizaje automático a señales electroencefalográficas. El objetivo es poder generar anotaciones pertinentes a señales con epilepsia, tales como duración, morfología y abundancia para segmentos preictales e interictales. Para lograr el objetivo, se propone desarrollar módulos de procesamiento de señales, selección de características de interés de las señales EEG, implementación de los algoritmos de aprendizaje automático supervisado, muestra visual de los resultados y anotaciones y la implementación en conjunto con la *Epileptic EEG Analysis Toolbox* desarrollada previamente por María Angulo y María Fernanda Pineda.

Antecedentes

Se tiene conocimiento que, para finales del año 2020, habían 65 millones de personas al rededor del mundo con epilepsia, y, solamente en Estados Unidos, se estima que al año se diagnostican 150,000 casos nuevos [1]. A pesar de que esta es una de las condiciones médicas con mayor antigüedad dentro del conocimiento humano, es persistente el miedo hacia las personas que la padecen. Si un análisis clínico muestra que un paciente puede estar padeciendo de este trastorno, es competente someterlo a un electroencefalógrafo y revisar su actividad cerebral; el resultado de esto es un conjunto de señales conocidas como electroencefalografía.

En Guatemala, existe el centro de epilepsia y Parkinson conocido como HUMANA. Esta es una organización formada por profesionales en Neurociencias, los cuáles trabajan en beneficio de los pacientes que padecen problemas Neurológicos de difícil control; como lo puede ser la epilepsia [2]. Este centro tiene las herramientas para obtener señales electroencefalográficas de pacientes sanos y pacientes con episodios epilépticos, por lo que también hacer las anotaciones competentes dentro de estos registros. Las anotaciones de un EEG se realizan de forma manual, siendo un especialista entrenado el encargado de observar la señal y resaltar segmentos de interés dentro del mismo. El tiempo de esta operación depende del tiempo del registro, por lo que, para registros de duraciones largas (desde 24 horas hasta un par de días), esta tarea puede consumir mucho tiempo y atrasar el diagnóstico [3].

Uno de estos trabajos, y la base de este trabajo de graduación es el realizado por María Jesús Angulo Tijerino, de nombre *Análisis y Reconocimiento de Patrones de Señales Biomédicas de Pacientes con Epilepsia* [4]. Entre los resultados de mayor importancia para el presente trabajo se encuentra una primera iteración en el uso de aprendizaje automático no supervisado para detectar, dentro de los registros del instituto HUMANA, aquellos registros que presenten epilepsia. Adicional a esto, se encuentra el desarrollo de una app de Matlab, con la capacidad de conectarse a una base de datos para analizar señales EEG y realizar anotaciones que pueden ser consultadas en cualquier momento.

En los resultados del aprendizaje automático no supervisado, se tiene que fue empleada la técnica de agrupamiento (conocida como *clustering*). Para este *clustering*, se definieron 2 grupos como parámetro para separar las señales; el resultado no fue satisfactorio. A pesar de que el método tenía clasificadores con alta exactitud, varios registros sin presencia de epilepsia eran marcados como registros con presencia del desorden; esto indica que se debe

de seleccionar otro vector de características para poder establecer los grupos deseados.

La aplicación de Matlab fue realizada en por María Angulo, en conjunto con Maria Fernanda Pineda Esmieu, y se presenta en su trabajo *Diseño e Implementación de una Base de Datos de Señales Biomédicas de Pacientes con Epilepsia* [5]. En esta app, es posible la conexión de Matlab con una base de datos; esto para poder leer y guardar datos de forma ordenada. La siguiente parte de la interfaz es la capacidad de esta de usar aprendizaje de máquina para generar un vector de características, visualización de un único canal de la señal electroencefalográfica, así como dos ventanas para la implementación de una máquina de vectores de soporte o de una red neuronal.

Justificación

En el contexto nacional, 325,000 personas padecen de algún tipo de epilepsia. Además de tratarse de un problema médico, es un problema social, ya que esta condición genera un alto impacto emocional tanto para el paciente como para su familia [6]. Los métodos para la detección de epilepsia datan desde 1970, y generalmente se basan en clasificadores binarios (estado epiléptico o estado normal)[7].

Al hablar de un evento de interés dentro de un electroencefalograma, los avisos de tipo binario se quedan cortos. Hay muchas características que se pueden anotar; por ejemplo, su morfología (la onda es irregular, sinusoidal, regular, rítmica, etc.), así como su abundancia (el segmento de interés es continuo, abundante, ocasional, esporádico, etc.) [3]. En el contexto de la epilepsia, es de interés conocer la morfología y duración de las etapas preictales (antes del evento) e interictales (entre eventos epilépticos). Lo anterior se debe a que, si se tiene un registro clínico que indique que hay epilepsia presente, un evento interictal en el electroencefalograma ayuda a asegurar el diagnóstico de epilepsia en el paciente; mientras que, al poder reconocer y anotar los eventos preictales, es posible conocer cuánto tiempo se tiene antes de que se ocurra el episodio ictal.

Este trabajo pretende automatizar el proceso manual de los especialistas, marcando ventanas de interés con anotaciones, las cuáles, ellos pueden validar o descartar a su criterio. Esto les permitiría centrarse más en la interpretación de la señal y no en usar el tiempo buscando dentro del registro completo. Además, es un primer paso a una automatización para el monitoreo en tiempo real de un paciente con un electroencefalograma.

Objetivos

Objetivo General

Desarrollar y validar un proceso de reconocimiento y anotación de posibles episodios ictales en señales electroencefalográficas (EEG) de pacientes con epilepsia, por medio de técnicas de aprendizaje automático.

Objetivos Específicos

- Adaptar la herramienta de software desarrollada en la fase previa para el reconocimiento de señales EEG con registros ictales de pacientes de HUMANA.
- Identificar automáticamente segmentos de interés dentro de las señales, de acuerdo a parámetros utilizados por el personal de HUMANA.
- Generar automáticamente archivos con anotaciones relevantes para los segmentos de interés identificados.
- Validar los segmentos de interés y las anotaciones con especialistas de HUMANA.

Marco teórico

Epilepsia

La epilepsia es un desorden crónico, el cuál tiene como síntoma conocido las convulsiones, las cuáles son recurrentes y espontáneas. Estas convulsiones pueden tener origen gracias a un daño cerebral, o una tendencia familiar a padecer el desorden; sin embargo, la causa es generalmente desconocida [8].

Algunas veces, el historial clínico y familiar, así como los exámenes realizados por medio de electroencefalogramas, son similares entre un grupo de personas con el desorden de epilepsia. Esto se le conoce como “Epilepsy Syndrome”, y son definidos por ciertas características como:

- Tipo o tipos de epilepsia
- Edad en la que comenzaron las convulsiones
- Causas de las convulsiones
- Ya sea sí las convulsiones son heredadas
- La parte del cerebro involucrada
- Los factores que provocan las convulsiones
- La severidad y frecuencia de las convulsiones
- Patrones específicos dentro de un electroencefalograma, durante y entre convulsiones
- Hallazgos en imágenes cerebrales, hechas con resonancia magnética (MRI) o tomografía por computadora (CT).

Cabe destacar que, a pesar de todas las características descritas, no todos los síndromes de epilepsia pueden ser catalogados.

Tipos de epilepsia

Dentro de la descripción de las características que definen a un síndrome de epilepsia, se denotó que existen distintos tipos de epilepsia, además de que una persona puede sufrir más de uno. La Liga Internacional Contra la Epilepsia estableció una clasificación para convulsiones y epilepsias en 2017, y en enero de 2021, actualizaron la clasificación anterior para modificar detalles en las convulsiones de neonatos. Esta clasificación es la siguiente [9]:

- Automatismos

Consta de una actividad motora más o menos coordinada, la cuál suele ocurrir cuando la cognición se encuentra alterada. Se asemeja a menudo a un movimiento involuntario.

- Clónico

Sacudidas simétricas o asimétrica, las cuales son generalmente repetitivas e involuntarias a los mismos grupos musculares.

- Espasmos epilépticos

Se caracteriza por repentinas flexiones, extensiones o una mezcla de ellas, ocurriendo en los músculos predominantemente proximales y troncales. Son de forma más sostenida que un movimiento mioclónico, pero no tan sostenidos como una convulsión tónica.

- Mioclónico

Contracciones repentinas e involuntarias en grupos de músculos de distintas topografías (axiales, miembros proximales, distales). Son de corta duración (menores a 100 milisegundos).

- Tónico

Consta de un aumento sostenido en la contracción muscular; puede durar desde unos segundos hasta minutos.

- Autonómico

Una alteración distinta de la función del sistema nervioso autónomo, que involucra las funciones cardiovasculares, pupilares, gastrointestinales, sudomotores, vasomotores y funciones termorreguladoras.

- Arresto conductual

Pausa (conocida como arresto) de actividades; consta de congelamiento o inmovilización como en el comportamiento de arresto convulsivo.

- Convulsión secuencial

Se clasifican aquí aquellas convulsiones que, en un electroencefalograma, presentan eventos con una secuencia de signos, síntomas y cambios en distintos momentos.

- Convulsión únicamente electroencefalográfica

Son las subclínicas, sin ninguna manifestación clínica.

Señales electroencefalográficas

Las señales electroencefalográficas (referidas el resto del documento como EEG) provienen de la lectura de la actividad eléctrica cerebral, producida por potenciales iónicos en las neuronas. Se generan al realizar la captación de potenciales en el cerebro por medio de electrodos. En dependencia del tipo de EEG, se colocan en tres posibles puntos [10]:

- EEG profundo: emplea microelectrodos implantados en el cerebro. Para este tipo de mediciones, se requiere una intervención quirúrgica para colocar los electrodos.

- Electrocortigrama (ECoG): emplea electrodos corticales; estos se ubican directamente en la corteza cerebral. Al igual que el EEG profundo, se requiere de intervención quirúrgica.
- EEG estándar: es un EEG completamente externo; involucra electrodos que se posicionan sobre el cuero cabelludo. Es de los más usados, debido a que no requiere intervención quirúrgica, lo que reduce tanto el riesgo como el tiempo de preparación del paciente.

Entre más cerca del cerebro se tome la medición, se obtienen señales con mejores relaciones al ruido (señales más claras). Una desventaja de una EEG son las magnitudes de potenciales que se manejan, las cuales se encuentran entre los 10 a 200 μV ; además, sus frecuencias varían desde un poco menos de 1 Hz a 100Hz (algunas documentaciones presentan que los voltajes van desde 0.2 a 300 μV y de DC a 150Hz). Para el posicionamiento de electrodos, se emplea el sistema 10/20, el cuál es el más utilizado internacionalmente [11].

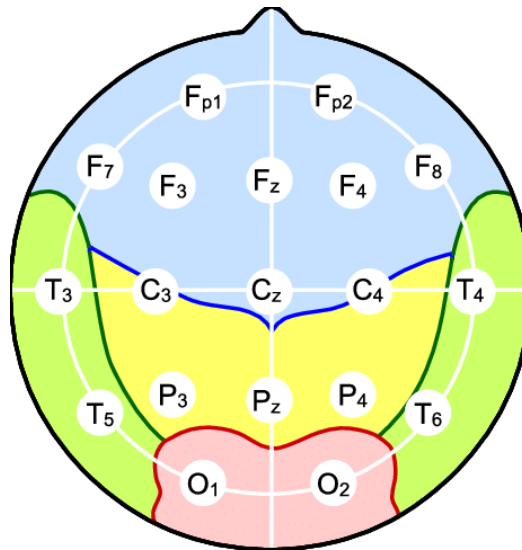


Figura 1: Posicionamiento de los electrodos en estándar 10/20 [11].

Las señales por medir se separan en ritmos cerebrales. Son caracterizadas por su frecuencia, localización y la asociación con su función con el cerebro. Algunas características importantes de las señales son:

- Sufren de un efecto conocido como “desincronización relacionada al evento” o ERD por sus siglas en inglés. Refiere a que, para prepararse para un evento que afecte la zona cortical donde se genera la señal, estas tienden a atenuarse o bloquearse.
- Luego de un ERD, reaparece el ritmo de forma intensificada. Esto se conoce como “sincronización relacionada al evento”, o ERS.
- Los ritmos cerebrales aparecen sobre un grupo de neuronas que no estén realizando su tarea asignada; las señales aparecen en condición de reposo.

Los ritmos cerebrales se separan por su frecuencia en señales alfa, beta, delta, theta, mu y gamma. Actualmente existen aún más ritmos, pero comparten bandas de frecuencia con los mencionados anteriormente. Las bandas de frecuencia son [12]:

- I. Ritmo delta (δ), de 0.5 a 2.5Hz.
- II. Ritmo theta (θ), de 4 a 7Hz.
- III. Ritmo alfa (α), de 8 a 13Hz.
- IV. Ritmo mu (μ), de 8 a 13Hz al igual que el ritmo α .
- V. Ritmo beta (β), de 13 a 30Hz.
- VI. Ritmo gamma (γ), frecuencias mayores a 30Hz.

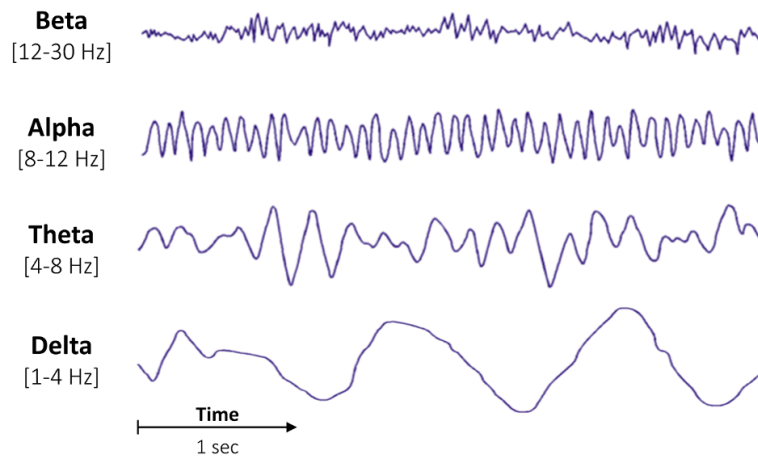


Figura 2: Ejemplo de ritmos cerebrales presentes en un EEG [12].

Procesamiento de señales EEG

Gracias a que muchos EEG pueden contar con relaciones de ruido grandes, para una buena visualización de la señal es necesario filtrar la señal proveniente de un EEG. El mejor filtro para emplearse es un filtro adaptativo. Sin embargo, estos requieren el perfil del ruido. Un ejemplo de su uso en EEG es la remoción de ruido generada por las señales cardíacas (señales que se obtienen de un electrocardiograma o ECG).[13].

Por otro lado, se tienen los filtros *Infinite Impulse Response* (IIR), los cuáles brindan una mejor respuesta en cuanto a la reducción de ruido en un ancho de banda determinado. Entre estos filtros IIR, se tienen los *Butterworth* y *Chebyshev*, los cuáles son de interés para el trabajo a realizar.

Un tercer tipo de filtro para señales biomédicas es el filtro de Wiener, el cuál es un filtro lineal que se aplica de forma adaptativa a la varianza local de los datos. Si la variación es grande, se realiza poco suavizado y cuando la variación es pequeña se realiza un mayor suavizado. Es un filtro más selectivo, ya que permite el paso de los bordes y de otras partes de alta frecuencia [13].

Obtención de características de un EEG

Gracias a la naturaleza orgánica de las señales EEG, son complejas de cuantificar. Esto hace que no sea adecuado caracterizarlas por métodos determinísticos; en cambio, es mejor el uso de métodos estocásticos. Es general que para el procesamiento de un EEG, se empleen tres clases de análisis [14]:

- Dominio del tiempo
- Dominio de frecuencia
- Dominio TS y TF (tiempo-escala y tiempo-frecuencia)

En el dominio del tiempo, se encuentra el análisis de medidas lineales univariadas (como lo son los momentos estadísticos, dígame varianza, asimetría, curtosis). En el dominio de frecuencia se encuentra la famosa transformada de Fourier, la cuál extrae el contenido espectral de una señal [15]. La tercera técnica emplea lo que se conoce como la transformada de *Wavelet*, la cuál es una transformación que proporciona la representación tiempo-frecuencia de una señal. Esta última es especialmente útil, ya la transformada de Fourier se emplea en señales con características estacionarias, lo cuál no es el caso en las señales EEG; por otro lado, la transformada de Wavelet toma en cuenta esto, y utiliza simultáneamente características de la convolución entre el tiempo y la frecuencia [14].

Anotaciones para un EEG

Las anotaciones para los electroencefalogramas siempre se han realizado de forma manual: un especialista observa todo el registro y va marcando los segmentos donde se observen comportamientos de interés. Las combinaciones de anotaciones que se pueden realizar son enormes, además que, el desconocimiento de características bien definidas puede generar una sobre interpretación. Por lo anterior, es necesario recordar: ¿qué responde realmente un registro de EEG? En interés de este trabajo de graduación, una de las repuestas que da es la probabilidad que tiene un paciente de sufrir epilepsia [3].

Tomando esto en cuenta, se tienen dos características importantes al momento de describir un evento dentro de un EEG: la cantidad de veces que se repite en un segundo, la cantidad de veces que esta aparece en todo el registro EEG y la forma que tiene; esto se conoce como frecuencia, abundancia y morfología respectivamente. Para la frecuencia, se anota con números en Hz, y pueden darse ejemplos en base a otras ondas (por ejemplo, un ritmo alfa tiene frecuencia de 8-12Hz, por lo que se puede dar de referencia si la frecuencia del segmento de interés se encuentra en el mismo rango). La abundancia es una anotación empírica, la cuál usa términos como continuo, abundante, frecuente, ocasional, esporádico y único. Por ultimo, la morfología responde a la forma de la señal, y puede ser sinusoidal, regular, irregular, monótona, polimorfa, rítmica o semi rítmica [3].

Machine Learning

Se conoce como *Machine Learning* o Aprendizaje Automático al conjunto de técnicas que forman parte de la inteligencia artificial. Estas técnicas (las cuáles son algoritmos de computadora) buscan el aprendizaje dentro de grandes grupos de datos. La característica de importancia de los algoritmos es que son capaces de predecir casos nuevos, todo en base a la experiencia adquirida con un conjunto de datos enfocados para esta tarea. A esa característica se le conoce como la “generalización”. Dentro de estas técnicas, se separan dos grandes grupos: las técnicas supervisadas y las no supervisadas [16].

Aprendizaje Automático supervisado

Este tipo de técnicas pretenden construir un patrón a partir de datos que ya son conocidos previamente, lo cuál da paso a una predicción del valor de salida de nuevos datos, con el conocimiento único de ciertas características. Se usa una gran cantidad de datos, los cuáles, el propio investigador separa en tres grupos [17]:

- I. Grupo de entrenamiento: En este grupo se encuentran todos los datos con los que se harán las iteraciones para que el algoritmo “aprenda”.
- II. Grupo de validación: En este grupo se colocan datos para que el algoritmo valide su aprendizaje. Forma parte de un proceso cíclico, ya que se va del entrenamiento a la validación, y, de necesitarse, de nuevo al entrenamiento. Busca que el algoritmo funcione para casos reales, y no únicamente para los datos de entrenamiento.
- III. Grupo de prueba: Este para por el algoritmo una única vez, y es la prueba para conocer qué tanto acierto tiene el algoritmo final.

Esta separación se realiza de forma aleatoria. Hay varios algoritmos para poder hacer la clasificación, los dos de interés dentro de este trabajo son [17]:

- Aprendizaje Profundo: Este algoritmo emplea redes neuronales para poder entrenar al algoritmo. En este tipo de aprendizaje automático, se configuran ciertos parámetros básicos, y se entrena al algoritmo en base a ellos. El entrenamiento debe de preparar al algoritmo para que genere una base de “conocimiento” propio, con el cuál puede reconocer patrones. El fundamento de esto se encuentra en el uso de capas de procesamiento, las cuáles están contenidas dentro de la red neuronal.
- Máquina de Vectores de Soporte: Este algoritmo se encarga de utilizar hiperplanos de dimensión definida para clasificar los distintos datos. La dimensión del hiperplano se define por la cantidad de características que serán computadas. Los vectores de soporte son puntos de datos, los cuáles determinan la posición y orientación del plano.

Aprendizaje Automático no supervisado

La característica de estas técnicas, es que al algoritmo no se le da a conocer cuáles son las salidas esperadas; el punto es buscar patrones. Para poder determinar qué es lo que

se desea predecir, se deben de encontrar estructuras dentro de los datos; para ello, se usa el agrupamiento (o *clustering*) y la asociación. El agrupamiento es un proceso donde se particiona un conjunto de datos en un conjunto de subclases que sean relevantes. Estas se conocen como grupos. La asociación consta en generar un conjunto de características dentro de los datos, las cuáles deben de ser relevantes para la correcta predicción [17].

Metodología

Las señales electroencefalográficas son complejas de identificar: son no periódicas y tienen una relación de señal-ruido bastante baja. En búsqueda de aplicar algoritmos de aprendizaje automático a las anotaciones de un electroencefalograma, la metodología se compone de 8 partes relevantes: el procesamiento de la señal, el entrenamiento de distintos algoritmos, las pruebas para determinar la confianza de cada uno y la implementación visual de los resultados.

Obtención de las señales EEG

Para poder desarrollar los algoritmos de aprendizaje automático que se pretenden emplear, se requieren de señales que cumplan con las anotaciones de interés. Gran parte de estas se obtendrán por parte de la base de datos del concurso *American Epilepsy Society Seizure Prediction Challenge*, encontrado en la página de *Kaggle*. Dentro de ella se encuentran registros de iEEG (electroencefalogramas intra craneales) de un paciente, en los cuáles ya se tienen catalogados en segmentos preictales e interictales, así como un grupo de señales de prueba para los algoritmos [18]. También se emplearán señales proveídas por parte de HUMANA, en donde no se tendrán etiquetas previas, las cuáles servirán para evaluar la funcionalidad del software a realizar para cumplir las expectativas de los encargados.

Procesamiento de las señales

En esta etapa, se empleará de base el reporte técnico *Filtrado Básico de Señales Biomédicas* [13], para poder explorar distintos métodos de filtrado, como el uso de filtros *Butterworth*, *Chebyshev* (filtros IIR) y filtro de Wiener. Se aplicarán y validarán los algoritmos especificados en el reporte técnico en señales EEG de interés. Se usará MATLAB como herramienta principal.

Selección de características de interés

Para que los algoritmos de aprendizaje supervisado tengan un buen desempeño, es requerido que se haga una correcta selección de características para ello. Estas características se obtendrán por medio de métodos en el dominio del tiempo y de tiempo-escala/tiempo-frecuencia. Se pretende explorar características de energía acumulada, ya que en un estudio del año 2001 se estableció que la energía acumulada durante los períodos de referencia es muy similar, mientras que la energía acumulada en los registros previos a la convulsión superan a los de referencia dentro de los 50 minutos antes del inicio de las crisis [19].

Implementación de los algoritmos de aprendizaje automático supervisado

Los algoritmos de aprendizaje automático a emplear serán las redes neuronales y la máquina de vectores de soporte, además de que se evaluarán otras opciones. Esto se realizará con Matlab y distintas toolboxes:

- *Statistics and Machine Learning Toolbox.*
- *Deep Learning Toolbox.*
- *Signal Processing Toolbox.*
- *Wavelet Toolbox.*

Muestra visual de los resultados y anotaciones

Con una clasificación satisfactoria, se generarán figuras de Matlab en donde se mostrará la gráfica del segmento de interés, con algún color a definir (representando el estado ictal, preictal e interictal) y una posible anotación bajo el gráfico.

Implementación en conjunto con la *Epileptic EEG Analysis Toolbox*

Al tener un algoritmo idóneo para la clasificación, y una muestra de resultados visual lo suficientemente clara y adecuada, se añadirá la opción del análisis a profundidad del canal de EEG seleccionado; esto para que el usuario decida si quiere únicamente revisar si se tiene presencia ictal o no (como en la iteración pasada de la *toolbox*) o si se desea obtener más información al respecto.

Validación por sesiones del progreso con HUMANA

Para mantener una retroalimentación activa y cumplir los objetivos, a lo largo del desarrollo del trabajo se agendarán reuniones con los encargados de HUMANA para mostrar y comentar los avances. No se tienen establecidas fechas exactas ni cercanas debido a que esta actividad queda a disponibilidad de horarios de los encargados. Dentro de estas reuniones se espera contrastar las anotaciones encontradas en el documento [3] con las anotaciones que HUMANA realiza rutinariamente con un EEG.

Implementación de los algoritmos de aprendizaje automático no supervisado

Para poder continuar con lo realizado en el trabajo de graduación anterior sobre el uso de aprendizaje no supervisado, se emplearán características diferentes, las cuales fueron cruces por cero, valor medio absoluto, curtosis y desviación estándar. Lo anterior es debido a que uno de los resultados fue que los algoritmos presentaron exactitud, pero no se separaron las

señales de la forma esperada. La idea para la segunda iteración es emplear características de wavelet junto con energía acumulada para realizar algoritmos de agrupamiento.

Cronograma de actividades

Para el cronograma, se establecieron **24** actividades previstas. Estas actividades conforman 8 grupos distintos, los cuales se caracterizan por una letra y un color:

- A. Corresponden a investigación del tema.
- B. Corresponden a la obtención de las señales a utilizar para los algoritmos.
- C. Estas actividades corresponden a la implementación de aprendizaje automático en Matlab.
- D. Corresponden a la generación de anotaciones.
- E. Estas corresponden a la integración de las anotaciones y el aprendizaje automático con la toolbox.
- F. En este grupo están las actividades enfocadas en emplear aprendizaje no supervisado.
- G. Esta actividad se enfoca en las reuniones con los encargados y especialistas de HUMANA para la validación y retroalimentación del trabajo.
- H. Esta actividad es la redacción del documento de tesis final.

Automatización del Proceso de Anotación de Señales EEG de Pacientes con Epilepsia por Medio de Técnicas de Aprendizaje Automático

| Actividad | Abril | | | | | | | Mayo | | | | | | | Junio | | | | | | | Julio | | | | | | | Agosto | | | | | | | Septiembre | | | | | | |
|--|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|--------|--|--|--------|--|--|--|--|--|--|------------|--|--|--|--|--|--|
| | 9-abr | 16-abr | 23-abr | 30-abr | 7-may | 14-may | 21-may | 28-may | 4-jun | 11-jun | 18-jun | 25-jun | 2-jul | 9-jul | 16-jul | 23-jul | 30-jul | 6-ago | 13-ago | 20-ago | 27-ago | 3-sep | 10-sep | 17-sep | 24-sep | 30-sep | | | | | | | | | | | | | | | | |
| A.1 Anotaciones interictales | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| A.2 Anotaciones preictales | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| A.3 Morfología, tiempo y abundancia de las señales | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| A.4 Machine Learning (Red Neuronal) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| A.5 Morfología, tiempo y abundancia de las señales | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| B.1 Selección de las señales pre/interictales | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| B.2 Obtención de EEGs sanos | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| B.3 Aplicación de filtros para las señales | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| C.1 Selección de las características | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| C.2 Generar un set de entrenamiento para el algoritmo | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| C.3 Obtención de las características | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| C.4 Hacer ML con una red neuronal | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| C.5 Hacer ML con una SVM | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| D.1 Crear ventana de matlab con la ventana de interés | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| D.2 Graficar con colores las partes importantes | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| D.3 Mostrar en texto el tipo de señal encontrada | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| E.1 Generar la ventana de la app en la toolbox | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| E.2 Integrar el ML a la ventana de la app | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| E.3 Integrar las anotaciones a la app | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| F.1 Establecer la cantidad de categorías a emplear | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| F.2 Selección de las características | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| F.3 Implementación y aplicación de técnicas de agrupamiento | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| G.1 Reuniones para compartir avances y obtener retroalimentación | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H.1 Creación y redacción del documento final de tesis | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Dentro del cronograma, el color verde indica que la actividad ha finalizado, y el color rojo indica que la actividad aún no ha sido terminada.

Para comprender el estudio, se detallan las actividades:

- A.1-A.4** Estas actividades constaron de una investigación y revisión bibliográfica para obtener los conocimientos requeridos para trabajar en las distintas etapas del trabajo de graduación.
- B.1-B.2** La selección de las señales a emplear es importante, puesto que deben de cumplir con los requisitos para ser útiles en el trabajo. Las señales seleccionadas fueron las brindadas por Kaggle, ya que son de registros preictales e interictales. En específico, las preictales se brindan en un rango de 60 minutos a 10 minutos antes del episodio ictal. La obtención de EEG sanos permite ser una referencia al momento del estudio, para evitar caracterizar una señal de un paciente con epilepsia de uno sano.
- B.3** Se pretende aplicar una etapa de filtrado a las señales EEG, removiendo frecuencias fuera de un rango de interés. Es de cuidado el empleo de filtros, puesto que el filtrar de más hace que se pierda información valiosa de la señal [13].
- C.1-C.3** Estas actividades constan en seleccionar las características que serán empleadas en el aprendizaje automático, estas se deben de seleccionar en base a bibliografías, así como tener un método de extraerlas. Para el entrenamiento del algoritmo, se debe de establecer un set de señales que sean para entrenar al algoritmo, por lo que se establecerán qué señales formarán parte de este grupo.
- C.4-C.5** Estas actividades se enfocan directamente en el entrenamiento y validación de los algoritmos de aprendizaje automático a emplear en el trabajo.
- D.1-D.3** En estas se pretende generar las ventanas en gráficos para su visualización, dando información pertinente para que el especialista lea la propuesta de anotación que el software ha generado.
- E.1-E.3** Estas actividades se enfocan en implementar los resultados obtenidos en el trabajo a la *Epileptic EEG Analysis Toolbox*, generando una ventana adicional con la misma información mostrada en los gráficos generados en las actividades E.
- F.1-F.3** Se proponen estas actividades para seguir con el estudio sobre la implementación de aprendizaje no supervisado a la clasificación de señales.
- G.1** Actividad referida a las reuniones esporádicas con los encargados de HUMANA para la retroalimentación.
- H.1** Actividad referida a la redacción del documento de tesis final.

Índice preliminar

| | No. Página |
|---|------------|
| Prefacio | ii |
| Lista de figuras | iii |
| Lista de cuadros | iv |
| Resumen | v |
| Abstract | vi |
| 1. Introducción | 1 |
| 2. Antecedentes | 2 |
| 3. Justificación | 3 |
| 4. Objetivos | 4 |
| 4.1. Objetivo General | 4 |
| 4.2. Objetivos Específicos | 4 |
| 5. Alcance | 5 |
| 6. Marco Teórico | 6 |
| 6.1. Señales Electroencefalográficas (EEG) | 6 |
| 6.1.1. Origen de la actividad eléctrica del cerebro | 7 |
| 6.1.2. Obtención de un EEG | 8 |
| 6.1.3. Posicionamiento de Electroodos | 8 |
| 6.1.4. Características de un EEG | 9 |
| 6.2. Epilepsia | 10 |
| 6.2.1. Clasificación y Tipos de Epilepsia | 10 |

| | |
|---|-----------|
| 6.3. Procesamiento de señales EEG | 11 |
| 6.3.1. Características de un EEG en el dominio del tiempo | 12 |
| 6.3.2. Características de un EEG en el dominio de la frecuencia | 13 |
| 6.3.3. Características de un EEG en el dominio de tiempo-frecuencia | 14 |
| 6.3.4. Filtrado de las señales EEG | 15 |
| 6.4. Anotaciones de un EEG | 16 |
| 6.5. Aprendizaje Automático (Machine Learning) | 17 |
| 6.5.1. Tipos de aprendizaje automático | 17 |
| 6.6. Aprendizaje Automático Supervisado | 18 |
| 6.6.1. Redes Neuronales (RN) | 18 |
| 6.6.2. Máquinas de Vectores de Soporte (SVM) | 19 |
| 6.7. Aprendizaje Automático No Supervisado | 18 |
| 11. Resultados | 19 |
| 12. Conclusiones | 20 |
| 13. Recomendaciones | 21 |
| Bibliografía | 22 |

Referencias

- [1] Epilepsy Foundation Communications, *International Epilepsy Day*, <https://www.epilepsy.com/make-difference/public-awareness/international-epilepsy-day>, Accedido: 23-03-2021, 2020.
- [2] HUMANA, <https://humanagt.org/>, Accedido: 05-05-2021, 2015.
- [3] L. E. Morillo, “Análisis Visual del Electroencefalograma,” en *Guia Neurologica 7 - Neuro Electro Diagnóstico*. Colombia: Asociación Colombiana de Neurología, 2005, págs. 143-163.
- [4] M. J. Angulo, “Análisis y Reconocimiento de Patrones de Señales Biomédicas de Pacientes con Epilepsia,” Tesis de licenciatura, Universidad Del Valle de Guatemala, 2020.
- [5] 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.

- [6] HUMANA, *Epilepsia*, <https://humanagt.org/epilepsia/>, Accedido: 02-05-2021, 2015.
- [7] S. Villazana, A. Eblen, G. Montilla y C. Seijas, "Detección de cambios abruptos en señales epilépticas en EEG," *Archivos de Neurociencias*, vol. 22, n.º 3, págs. 6-18, 2017.
- [8] Epilepsy Foundation, *What is Epilepsy?* <https://www.epilepsy.com/learn/about-epilepsy-basics/what-epilepsy>, Accedido: 28-03-2021, 2014.
- [9] P. RM, C. MR, M. EM y et al, "The ILAE classification of seizures and the epilepsies: Modification for seizures in the neonate," *Epilepsia*, n.º 00, págs. 1-14, 2021.
- [10] T. M. Vaughan, J. R. Wolpaw y E. Donchin, "EEG-based communication: prospects and problems," *IEEE Transactions on Rehabilitation Engineering*, vol. 4, n.º 4, págs. 425-430, 1996. DOI: 10.1109/86.547945.
- [11] D. Plass-Oude Bos, "EEG-based Emotion Recognition," *The Influence of Visual and Auditory Stimuli*, ene. de 2006.
- [12] Vallar, R., *Compute the average bandpower of an EEG signal*, <https://raphaelvallat.com/bandpower.html>, Accedido: 28-03-2021, 2018.
- [13] J. González Murillo, *Filtrado Básico de Señales Biomédicas*, mayo de 2014. DOI: 10.13140/2.1.1757.2168.
- [14] F. D. Q. Quinga, "Predicción de crisis epilépticas utilizando técnicas de procesamiento de señales electroencefalográficas y aprendizaje de máquina," Tesis de licenciatura, Universidad de las Fuerzas Armadas, Ecuador, 2017.
- [15] A. N. B. Cicchino, "Técnicas de procesamiento de EEG para detección de eventos," Tesis de maestría, Universidad Nacional de la Plata, 2013.
- [16] G. I. G. Gazabón, "Modelo de Machine Learning para la Clasificación de pacientes en términos del nivel asistencial requerido en una urgencia pediátrica con Área de Cuidados Mínimos," Tesis de maestría, Universidad Tecnológica de Bolívar, 2014.
- [17] J. F. V. García, "Modelo de pronóstico de rendimiento académico de alumnos en los cursos del programa de estudios básicos de la Universidad Ricardo Palma usando algoritmos de Machine Learning," Tesis de maestría, Universidad Ricardo Palma, Perú, 2019.
- [18] Kaggle, *American Epilepsy Society Seizure Prediction Challenge*, <https://www.kaggle.com/c/seizure-prediction/data>, Accedido: 10-05-2021, 2015.
- [19] B. Litt, R. Esteller, J. Echaz y et al., "Epileptic seizures may begin hours in advance of clinical onset: a report of five patients.," *Neuron*, vol. 30, n.º 1, págs. 51-64, 2001.