

Detector Multimodal de Crisis Epilépticas Motoras y no Motoras para Dispositivos Wearable Basado en Actividad Electrodermal, Acelerometría y Fotopletismografía



Edward Alexis Molina Rivera

Tesis de Maestría en Ingeniería Telemática

Director:

PhD. Diego Mauricio López Gutiérrez

Universidad del Cauca

Facultad de Ingeniería Electrónica y Telecomunicaciones

Departamento de Telemática

Línea de investigación en e-Salud

Popayán, Noviembre de 2019

Edward Alexis Molina Rivera

Detector Multimodal de Crisis Epilépticas Motoras y no Motoras para Dispositivos Wearable Basado en Actividad Electrodermal, Acelerometría y Fotopletismografía

Tesis presentada a la Facultad de Ingeniería
Electrónica y Telecomunicaciones de la
Universidad del Cauca para la obtención del
Título de

Magíster en:
Ingeniería Telemática

Director:
PhD. Diego Mauricio López Gutiérrez

Popayán
2020

Agradecimientos

“Confía en el Señor con todo tu corazón; no dependas de tu propio entendimiento. Busca su voluntad en todo lo que hagas, y él te mostrará cuál camino tomar.” – Proverbios 3:5-6

No puedo pasar por alto la oportunidad de agradecer a Dios, quien me ha guiado no sólo en todo este proceso de Maestría, sino también a lo largo de mi vida. A Él gracias por alcanzarme con su amor y su gracia, y por permitirme descubrir en Él el verdadero significado y propósito de la vida. Gracias a Diego López, PhD, por su tutoría y por la confianza depositada desde el principio. Gracias a María Eugenia Miño, Neuropediatra, por su asesoría en temas de Neurología y Epilepsia. Gracias a Axon Ped S.A.S. de Popayán, Colombia, por facilitar la recolección de señales de pacientes para experimentación. Gracias a Björn Eskofier, Prof. Dr., y a su equipo del *Machine Learning and Data Analytics Lab* de Erlangen, Alemania, por hacer posible mi estancia de investigación, por su retroalimentación en temas de Inteligencia Artificial y por proveer los datos utilizados en la investigación. Gracias a la Universidad del Cauca, y en específico, al Grupo de Ingeniería Telemática y a los integrantes del proyecto NeuroMoTIC por su apoyo y por hacer posible la investigación. Gracias a Colciencias por la financiación proporcionada en el marco de la Convocatoria 715-2015, *Convocatoria para Proyectos de Investigación y Desarrollo en Ingenierías*, proyecto NeuroMoTIC: *Sistema móvil para el Apoyo Diagnóstico de la Epilepsia*, contrato número FP44842-154-2016. Finalmente, gracias a mis padres y hermanos, por ser el motor en cada etapa de mi vida; y a mis amigos, por estar presentes y por la motivación dada a lo largo de este proceso. A cada uno de los mencionados, mi más sincera gratitud.

Resumen estructurado

Antecedentes: El diagnóstico y tratamiento de la epilepsia se basa principalmente en el conteo y seguimiento de las crisis epilépticas. Los dos métodos más utilizados para este propósito son los auto-informes y la video-Electroencefalografía (v-EEG). Las limitaciones de estos enfoques incluyen la baja calidad de los informes realizados por los pacientes y la incomodidad percibida por los pacientes en los estudios de v-EEG, así como sus altos costos. Debido a lo anterior, varios sensores y dispositivos portátiles han surgido como tecnologías alternativas para la detección de crisis epilépticas; sin embargo, los enfoques existentes presentan bajos resultados en la detección de crisis epilépticas no motoras, en términos de precisión y tasa de falsos positivos.

Objetivo: El propósito de esta disertación es implementar un modelo de aprendizaje automático multimodal para la detección de crisis epilépticas motoras y no motoras basado en datos obtenidos de sensores de acelerometría (ACC), actividad electrodermal (EDA) y fotopleletismografía (PPG).

Métodos: Siguiendo un proceso de análisis de datos, se describieron, pre-procesaron y caracterizaron señales de 11 pacientes epilépticos con 44 crisis epilépticas. Diez algoritmos de aprendizaje automático fueron entrenados para clasificar segmentos de 10 segundos en las clases crisis epiléptica o no crisis epiléptica. Se tuvieron en cuenta crisis epilépticas de diferentes semiologías (síntomas), incluidas las motoras y no motoras. El mejor modelo se evaluó de acuerdo a métricas de rendimiento y de acuerdo a estándares para la validación clínica de dispositivos de detección de crisis epilépticas. Finalmente, se llevó a cabo el despliegue del mejor modelo de aprendizaje automático.

Resultados: Los principales resultados obtenidos fueron: 1) Un dataset de señales EDA, ACC y PPG de pacientes con crisis epilépticas de diferente semiología. 2) Un modelo de aprendizaje automático para la detección de crisis epilépticas motoras y no motoras, basado en señales EDA, ACC y PPG. 3) Evaluación del modelo de aprendizaje automático: El algoritmo Random Forest logró el mejor rendimiento y,

después de ser evaluado, obtuvo una exactitud del 99%, una precisión del 97%, una sensibilidad del 98% y una medida F del 97%. De acuerdo con los estándares para la validación clínica de dispositivos de detección de crisis epilépticas, el modelo presentado clasificó como un modelo de Fase 2. 4) Aplicación móvil de Android para la detección automática de eventos epileptiformes. Esta aplicación está conectada a una pulsera de tipo wearable y está integrada al modelo de aprendizaje automático.

Conclusiones: El modelo de aprendizaje automático propuesto puede clasificar crisis epilépticas motoras y no motoras con un nivel aceptable de precisión y con una tasa baja de falsos positivos. El modelo presentado podría contribuir al desarrollo de sistemas de circuito cerrado (*closed-loop*) que podrían mejorar el monitoreo continuo de las crisis epilépticas en el futuro. La aplicación móvil desarrollada para la detección automática de eventos epileptiformes tiene el potencial de contribuir al monitoreo de las crisis epilépticas, como una alternativa al método tradicional de auto-informe. De esta manera, la aplicación tiene el potencial de proporcionar retroalimentación a pacientes, cuidadores y médicos.

Palabras clave: detección de crisis epilépticas, aprendizaje automático, wearable, acelerometría, actividad electrodermal, fotopletismografía.

Structured abstract

Background: Epilepsy diagnosis and treatment is based mainly on counting and monitoring seizures. The two most commonly used methods for this purpose are self-reports and video-Electroencephalography (v-EEG). Limitations of these approaches include low quality of the reports made by patients, and the discomfort that a v-EEG study represents for patients, as well as its high costs. Due to the above, various wearable sensors and portable devices have emerged as alternative technologies for seizure detection; nevertheless, existing approaches show low results in the classification in terms of accuracy and false positive rate when they try to detect non-motor seizures.

Aim: The purpose of this dissertation is to implement a multimodal machine learning model for motor and non-motor seizure detection based on data obtained from accelerometry (ACC), electrodermal activity (EDA), and photoplethysmography (PPG) sensors.

Methods: Following a process of data analytics, signals of 11 epileptic patients with 44 epileptic seizures were described, preprocessed, and characterized. Ten machine learning algorithms were trained to classify 10-second segments in non-seizure or seizure classes. Seizures of different semiology (symptoms), including motor and non-motor ones, were taken into account. The best model was evaluated according to performance metrics and according to standards for clinical validation of epileptic seizure detection devices. Finally, the deployment of the best machine learning model was carried out.

Results: The main results obtained were: 1) A dataset of EDA, ACC, and PPG signals from patients with seizures of different semiology. 2) A machine learning model for the detection of motor and non-motor epileptic seizures based on EDA, ACC, and PPG signals. 3) Evaluation of the machine learning model: The Random Forest algorithm achieved the best performance, and after being evaluated, it obtained an accuracy of 99%, precision of 97%, recall of 98%, and F-Measure of 97%. According to standards for clinical validation of epileptic seizure detection

devices, the presented model was classified as a Phase 2 model. 4) Android mobile application for automatic detection of epileptiform events. This application is connected to a wearable wristband and it is integrated to the machine learning model.

Conclusions: The proposed machine learning model can classify motor and non-motor epileptic seizures with an acceptable level of precision and with a low false-positive rate. The model herein presented could contribute to the development of closed-loop systems that could improve the continuous monitoring of seizures in the future. The developed mobile application for automatic detection of epileptiform events has the potential to contribute to the monitoring of epileptic seizures, as an alternative to the traditional method of self-report. In such a way, it has the potential to provide feedback to patients, caregivers, and physicians.

Keywords: seizure detection, machine learning, wearable, accelerometry, electrodermal activity, photoplethysmography.

Contenido

Capítulo 1	1
1. Introducción	1
1.1. Planteamiento del problema.....	1
1.2. Pregunta de investigación e hipótesis	3
1.3. Motivación	4
1.4. Objetivos	5
1.4.1. Objetivo General.....	5
1.4.2. Objetivos Específicos	5
1.5. Metodología	5
1.6. Contenido de la Disertación.	7
Capítulo 2.....	9
2. Comprensión del Negocio.....	9
2.1. Identificación del objetivo del negocio	10
2.1.1. Antecedentes (<i>background</i>).....	10
2.1.2. Objetivo del negocio	16
2.1.3. Criterios de éxito del negocio	16
2.2. Evaluación de la situación.....	16
2.2.1. Inventario de recursos	16
2.2.2. Riesgos y contingencias	18
2.3. Identificación del objetivo de la minería de datos	18
2.3.1. Objetivo de la minería de datos	18
2.3.2. Criterios de éxito de la minería de datos	19
2.4. Realización del plan del proyecto.....	19
2.4.1. Plan del proyecto	19

2.4.2. Evaluación inicial de herramientas	20
Capítulo 3.....	22
3. Comprensión de los datos	22
3.1. Definición del marco conceptual	23
3.1.1. Marco conceptual	23
3.2. Recolección de los datos iniciales.....	25
3.2.1. Reporte de recolección de los datos iniciales.....	25
3.3. Descripción de los datos	27
3.3.1. Reporte de descripción de los datos.....	27
3.4. Exploración de los datos	31
3.4.1. Reporte de exploración de los datos	31
3.5. Verificación de la calidad de los datos	33
3.5.1. Reporte de la calidad de los datos.....	33
Capítulo 4.....	36
4. Preparación de los datos	36
4.1. Selección de los datos	37
4.1.1. Datos seleccionados.....	37
4.1.2. Características seleccionadas.....	39
4.2. Limpieza y pre-procesamiento de los datos	41
4.3. Construcción de los datos	43
4.3.1. Caracterización de señales.	43
4.4. Integración y formateo de los datos	46
4.4.1. Descripción del dataset preparado	47
Capítulo 5.....	49
5. Modelado y Evaluación.....	49
5.1. Selección de los algoritmos de clasificación	49
5.2. Modelado y resultados	51
5.3. Evaluación del modelo	53

5.3.1. Métricas de rendimiento	53
5.3.2. Evaluación del rendimiento.....	55
5.4. Evaluación de la calidad del proceso y del modelo desde la perspectiva clínica.....	57
5.4.1. Evaluación de acuerdo a estándares de validación clínica.....	57
Capítulo 6.....	62
6. Despliegue.....	62
6.1. Arquitectura de despliegue.....	63
6.1.1. Descripción del funcionamiento del sistema detector	63
6.1.2. Arquitectura de despliegue del sistema detector.	64
6.2. Aplicación móvil para la detección automática de eventos epileptiformes	65
Capítulo 7	69
7. Informe final	69
7.1. Contribuciones	69
7.2. Conclusiones.....	71
7.3. Trabajos futuros	72
Referencias	73
Anexo A.....	79
Consentimiento Informado	79
Anexo B.....	83
Protocolo de recolección de señales EDA, ACM, HR y EEG	83

Lista de Figuras

Figura 1. Clasificación de los tipos de crisis epilépticas, versión expandida.	3
Figura 2. Fases de CRISP-DM.	6
Figura 3. Recolección inicial de señales de EDA, ACC, PPG, TEMP y EEG, pertenecientes a un paciente pediátrico.	26
Figura 4. Pulsera E3 wristband.	28
Figura 5. Exactitud alcanzada por cada algoritmo entrenado con parámetros por defecto en una configuración de validación cruzada.	52
Figura 6. Matriz de confusión de la evaluación del modelo de <i>Machine Learning</i> (Random Forest).	56
Figura 7. Componentes y funcionamiento general del sistema automático para la identificación de eventos epileptiformes.	63
Figura 8. Arquitectura de despliegue del sistema detector.	65
Figura 9. Capturas de pantalla de las interfaces iniciales de la aplicación móvil. .	66
Figura 10. Capturas de pantalla de las interfaces para la recolección de señales ACC, EDA y PPG.	67
Figura 11. Capturas de pantalla de las interfaces de visualización de resultados de la detección.	68

Lista de Tablas

Tabla 1. Tareas ejecutadas y resultados alcanzados en la fase Comprensión del negocio.....	9
Tabla 2. Detectores de crisis epilépticas que pueden ser utilizados con dispositivos wearable o dispositivos portables.....	15
Tabla 3. Recursos hardware disponibles en el transcurso de la investigación.	17
Tabla 4. Plan general del proyecto considerando fases y tareas de la metodología CRISP-DM.....	20
Tabla 5. Tareas ejecutadas y resultados alcanzados en la fase Comprensión de los datos.	22
Tabla 6. Frecuencia de muestreo y rango de operación de los sensores de la pulsera E3.....	29
Tabla 7. Archivos contenidos en cada carpeta del dataset de señales de la pulsera E3 wristband.....	31
Tabla 8. Principales características de los pacientes epilépticos contenidos en el dataset	32
Tabla 9. Número de crisis epilépticas identificadas por cada paciente, discriminadas por semiología.....	32
Tabla 10. Número total de crisis epilépticas por cada semiología.	33
Tabla 11. Tareas ejecutadas y resultados alcanzados en la fase Preparación de los datos.	37
Tabla 12. Número de crisis epilépticas incluidas en el dataset preparado, discriminadas por semiología	38

Tabla 13. Número de crisis epilépticas incluidas en el dataset preparado, agrupadas por semiología.....	38
Tabla 14. Características seleccionadas con base en el trabajo de Knoth [41]. ...	41
Tabla 15. Estructura de carpetas de las señales extraídas.	43
Tabla 16. Estructura de archivo de datasets de entrenamiento y prueba (formato CSV).....	48
Tabla 17. Tareas y resultados de las fases Modelado y Evaluación.....	49
Tabla 18. Algoritmos de clasificación seleccionados y sus parámetros por defecto.	51
Tabla 19. Resultados de métricas de entrenamiento de cada uno de los algoritmos seleccionados, utilizando el método validación cruzada estratificada con 10 pliegues.....	52
Tabla 20. Matriz de confusión y valores calculados.....	53
Tabla 21. Métricas de evaluación del modelo de <i>Machine Learning</i> (Random Forest).....	56
Tabla 22. Fases del desarrollo y validación clínica de SDD, según características clave.....	58
Tabla 23. Evaluación de la calidad del proceso y del modelo final desde la perspectiva clínica según los estándares propuestos por Beniczky y Ryvlin [67].	59
Tabla 24. Medidas de resultado correspondientes al proceso de minería de datos y al modelo final.	60
Tabla 25. Tareas y resultados de la fase de despliegue.....	62

Lista de Acrónimos

ILAE: International League Against Epilepsy / Liga Internacional Contra la Epilepsia

LMIC: Low and Medium Income Countries / Países de Bajos y Medianos Ingresos

v-EEG: Video-Electroencephalography / Video-Electroencefalografía

ACC: Accelerometry / Acelerometría

EDA: Electrodermal Activity / Actividad Electrodermal

PPG: Photoplethysmography / Fotopletismografía

EEG: Electroencephalography / Electroencefalografía

CRISP-DM: Cross Industry Standard Process for Data Mining / Proceso Estándar de la Industria Cruzada para Minería de Datos

PNES: Psychogenic Non-Epileptic Seizures / Convulsiones Psicogénicas no Epilépticas

ECG: Electrocardiography / Electrocardiografía

PPV: Positive Predictive Value / Valor Predictivo Positivo

CPS: Complex Partial Seizures / Crisis Parciales Complejas

GPS: Generalized Partial Seizures / Crisis Parciales Generalizadas

GTCS: Generalized Tonic Clonic Seizures / Crisis Tónico Clónicas Generalizadas

TLE: Temporal Lobe Epilepsy / Epilepsia del Lóbulo Temporal

HR: Heart Rate / Ritmo Cardíaco

SpO₂: Arterial oxygenation / Oxigenación arterial

TEMP: Temperature / Temperatura

IPS: Health care provider / Instituto Prestador de Servicios de Salud

CSV: Comma Separated Values / Valores Separados por Comas

IBI: Inter Beat Interval / Intervalo entre latidos cardíacos

SCL: Skin Conductance Level / Nivel de Conductividad de la Piel

SCR: Skin Conductance Response / Respuesta de la Conductividad de la Piel

FPR: False Positive Rate / Tasa de Falsos Positivos

BLE: Bluetooth Low Energy / Bluetooth de Baja Energía

Capítulo 1

1. Introducción

1.1. Planteamiento del problema

De acuerdo a la Liga Internacional Contra la Epilepsia (ILAE, por sus siglas en inglés), la epilepsia es una enfermedad cerebral que se define por cualquiera de las siguientes circunstancias: (1) Al menos dos crisis no provocadas (o reflejas) con más de 24 horas de separación; (2) una crisis no provocada (o refleja) y una probabilidad de presentar nuevas crisis durante los 10 años siguientes que sea similar al riesgo general de recurrencia de pacientes que hayan sufrido dos crisis no provocadas (al menos el 60%); (3) diagnóstico de un síndrome de epilepsia [1]. Una crisis epiléptica es una aparición transitoria de signos y/o síntomas provocados por una actividad neuronal anómala excesiva o simultánea en el cerebro [2].

Aproximadamente 50 millones de personas en el mundo tienen epilepsia y cerca del 80% de estas personas viven en países de bajos y medianos ingresos (LMIC, por sus siglas en inglés); además, alrededor de las tres cuartas partes de quienes viven en LMIC no reciben el tratamiento que necesitan [3]. En Colombia la prevalencia general de la epilepsia es de 11,3 por cada 1000 habitantes [4], con solo un 30% de la población controlada [5].

Una vez diagnosticada, la epilepsia puede controlarse con fármacos antiepilépticos en cerca del 70% de los pacientes, mientras que el 30% restante son resistentes al tratamiento farmacológico [6]. La práctica clínica de la epilepsia se basa en gran parte en el conteo y seguimiento de las crisis epilépticas [7], [8]. De acuerdo a J. Bidwell et al., la información más relevante que los neurólogos necesitan conocer

para el tratamiento de la epilepsia (selección inicial de un medicamento anti epiléptico y ajustes posteriores) es: tipo de crisis epiléptica, caracterización del movimiento y frecuencia de las crisis epilépticas [9]. Por lo general se les pide a los pacientes o cuidadores llevar un diario de las crisis epilépticas con esta información, el cual ayudará al neurólogo en la toma de decisiones clínicas [6], [9]; este método es conocido también como auto-informe.

Los auto-informes producen fallos entre el 30% y 50% al reportar crisis epilépticas ocurridas en el día, y superiores al 85% en crisis nocturnas; además, se ha reportado que no hay un consenso en los testigos en datos importantes de las crisis y se evidencia incomodidad en los pacientes al mantener un registro de sus episodios y síntomas [9]. El *gold-standard* para obtener información precisa de las crisis epilépticas es la monitorización de video-Electroencefalografía (v-EEG) por parte de un neurólogo o epileptólogo. Para realizar una v-EEG el paciente debe permanecer en el hospital con electrodos en la cabeza y este procedimiento puede tomar hasta 3 días en completarse. Esto causa que sea un procedimiento molesto para los pacientes, además es muy costoso en términos económicos tanto para las clínicas como para los pacientes, por último pocos neurólogos tienen acceso a los dispositivos e infraestructura necesaria [10]. En Colombia, por ejemplo, una hora de v-EEG puede costar de \$300.000 a \$700.000.

Debido a la baja precisión de los auto-informes al reportar crisis epilépticas y a las dificultades relacionadas con la v-EEG (dentro de las que se encuentran la necesidad de un neurólogo o epileptólogo y la dificultad para ser usada en contextos de la vida cotidiana), han surgido sistemas detectores de crisis epilépticas alternativos a los basados en Electroencefalografía (EEG), los cuales tienen el potencial de ser automáticos, portables, ambulatorios y no invasivos [11]–[13]. Éstos se usan para detectar crisis epilépticas por medio de la medición de señales biológicas relacionadas con cambios fisiopatológicos producidos por las crisis epilépticas [13] y pueden ser usados como herramientas de diagnóstico o de monitoreo continuo. Estos sistemas se han basado principalmente en la medición modal (un solo sensor o variable) o multimodal (más de un sensor o variable) de acelerometría (ACC, por sus siglas en inglés), movimiento, actividad electrodermal (EDA, por sus siglas en inglés), sonido, electromiografía, frecuencia cardíaca y/o video [14].

Las crisis epilépticas más específicas están divididas en dos grupos: crisis de inicio motor (*motor onset*) y crisis de inicio no motor (*nonmotor onset*) (Figura 1). Un

sistema ideal alternativo a los auto-informes y a la v-EEG para la detección de crisis epilépticas debería identificar todos los tipos posibles de crisis.

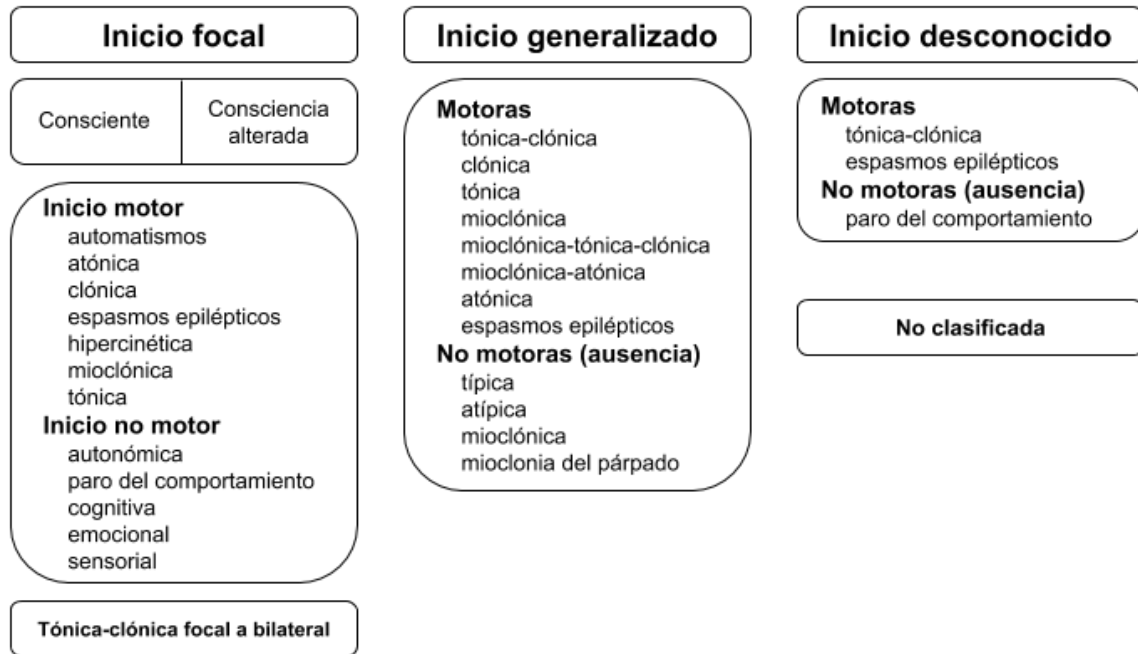


Figura 1. Clasificación de los tipos de crisis epilépticas, versión expandida.

Fuente: Traducida al español de Liga Internacional Contra la Epilepsia [15]

A pesar de que las crisis no motoras representan una cantidad importante de los tipos de crisis, la mayoría de los sistemas detectores de crisis epilépticas alternativos a los basados en EEG están contruidos para la detección de crisis de tipo motor; adicionalmente, los sistemas que consideran tanto crisis motoras como no motoras presentan bajos resultados de detección en términos de falsos positivos y precisión, tal como se muestra en el estado del arte presentado en la Sección 2.1.1.

1.2. Pregunta de investigación e hipótesis

Teniendo en cuenta lo mencionado anteriormente y el estado del arte presentado en la Sección 2.1.1, esta disertación aborda la siguiente pregunta de investigación: *¿Cómo detectar crisis epilépticas motoras y no motoras de una forma automática, ambulatoria y no invasiva?*

Con el fin de abordar la pregunta de investigación planteada, esta disertación de maestría establece como hipótesis que es posible identificar de forma automática crisis epilépticas motoras y no motoras por medio de un detector multimodal basado en técnicas de *Machine Learning* para dispositivos *wearable* que incorpore la medición de EDA, ACC y Fotopleletismografía (PPG, por sus siglas en inglés).

Diferentes estudios muestran el potencial de la EDA y la ACC [7], [12], [16], [14], así como de la PPG [12], [17], [13], para la detección automática de crisis epilépticas. En la mayoría de los casos, estas señales han sido medidas por medio de dispositivos *wearable* (pulseras, por ejemplo), los cuales cumplen con las características deseadas para la construcción de sistemas automáticos, portables, ambulatorios y no invasivos.

1.3. Motivación

En la comunidad científica internacional existe un gran interés en la construcción de sistemas detectores de crisis epilépticas que sean alternativos a los basados en EEG y, además, que sean automáticos, portables, ambulatorios y no invasivos. Por otro lado, en el contexto regional y nacional, el proyecto NeuroMoTIC identificó la necesidad del monitoreo continuo de crisis epilépticas como un mecanismo de apoyo a los neurólogos en la gestión de tratamientos farmacológicos de sus pacientes. El monitoreo cobra importancia en pacientes de difícil manejo de la epilepsia, ya que permitiría la identificación automática de crisis en un periodo largo de tiempo, en contraste con las crisis que se pueden identificar en el corto tiempo que toma una lectura convencional de EEG (30 minutos generalmente). En este sentido, la investigación desarrollada y reportada en esta disertación es relevante tanto para la comunidad internacional, como para NeuroMoTIC, ya que posibilita la construcción de sistemas de monitoreo continuo y automático de pacientes epilépticos, los cuales podrían impactar positivamente en la efectividad de los tratamientos y en la calidad de vida de los pacientes.

Esta disertación de maestría está enmarcada en el proyecto NeuroMoTIC (NeuroMoTIC: Sistema móvil para el Apoyo Diagnóstico de la Epilepsia), el cual es un proyecto de investigación financiado por Colciencias (Convocatoria 715 de 2015, Contrato FP44842-154-2016) y ejecutado por la Universidad del Cauca (VRI ID

4427), el cual tiene como objetivo construir un sistema que permita la recolección, gestión y clasificación de información clínica y señales EEG para soportar el diagnóstico remoto de la epilepsia. Este sistema está pensado principalmente para su uso en áreas rurales o de difícil acceso. Hace uso de técnicas de Inteligencia Artificial para identificar crisis epilépticas en exámenes EEG y de esta manera, ayudar a un neurólogo en la toma de decisiones para que pueda diagnosticar la epilepsia de una forma precisa y rápida.

1.4. Objetivos

1.4.1. Objetivo General

Proponer un detector multimodal de crisis epilépticas motoras y no motoras basado en técnicas de *Machine Learning* para dispositivos wearable que incorporen EDA, ACC y PPG.

1.4.2. Objetivos Específicos

1. Construir un dataset de señales EDA, ACC y PPG de pacientes con crisis epilépticas.
2. Construir un clasificador para la detección de crisis epilépticas motoras y no motoras basado en EDA, ACC y PPG.
3. Evaluar la precisión del clasificador en la detección de crisis epilépticas.

1.5. Metodología

Esta disertación siguió formalmente la metodología *Cross Industry Standard Process for Data Mining* (CRISP-DM) [18], teniendo en cuenta que el objetivo general y los objetivos específicos pueden desarrollarse en el contexto de la minería de datos. CRISP-DM está descrita en términos de un modelo de proceso jerárquico,

consistente en un conjunto de tareas descritas en cuatro niveles de abstracción (de lo general a lo específico): fase, tarea genérica, tarea especializada e instancia de proceso. En el nivel superior, el proceso de minería de datos está organizado en seis fases (Figura 2).

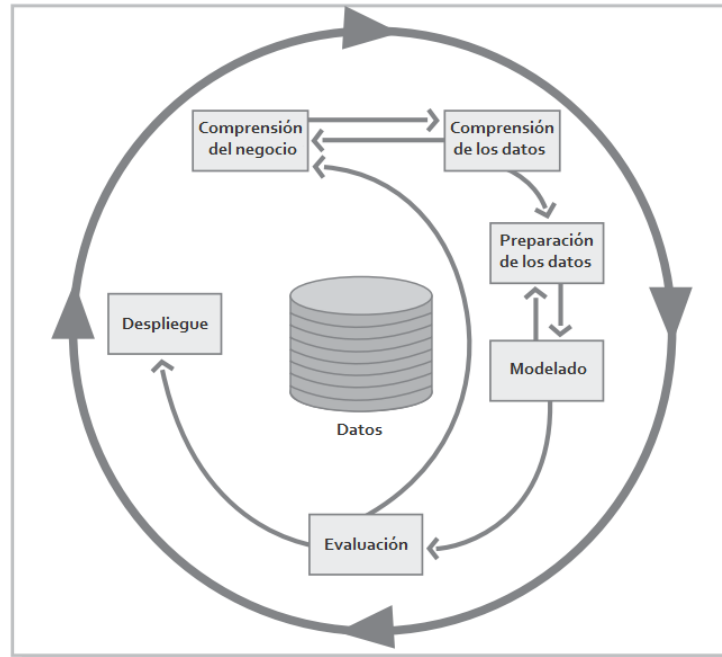


Figura 2. Fases de CRISP-DM.

Fuente: Traducido al español de CRISP-DM 1.0 [18]

Comprensión del negocio (*Business Understanding*): Incluye la comprensión de los objetivos desde un punto de vista del negocio, la identificación del objetivo de minería de datos y la realización de un plan para la ejecución de las fases siguientes.

Comprensión de los datos (*Data Understanding*): En esta fase se describe la estructura de los datos a ser utilizados, se realiza una exploración inicial de los mismos y una verificación de la calidad.

Preparación de los datos (*Data Preparation*): Esta fase comprende la selección, limpieza y pre-procesamiento de los datos a ser utilizados, con el fin de adaptarlos para aplicar diferentes algoritmos de *Machine Learning*.

Modelado (*Modeling*): Durante esta fase se aplican diferentes algoritmos de clasificación sobre los datos, con el fin de seleccionar el algoritmo y la configuración de parámetros que presente el mejor rendimiento.

Evaluación (*Evaluation*): En esta fase se evalúa el algoritmo seleccionado en la fase anterior, en términos de su rendimiento y del cumplimiento del objetivo de la minería de datos.

Despliegue (*Deployment*): El propósito de esta fase es presentar el modelo de una forma agradable al usuario final.

Teniendo en cuenta que CRISP-DM está diseñada para utilizarse y adaptarse para diferentes contextos del análisis de datos, en esta disertación las tareas genéricas pertenecientes a las seis fases de CRISP-DM fueron adaptadas de acuerdo a las necesidades específicas del análisis de datos para la problemática planteada.

1.6. Contenido de la Disertación.

El Capítulo 1 corresponde al capítulo introductorio. Los Capítulos 2 – 6 están divididos de acuerdo a las seis fases de la metodología CRISP-DM. Cada uno de ellos contiene una breve descripción de la fase y las tareas realizadas en ella.

- El Capítulo 1 presenta el planteamiento del problema, la pregunta de investigación e hipótesis, motivación, objetivos, metodología y la estructura general de esta disertación.
- El Capítulo 2 describe las tareas realizadas en la fase Comprensión del negocio, dentro de las que se encuentran la identificación de los objetivos desde el punto de vista del negocio y la minería de datos, los antecedentes (estado del arte) y el plan del proyecto.
- El Capítulo 3 presenta la fase Comprensión de los datos, la cual comprende el marco conceptual, la descripción y exploración de los datos utilizados en el análisis, así como una verificación de la calidad de los mismos.
- El Capítulo 4 comprende la fase de Preparación de los datos, en la cual se realiza la selección, limpieza y pre-procesamiento de los datos, así como también la caracterización de las señales e integración de los datos.

- El Capítulo 5 describe las fases Modelado y Evaluación en la cual se presenta el entrenamiento de algoritmos de clasificación y sus resultados, así como la evaluación del mejor algoritmo en términos de su rendimiento y de la calidad desde la perspectiva clínica.
- El Capítulo 6 presenta la fase de Despliegue, en la cual se describe la integración del modelo seleccionado con una aplicación móvil y con una pulsera.
- El Capítulo 7 presenta el Informe final, el cual contiene las contribuciones de esta disertación, así como también las conclusiones y el trabajo futuro.

Capítulo 2

2. Comprensión del Negocio

La fase inicial del proceso de minería de datos se enfoca en la comprensión de los objetivos desde un punto de vista del negocio, es decir, desde el punto de vista del dominio de conocimiento, entidad, organización u otro, en el cual se vayan a utilizar los resultados del proceso de minería de datos en un contexto de producción. Lo anterior es el insumo base para la identificación de los objetivos desde el punto de vista técnico, es decir, desde el punto de vista de la minería de datos; a su vez, permite definir el plan de trabajo necesario para alcanzar los objetivos. La Tabla 1 muestra las tareas ejecutadas y resultados alcanzados en la fase Comprensión del negocio.

Comprensión del negocio	Identificación del objetivo del negocio
	- Antecedentes
	- Objetivos del negocio
	- Criterios de éxito del negocio
	Evaluación de la situación
	- Inventario de recursos
	- Riesgos y contingencias
	Identificación del objetivo de la minería de datos
	- Objetivos de la minería de datos
	- Criterios de éxito de la minería de datos
	Realización del plan del proyecto
	- Plan del proyecto
	- Evaluación inicial de herramientas

Tabla 1. Tareas ejecutadas y resultados alcanzados en la fase Comprensión del negocio.

2.1. Identificación del objetivo del negocio

Como puede inferirse del planteamiento del problema presentado en la Sección 1.1, el negocio está enmarcado en la identificación y monitoreo de crisis epilépticas por medio de tecnologías alternativas a la EEG, que sean cómodas de vestir (*wearable*) y que permitan un seguimiento automático y ambulatorio de las crisis. Dado que el contexto de esta disertación es la investigación científica, el primero paso para identificar los objetivos del negocio y los criterios de éxito es establecer los antecedentes, es decir, el estado del arte (*background*) relacionado con la identificación de crisis epilépticas por medio de tecnologías alternativas a la EEG.

2.1.1. Antecedentes (*background*)

Los antecedentes fueron determinados a partir de una revisión de la literatura. La búsqueda de literatura se realizó en tres bases de datos electrónicas, a saber, Science Direct, IEEE Xplore Digital Library y PubMed. La estrategia de búsqueda fue construida utilizando tres grupos de palabras clave: detección de crisis epilépticas (epileptic seizure detection, seizure monitoring, convulsion detection, convulsion monitoring, seizure classification), método de monitoreo (non-EEG, wearable sensor, wearable device, biosensor) y sensores utilizados (accelerometry, electrodermal activity, photoplethysmography, temperature, heart rate). Las palabras se seleccionaron teniendo en cuenta los sinónimos de los términos médicos contenidos en el MeSH (Medical Subject Headings), el cual es el diccionario de vocabulario utilizado para indexar artículos para PubMed (posiblemente la base de datos de bibliografía médica más amplia que existe).

Se seleccionaron para revisión capítulos de libros, artículos, reportes técnicos, revisiones de literatura, entre otros, que cumplieron los siguientes criterios de inclusión: (1) estar escrito en inglés o español; (2) describir un sistema o algoritmo que identifique crisis epilépticas a través de señales capturadas por medio de dispositivos wearable o dispositivos portables; (3) reportar por lo menos una métrica de rendimiento del sistema o algoritmo para la detección de crisis epilépticas.

A continuación se presentan los detectores de crisis epilépticas propuestos en los artículos seleccionados para la revisión, los cuales están agrupados por el sensor utilizado. Posteriormente se presenta una tabla comparativa con las características más relevantes de los detectores y las brechas de investigación identificadas.

Detectores basados en acelerometría (ACC)

Dalton et al. [19] presentan el desarrollo de un sistema de monitoreo remoto para identificar crisis epilépticas con componente motor simple, utilizando sensores cinemáticos basados en acelerometría. Recolectaron datos de pacientes realizando diferentes actividades de la vida cotidiana para diferenciar las crisis epilépticas de actividades diarias. Durante el monitoreo, tomaron datos de v-EEG como referencia para la detección de crisis y desarrollaron un algoritmo de tipo *template matching* para diferenciar crisis de actividades diarias. Finalmente desarrollaron una *body sensor network* en la plataforma mercurio. La sensibilidad alcanzada por el algoritmo fue 0.91 y la especificidad fue 0.84.

Por otro lado, en el trabajo de Gubbi et al. [20] se presenta un dispositivo wearable con un sensor de acelerometría como una solución para la detección y diagnóstico de convulsiones psicogénicas no epilépticas (PNES, por sus siglas en inglés). Se desarrolló un algoritmo de detección de crisis epilépticas y un algoritmo de clasificación de PNES. Los algoritmos fueron probados con datos de pacientes epilépticos, obteniendo 100% de sensibilidad y un número bajo (no especificado) de falsas alarmas.

Por su parte, Velez et al. [21] muestra el desarrollo de un sistema detector y de seguimiento de crisis epilépticas, en el que un acelerómetro de pulsera se conecta vía Bluetooth a una tableta electrónica para transmitir información sobre los eventos epilépticos identificados a un portal web. La pulsera registra fecha, hora, audio, duración, frecuencia y amplitud de los eventos, con el fin de ayudar en el seguimiento de la epilepsia y evitar los errores humanos en los auto-informes. Los eventos capturados fueron verificados por medio de v-EEG. La pulsera capturó el 92.3% de las crisis epilépticas tónico clónicas generalizadas (GTCS, por sus siglas en inglés).

Schulc et al. [22] presentan el diseño y una validación exploratoria de un sistema de detección de GTCS en tiempo real basado en acelerometría. Los sensores de ACC registraron un aumento de actividad en los tiempos de ocurrencia de crisis epilépticas, actividad que difería claramente de los movimientos diarios. Los

resultados alcanzados fueron: sensibilidad 100%, especificidad $\geq 88\%$ y valor predictivo positivo $\geq 75\%$. El estudio resalta que la validez del dispositivo de detección debe evaluarse con más datos para poder confirmar los resultados positivos y desarrollar un algoritmo de corte para la detección automática de crisis.

Detectores basados en ritmo cardiaco

De Cooman et al. [23] describen un método que clasifica diferentes características de frecuencia cardiaca para detectar crisis de pacientes con epilepsia del lóbulo temporal (TLE, por sus siglas en inglés). Utilizan un clasificador de máquina de vectores de soporte. El algoritmo fue probado en un dataset de más de 918 horas de datos procedentes de 17 pacientes, el cual contiene 127 crisis parciales complejas (CPS, por sus siglas en inglés) y crisis parciales generalizadas (GPS, por sus siglas en inglés). El algoritmo fue capaz de detectar 81.89% de las crisis, con un promedio de 1.97 falsas alarmas por hora.

El trabajo presentado por Varon et al. [24] estudia las señales de electrocardiografía (ECG) registradas en pacientes con epilepsia, cuyas edades oscilaron entre 3 y 48 años. Únicamente consideraron crisis focales y generalizadas. El estudio se hizo mediante el promediado de señal rectificadas de fase, que es una técnica que encuentra cuasi-periodicidades en señales ruidosas y no estacionarias. Obtuvo un valor predictivo positivo (PPV, por sus siglas en inglés) del 86.21% con una sensibilidad del 100% para las crisis focales, y un PPV del 84.3% con una sensibilidad del 93.1% para las crisis generalizadas.

Por su parte, Van Elmpt et al. [25] presentan un modelo para la detección y caracterización automática de los cambios en la frecuencia cardíaca durante las crisis epilépticas en epilepsia severa. En este trabajo se controlaron los cambios en la frecuencia cardíaca en 10 pacientes con 104 crisis, en su mayoría tónicas y mioclónicas, para evaluar la detección de crisis basadas en la frecuencia cardíaca. El monitoreo de v-EEG sirvió como referencia. Dos algoritmos fueron desarrollados, primero se utilizó un algoritmo de ajuste de curva para caracterizar los patrones de frecuencia cardíaca. Un segundo algoritmo basado en un filtro de mediana móvil, permitía la detección automática del inicio del cambio de la frecuencia cardíaca. Se encontraron cambios en la frecuencia cardíaca en 8 de los 10 pacientes y en 50 de 104 crisis. En dos de cada tres pacientes con más de 10 crisis, se obtuvo un PPV de al menos el 50% y una sensibilidad superior al 90%.

Detectores basados en actividad electrodermal (EDA)

Heldberg et al. [10] analizaron señales EDA y ACC de 8 pacientes incluyendo también crisis epilépticas sin actividad motora. Implementaron un enfoque de clasificación jerárquica para clasificar segmentos de 10 segundos de estas señales en crisis predominantemente motoras o crisis predominantemente no motoras. Utilizando un clasificador Random Forest obtuvieron una sensibilidad de 90.5%, especificidad de 93.3% y precisión de 5.6% para las crisis con actividad motora. Para las crisis sin actividad motora, la sensibilidad fue 85.3%, la especificidad fue de 96.8% y la precisión de 12.3%.

El trabajo realizado por Cogan et al. [26] describe una metodología y un sistema portátil para la detección de crisis epilépticas, el cual alerta a un cuidador y registra el momento de la crisis en un diario electrónico. El sistema está basado en EDA, ritmo cardíaco (HR, por sus siglas en inglés), oxigenación arterial (SpO₂), ACC, temperatura (TEMP) y está definido en tres etapas. La etapa I busca la actividad concurrente en el HR, la SpO₂ y la EDA. La etapa II busca un patrón específico creado por estas tres bioseñales. Para los pacientes cuyas crisis no pueden ser detectadas por la Etapa II, la Etapa III detecta las crisis usando un EEG de canal limitado con un máximo de tres electrodos. De diez pacientes, la Etapa I reconoció 11 crisis de siete pacientes, la Etapa II detectó diez crisis de seis pacientes y la Etapa III detectó todas las crisis de dos de tres pacientes.

Onorati et al. [27] recogieron eventos anotados de v-EEG, los cuales pertenecían a 69 pacientes de seis centros médicos. Utilizaron tres pulseras diferentes para registrar la EDA y la ACC. Con estas señales se entrenaron y probaron dos clasificadores de aprendizaje automático y un clasificador publicado basado en EDA y ACC. Los datos de pulseras se analizaron para estimar la duración del movimiento de las crisis y las respuestas autonómicas. Como resultado, obtuvo una mejora en la clasificación respecto al clasificador anteriormente publicado. El más eficiente produjo una sensibilidad de 94.55% y una tasa de falsos positivos de 0.2 eventos por día.

Poh et al. [28] desarrollaron un algoritmo para la detección automática de crisis GTCS basado en EDA y ACC. Estas señales fueron capturadas por un biosensor de pulsera construido por los autores. El problema de la detección de crisis fue planteado como una tarea de aprendizaje supervisado en la que el objetivo era clasificar segmentos de 10 segundos como crisis o no-crisis. El algoritmo utilizó 19

características extraídas de EDA y ACC, y una máquina de vectores de soporte para la clasificación. El rendimiento se evaluó utilizando un método de doble validación cruzada. El algoritmo detectó 15 (94%) de 16 GTCS con 130 falsas alarmas (0.74 por 24 horas).

Brechas de investigación identificadas

A continuación se presenta una tabla comparativa con datos complementarios de los trabajos incluidos en la revisión (Tabla 2).

Trabajo	Tipos de crisis consideradas	Presencia de movimiento en crisis	Sensores	Resultados	Integrado a sistema real
[19]	No especificado	Componente motor simple	ACC	Sensibilidad: 91% Especificidad: 84%	Sí
[20]	PNES	Crisis motoras	ACC	Sensibilidad: 100%	No
[21]	GTCS	Crisis motoras	ACC	Sensibilidad: 92.3%	Sí
[22]	GTCS	Crisis motoras	ACC	Sensibilidad: 100% Especificidad: ≥88%	No
[23]	CPS, GPS (únicamente de pacientes con TLE)	Crisis motoras y no motoras	ECG	Sensibilidad: 81.89% Falsas alarmas: 197 por hora	No
[24]	Focales y generalizadas	No especificado	ECG	<u>Focales:</u> Sensibilidad: 100% PPV: 86.21% <u>Generalizadas:</u> Sensibilidad: 93.1% PPV: 84.3%	No
[25]	Tónicas, mioclónicas	Crisis motoras	ECG	Sensibilidad >90%. PPV ≥ 50%	No

Trabajo	Tipos de crisis consideradas	Presencia de movimiento en crisis	Sensores	Resultados	Integrado a sistema real
[10]	GTCS, hipermotoras, motoras complejas, dialépticas, automotoras e hipomotoras	Crisis motoras y no motoras	EDA y ACC	<u>Motoras:</u> Sensibilidad: 90.5%. Precisión: 5.6% <u>No motoras:</u> Sensibilidad: 85.3% Precisión: 12.3%	No
[26]	GTCS, CPS	No especificado	EDA, HR, SpO2, ACC, TEMP	Sensibilidad > 60% (sin EEG). Sensibilidad > 90% (involucrando 3 electrodos de EEG)	No
[27]	GTCS	Crisis motoras	EDA, ACC	Sensibilidad: 94.55%	Sí
[28]	GTCS	Crisis motoras	EDA, ACC	Sensibilidad: 94%	No

Tabla 2. Detectores de crisis epilépticas que pueden ser utilizados con dispositivos wearable o dispositivos portables.

A partir de la revisión de la literatura se identifican las siguientes brechas de investigación:

1. La mayoría de trabajos sólo tienen en cuenta crisis epilépticas de tipo motor.
2. Aquellos trabajos que consideran crisis motoras y no motoras presentan bajos resultados en la detección en términos de falsas alarmas y precisión.
3. La mayoría de los trabajos presentan detectores o algoritmos aislados, es decir no integrados a ninguna plataforma o sistema para el monitoreo de crisis epilépticas.
4. Los algoritmos o sistemas desarrollados carecen de pruebas en entornos cotidianos al paciente.

Las brechas de investigación que abordó esta disertación son las número 1 y 2, como se evidencia en las Secciones 2.1.2. y 2.1.3.

2.1.2. Objetivo del negocio

El objetivo del negocio que se identificó es similar al objetivo general presentado en el Capítulo 1: Construir un detector de crisis epilépticas motoras y no motoras para dispositivos wearable basado en EDA, ACC y PPG.

2.1.3. Criterios de éxito del negocio

Para darse por cumplido el objetivo del negocio, el algoritmo de detección de crisis epilépticas debió cumplir con los siguientes criterios de éxito:

1. Considerar en la detección, tanto crisis epilépticas motoras como no motoras.
2. En la etapa de evaluación, identificar la mayor cantidad posible de crisis epilépticas, las cuales deben ser identificadas de una forma precisa.

2.2. Evaluación de la situación

Esta tarea tiene como fin el reconocimiento de los recursos, riesgos y contingencias que deben ser considerados para la determinación del objetivo de la minería de datos y el plan del proyecto.

2.2.1. Inventario de recursos

Recursos humanos

El principal recurso humano que estuvo disponible para la investigación fue el estudiante de maestría Edward Alexis Molina Rivera, Ing., quien estuvo a cargo de la ejecución de las tareas de minería de datos de cada fase de la metodología CRISP-DM. El estudiante contó con la supervisión de Diego Mauricio López

Gutiérrez, PhD., y con el apoyo del Grupo de Investigación en Ingeniería Telemática (GIT) de la Universidad del Cauca. Adicionalmente, desde el dominio de la salud, se contó con la asesoría de la médico especialista en Neuropediatría, María Eugenia Miño Arango, quien aportó sus conocimientos en el área de la Neurología y la Epilepsia.

Datos

Al inicio de la investigación no se contó con los datos necesarios para la implementación del detector de crisis epilépticas que cumpliera con el objetivo del negocio, tampoco se encontraron datos públicos para este fin. Como se describe en la Sección 3.2.1, en las etapas iniciales de la investigación se realizó una recolección de señales de pacientes epilépticos y se realizaron pruebas iniciales, sin embargo, no se contó con un número significativo de crisis epilépticas. Durante la estancia de investigación realizada por el estudiante de maestría en el Machine Learning and Data Analytics Lab de la Universidad Friedrich-Alexander-Universität Erlangen-Nürnberg (Erlangen, Alemania), se obtuvo el dataset utilizado en esta disertación, el cual será descrito a lo largo de las Secciones 3.3 y 3.4.

Recursos hardware

Durante el transcurso de la investigación se contó con los recursos hardware listados en la Tabla 3.

Recurso	Descripción
Computador personal	Dell Inspiron 14, procesador i5 de cuarta generación y 8GB de RAM
Servidor físico	Dell Power Edge T320
Servidor virtual	Máquina VMware, procesador Intel Xeon (8 núcleos), 16GB de RAM
Pulsera	E4 Wristband (Empatica Inc., Cambridge, MA, USA): Pulsera con certificación médica para la adquisición de señales fisiológicas. Sensores: EDA, ACC, PPG y TEMP.
Electroencefalógrafo	BW-II EEG (Neurovirtual, Fort Lauderdale, FL, USA): Dispositivo médico para la toma de señales de EEG. Compatible con el sistema de posicionamiento internacional de electrodos 10-20.
Smartphone	Sony Xperia XA

Tabla 3. Recursos hardware disponibles en el transcurso de la investigación.

2.2.2. Riesgos y contingencias

A continuación se presentan los riesgos que se identificaron al inicio de la investigación y las respectivas acciones de contingencia planificadas.

1. Riesgo de que los datos recolectados no cuenten con un número suficiente de crisis epilépticas: En caso de ser necesario, buscar datos a través de grupos de investigación que abordan temas de investigación similares.
2. Riesgo de que la capacidad de procesamiento de los recursos hardware no sea suficiente: Si el computador personal y el servidor físico no cuentan con los recursos hardware suficientes para realizar el proceso de análisis de datos, adicionar más núcleos de procesador y más RAM al servidor virtual.
3. Riesgo de pérdida de los datos de análisis: Realizar un almacenamiento redundante de los datos utilizados.

2.3. Identificación del objetivo de la minería de datos

Mientras que el objetivo del negocio presentado en la Sección 2.1.2. está descrito en términos del dominio, el objetivo de la minería de datos está descrito en términos técnicos, es decir, desde la perspectiva del análisis de datos.

2.3.1. Objetivo de la minería de datos

El tipo de problema de minería de datos abordado es clasificación, en específico, aprendizaje supervisado (*supervised learning*). Por lo anterior el objetivo de la minería de datos se definió como:

Clasificar segmentos de 10 segundos, los cuales incluyan señales EDA, ACC y PPG como crisis epiléptica (*seizure*) o no crisis epiléptica (*non-seizure*).

2.3.2. Criterios de éxito de la minería de datos

Desde la perspectiva de la minería de datos, el clasificador debió cumplir los siguientes criterios de éxito:

1. Incluir en los datos de entrenamiento instancias de crisis epilépticas de tipo motor, e instancias de crisis epilépticas de tipo no motor.
2. En la fase de evaluación, identificar crisis epilépticas con valores de sensibilidad y precisión mayores o iguales al 90%.

2.4. Realización del plan del proyecto

En esta sección se describe el plan del proyecto, elaborado con el fin de listar las fases y tareas necesarias para alcanzar el cumplimiento de los objetivos de la minería de datos, lo cual, a su vez llevará al cumplimiento del objetivo del negocio.

2.4.1. Plan del proyecto

La Tabla 4 muestra el plan general del proyecto. El detalle de la ejecución de cada tarea es descrito en su momento en los capítulos siguientes.

Fase / Tareas	Duración (meses)
Comprensión del negocio: <ul style="list-style-type: none">- Identificación del objetivo del negocio- Evaluación de la situación- Identificación del objetivo de la minería de datos- Realización del plan del proyecto	1
Comprensión de los datos: <ul style="list-style-type: none">- Definición del marco conceptual- Recolección de los datos iniciales- Descripción de los datos- Exploración de los datos- Verificación de la calidad de los datos	2

Fase / Tareas	Duración (meses)
Preparación de los datos: - Selección de los datos - Limpieza y pre-procesamiento de los datos - Construcción de los datos - Integración y formateo de los datos	3
Modelado: - Selección de los algoritmos de clasificación - Modelado y resultados	1
Evaluación: - Evaluación del modelo - Evaluación de la calidad del proceso y del modelo desde la perspectiva clínica	1
Despliegue: - Arquitectura de despliegue - Aplicación móvil para la detección automática de eventos epileptiformes	1

Tabla 4. Plan general del proyecto considerando fases y tareas de la metodología CRISP-DM.

2.4.2. Evaluación inicial de herramientas

Existen diferentes herramientas y lenguajes de programación que permiten realizar tareas de minería de datos, tales como pre-procesamiento de datos, entrenamiento y evaluación de algoritmos de clasificación, entre otras. Entre las herramientas y lenguajes de programación más utilizados en ámbitos académicos se encuentran MATLAB [29], R [30], WEKA [31] y Scikit-learn (Python) [32].

Al momento de evaluar las posibles herramientas se tuvieron en cuenta criterios como las restricciones de uso (uso libre, software de pago), la cantidad y calidad de documentación disponible, la curva de aprendizaje, su uso por la comunidad científica, existencia de módulos o paquetes adicionales que se puedan integrar a las herramientas para resolver necesidades puntuales y su nivel de soporte a las demás fases de la metodología CRISP-DM.

Teniendo en cuenta lo anterior, se seleccionaron como las herramientas principales para realizar tareas de minería de datos el lenguaje de programación Python y los módulos Scikit-learn y SciPy [33]. Scikit-learn es un módulo de código abierto de

amplio uso para Python desarrollado para realizar tareas de aprendizaje automático. SciPy también es un módulo de código abierto para Python, el cual permite, entre otros, cálculos matemáticos complejos y el procesamiento de señales. Además, cuando se vio la necesidad, se seleccionaron librerías adicionales para Python, las cuales serán mencionadas en su momento.

Capítulo 3

3. Comprensión de los datos

Una vez claro el objetivo del negocio y el objetivo de la minería de datos, se procede a recolectar y comprender los datos necesarios para el análisis. En este sentido, la fase de Comprensión de los datos inicia con la definición del marco conceptual, la recolección inicial de datos y continúa con tareas que permiten describir la estructura de los datos, identificar problemas de calidad y descubrir primeras impresiones de los datos. La Tabla 5 muestra las tareas ejecutadas durante esta fase.

Comprensión de los datos	Definición del marco conceptual
	- Marco conceptual
	Recolección de los datos iniciales
	- Reporte de recolección de los datos iniciales
	Descripción de los datos
	- Reporte de descripción de los datos
Exploración de los datos	
- Reporte de exploración de los datos	
Verificación de la calidad de los datos	
- Reporte de calidad de los datos	

Tabla 5. Tareas ejecutadas y resultados alcanzados en la fase Comprensión de los datos.

3.1. Definición del marco conceptual

Para comprender de una forma adecuada el proceso de recolección de datos iniciales, así como la descripción de los datos a ser utilizados, es necesario presentar la definición de diferentes conceptos teóricos, los cuales son relevantes para esta disertación.

3.1.1. Marco conceptual

A continuación se presentan los conceptos más relevantes relacionados con la epilepsia, la clasificación de crisis epilépticas y las señales fisiológicas utilizadas para la identificación de crisis epilépticas.

Epilepsia y crisis epiléptica

La epilepsia es un término colectivo para diferentes enfermedades del sistema nervioso en las que un trastorno en la actividad de las células nerviosas produce crisis epilépticas por una descarga síncrona anormal y repentina de una población de neuronas. En todo el mundo, entre el 5 y el 10% de la población mundial tiene al menos una crisis durante su vida. Alrededor del 1% sufre de epilepsia, que corresponde a unos 50 millones de personas en todo el mundo. Por lo tanto, la epilepsia es una de las enfermedades más prevalentes del sistema nervioso. La epilepsia afecta a personas de todas las edades, pero los niños y las personas mayores tienen mayor incidencia [34].

Dependiendo del área cerebral involucrada, una crisis epiléptica se puede expresar de maneras muy diferentes. El espectro de las crisis va desde extrañas percepciones sensoriales hasta espasmos musculares y pérdida de conciencia. Las causas de la epilepsia también son bastante variadas. Los recién nacidos y los niños a menudo sufren de epilepsia debido a traumatismos de nacimiento o desarrollos anormales relacionados con el cerebro. Heridas en la cabeza, los accidentes cerebrovasculares, los tumores o los traumas también pueden desencadenar crisis epilépticas en adultos [35].

Fases de las crisis epilépticas

La secuencia de una crisis epiléptica a menudo se puede asignar a tres fases distintas. La primera etapa, denominada la fase pre-ictal, describe el tiempo antes de que ocurra una crisis epiléptica. En esta fase, algunas personas notan cambios en el comportamiento, el sueño, el habla y trastornos cognitivos [36]. Sin embargo, no todas las personas con epilepsia experimentan cambios en esta fase. La fase ictal representa el periodo de ocurrencia de una crisis. Durante esta fase aparecen los síntomas (semiología) que se describirán más adelante. Para los neurólogos estos síntomas son los más importantes ya que la observación del comportamiento y las diferentes señales eléctricas del cuerpo le permiten al médico diagnosticar al paciente. La última fase se denomina post-ictal, en la cual se pueden observar anomalías clínicas transitorias del sistema nervioso central [37].

Tipos de crisis epilépticas según su semiología

Existen diferentes clasificaciones de crisis epilépticas. A continuación se describen los tipos de crisis epilépticas según su semiología de acuerdo a la propuesta de H. Lüders et al. [38]. Ya que la clasificación es extensa, en esta sección sólo se presentan los tipos de semiología considerados en esta disertación.

Las perturbaciones sensoriales sin signos objetivos se denominan aura. Se pueden distinguir en diferentes subcategorías, dependiendo de los síntomas subjetivos. Por ejemplo, un hormigueo o la sensación de sordera pertenece a las auras somato-sensoriales, y una alucinación visual se atribuye a la categoría de auras visuales. La duración suele ser de unos segundos a unos minutos.

Las crisis con actividad motora predominante se identifican por dos subgrupos, crisis motora simple y crisis motora compleja. Las crisis motoras simples consisten en movimientos antinaturales simples. La articulación de estos movimientos simples se realiza típicamente en una dirección. Por el contrario, las crisis motoras complejas se caracterizan por movimientos complejos pero naturales que no son apropiados en la situación [39]. El espasmo epiléptico y la crisis tónico-clónica pertenecen al grupo motor simple. Las crisis tónicas desencadenan la contracción continua de los músculos. De este modo, los grupos musculares involucrados pueden afectar un lado del cuerpo o ambos.

Una crisis hipermotora está marcada por movimientos rápidos y violentos. La ocurrencia es principalmente durante la fase de sueño [39]. Las crisis automotoras

se caracterizan por automatismos de movimiento de las extremidades distales y la boca. A menudo se asocian con estados de conciencia alterados [39].

Un caso especial es la crisis hipomotora. Se identifica por una actividad motora reducida o completamente inexistente. Se puede definir como la contraria a una crisis motora compleja [38].

Señales fisiológicas

A continuación se describen brevemente las señales fisiológicas consideradas en esta disertación y la forma en la que cada una se relaciona con la identificación de crisis epilépticas.

La EDA refleja las características eléctricas de la piel en respuesta a la secreción de sudor. Se mide mediante dos electrodos cercanos ubicados en contacto con la piel. Durante la activación fisiológica o el estrés emocional en una crisis epiléptica, el sistema nervioso simpático controla las glándulas sudoríparas, lo que aumenta la producción de sudor. [40], [41]

La ACC se mide con un sensor de aceleración 3D, el cual mide la aceleración que actúa en el dispositivo. Si el sensor no se mueve, solo la gravedad tiene un efecto en el dispositivo. Muchas crisis epilépticas están acompañadas de movimiento, por lo tanto, la ACC puede ayudar a distinguir entre una crisis epiléptica de tipo motor, y una de tipo no motor. [41]

La PPG es una técnica óptica simple y de bajo costo que se puede utilizar para detectar cambios en el volumen sanguíneo en el lecho micro-vascular de un tejido. Por lo general se usa de forma no invasiva realizando mediciones en la superficie de la piel y aporta información valiosa sobre el sistema cardiovascular. [42] La frecuencia cardíaca a menudo aumenta durante una crisis epiléptica [41].

3.2. Recolección de los datos iniciales

3.2.1. Reporte de recolección de los datos iniciales

La recolección de los datos iniciales se realizó entre marzo y agosto del año 2018, es decir, durante el transcurso de seis meses, en la IPS (Instituto Prestador de Servicios de Salud) AXON PED S.A.S. de la ciudad de Popayán (Cauca, Colombia).

Esta IPS realiza tomas de exámenes de EEG a pacientes pediátricos en busca de eventos anormales que estén relacionados con la epilepsia.

La recolección se realizó sobre pacientes pediátricos que estaban programados para ser atendidos por la IPS. Antes de la participación de los pacientes, sus padres o acudientes firmaron un consentimiento informado (Anexo A) el cual contiene la aprobación del Comité de Ética de la Universidad del Cauca (Popayán, Colombia), siguiendo la Declaración de Helsinki [43]. Además, la recolección de las señales EDA, ACC, PPG, TEMP y EEG estuvo dirigida por un protocolo de recolección (Anexo B). Las primeras cuatro señales fueron recolectadas por el dispositivo E4 wristband y la última, por el dispositivo BW-II EEG (Sección 2.2.1.). Los dos dispositivos estaban sincronizados con la hora del computador utilizado para almacenar las señales.

Se recolectaron alrededor de 200 exámenes EEG de 30 minutos de duración (Figura 3). Cada examen fue inspeccionado por la Neuropediatra María E. Miño en busca de eventos epilépticos en la señal EEG. Cuando estos eventos eran encontrados, se realizaba la anotación del tiempo de inicio y fin de cada uno. Con la señal de referencia (EEG) fue posible identificar los eventos en las señales de la pulsera E4.



Figura 3. Recolección inicial de señales de EDA, ACC, PPG, TEMP y EEG, pertenecientes a un paciente pediátrico.

A pesar del esfuerzo invertido en la recolección de datos, sólo se identificaron 24 crisis epilépticas pertenecientes a cuatro pacientes, con una duración muy corta por

crisis, lo que representó aproximadamente un total de 70 segundos de crisis epilépticas. Además, ninguna de las crisis era de tipo motor, con lo cual no se podrían cumplir los criterios de éxito del objetivo de la minería de datos (Sección 2.3.2).

Teniendo en cuenta que la obtención de un dataset de calidad y con suficientes datos representativos es de vital importancia para cualquier investigación enmarcada en la minería de datos, se decidió no utilizar el dataset recolectado inicialmente. Por lo tanto, se identificó la necesidad de contar con un dataset ya recolectado por grupos de investigación externos al GIT.

3.3. Descripción de los datos

Con la motivación de encontrar un dataset de señales apropiado y de recibir asesoría por parte de investigadores expertos, el estudiante de maestría realizó una estancia de investigación en el Machine Learning and Data Analytics Lab (MaD Lab) de la Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), ubicada en la ciudad de Erlangen (Alemania). Investigadores del MaD Lab han realizado trabajos previos enmarcados en la detección de crisis epilépticas utilizando bioseñales y *Machine Learning* [10], [41], [44], lo que permitió que el laboratorio fuera apropiado para la estancia de investigación.

Durante la estancia de investigación se obtuvo un dataset de señales de 11 pacientes epilépticos, recolectado con el dispositivo E3 wristband (equipado con sensores ACC, EDA, PPG y TEMP). Adicionalmente, se obtuvo un registro de anotaciones con el tipo de crisis (semiología) y los tiempos de inicio y fin de las crisis epilépticas de cada paciente. Este dataset fue recolectado por trabajos previos en el MaD Lab. Por tratarse de un dataset con una buena cantidad de crisis epilépticas identificadas, éste fue utilizado para el análisis de datos. La descripción de este dataset se presenta en detalle en las secciones 3.3.1 y 3.4.1.

3.3.1. Reporte de descripción de los datos

El dataset utilizado fue recolectado en trabajos previos del MaD Lab por Heldberg et al. [44] y Knoth [41]. Contiene datos de 11 pacientes epilépticos, los cuales fueron

recolectados en el Centro de Epilepsia del Departamento de Neurología en el Hospital Universitario de Erlangen (Erlangen, Alemania). La comisión de Ética de la Facultad de Medicina de la Friedrich-Alexander-Universität Erlangen-Nürnberg verificó y aprobó los trabajos responsables de la recolección de los datos.

Los datos fueron recolectados utilizando el dispositivo E3 wristband (Empatica Inc., Cambridge, MA, USA) (Figura 4), el cual es la versión anterior de la pulsera utilizada en la recolección inicial de datos (Sección 3.2.1.) e incorpora los mismos sensores. La pulsera E3 wristband es un dispositivo comercial diseñado para investigadores, la cual incluye varios sensores para medir señales fisiológicas de un paciente. Está equipada con un acelerómetro de tres ejes para medir la fuerza gravitacional (g) en cada dirección, un sensor de EDA para medir la conductividad eléctrica de la piel, un sensor de PPG, el cual permite medir el pulso del volumen sanguíneo (BVP), y un sensor de temperatura [45].



Figura 4. Pulsera E3 wristband.
Fuente: Garbarino et al. [45]

La Tabla 6 muestra la frecuencia de muestreo y el rango de operación de los sensores ACC, EDA, PPG y TEMP, los cuales están incorporados en la pulsera E3 wristband.

Sensor	Frecuencia de muestreo (Hz)	Rango de operación
ACC	32	± 2 g
EDA	4	0.01–100 μ S
PPG (BVP)	64	-500–500 (variación de la intensidad de la luz)

Sensor	Frecuencia de muestreo (Hz)	Rango de operación
TEMP	4	-40–115 °C

Tabla 6. Frecuencia de muestreo y rango de operación de los sensores de la pulsera E3.

La recolección de datos se realizó con pacientes que acudían a estudios v-EEG regulares de larga duración en una unidad de monitoreo v-EEG. A cada paciente se le colocó una pulsera en cada brazo para la recolección de señales. Las señales EEG fueron recolectadas siguiendo el sistema internacional de posicionamiento de electrodos 10-20. Además, se recolectaron señales de Electrocardiograma (ECG) por medio de dos electrodos ubicados en el pecho y un electrodo de referencia ubicado en el cuero cabelludo. Los registros de v-EEG y ECG fueron tomados como señales de referencia para la detección de crisis epilépticas por parte de un médico especialista en epilepsia.

Todos los cables de los electrodos fueron conectados a una caja amplificadora, la cual estaba localizada en la espalda del paciente, y un cable conectaba la caja amplificadora a la unidad de monitoreo. De esta manera, el paciente se podía mover libremente en el cuarto de monitoreo para realizar algunas actividades. Aunque los datos recolectados no reflejan las condiciones reales de la vida cotidiana, están recolectados en un ambiente más realista. El tiempo de recolección de señales de un paciente varió entre uno a nueve días.

Los pacientes fueron video monitoreados por un asistente médico durante todo el tiempo de la recolección de señales. Cada vez que se observaba un evento anormal en un paciente, se realizaba una anotación del tiempo de inicio y fin del evento. Después de 24 horas de recolección de datos, un especialista en epilepsia revisaba en los datos de v-EEG y ECG, todos los eventos detectados. Cuando un evento era confirmado como crisis epiléptica por el especialista, se registraba el tiempo de inicio y fin del mismo, y se clasificaba la crisis de acuerdo con la semiología de crisis epilépticas propuesta por H. Lüders et al. [38] (Sección 3.1.1), la cual está basada en los síntomas de las crisis. La extracción de crisis epilépticas en las señales de las pulseras E3 requirió una sincronización de tiempo entre las pulseras y el dispositivo EEG. Esta sincronización fue realizada creando un artefacto específico en la señal de aceleración de las dos pulseras y al mismo momento, estableciendo un marcador de tiempo en el entorno software del EEG.

Once pacientes epilépticos fueron monitoreados, siete hombres y cuatro mujeres, de edades entre los 18 y los 57 años. En total, se recolectaron alrededor de 1000

horas de datos. Los once pacientes experimentaron 46 crisis epilépticas, y el número de crisis por pacientes varió de 1 a 12.

El tamaño del dataset (señales del dispositivo E3 wristband) es de 2.59 GB y está dividido en 38 carpetas, cada una con registros de señales de aproximadamente 24 horas. Las carpetas están nombradas de acuerdo al siguiente formato:

YYYY_MM_DD_d[deviceID]_p[patientID]_leftright

En seguida se presenta el significado de cada elemento del formato.

YYYY: Año del registro de señales.

MM: Mes del registro de señales.

DD: Día del registro de señales.

deviceID: Identificador de la pulsera utilizada para la recolección.

patientID: Identificador del paciente anónimo a quien se recolectaron las señales.

leftright: “l” si el registro de señales se realizó en el brazo izquierdo, “r” si se realizó en el brazo derecho.

La Tabla 7 muestra los archivos de señales contenidos en cada carpeta del dataset. Todas las señales provenientes de los sensores de la pulsera E3 están almacenadas en archivos de valores separados por comas (CSV, por sus siglas en inglés). La primera fila de cada archivo de señales representa el tiempo inicial del registro expresado como una marca de tiempo Unix en Tiempo Universal Coordinado (UTC, por sus siglas en inglés). La segunda fila representa la frecuencia de muestreo del sensor expresada en Hz. Desde la tercera fila en adelante los archivos contienen los valores registrados de cada sensor según su frecuencia de muestreo y su rango de operación (Tabla 6).

Nombre de archivo	Descripción
ACC.csv	Datos del sensor ACC de 3 ejes. Los datos de los ejes x, y, z, están en la primera, segunda y tercera columna, respectivamente. Unidad expresada como 1/64 g.
BVP.csv	Datos del sensor PPG, ubicados en una sola columna. Unidad: variación de la intensidad de la luz.
EDA.csv	Datos del sensor EDA, ubicados en una sola columna. Unidad: Micro siemens (μ S).

Nombre de archivo	Descripción
IBI.csv	Intervalo entre latidos cardíacos individuales (IBI, por sus siglas en inglés) extraídos de la señal BVP, ubicados en una sola columna. Estos datos son registrados por el dispositivo sólo cuando es posible, por lo cual no tienen una frecuencia de muestreo especificada.
TEMP.csv	Datos del sensor de temperatura. Unidad: Grados Celsius (°C).
info.txt	Archivo con una corta descripción de la estructura de datos de cada archivo de señales.

Tabla 7. Archivos contenidos en cada carpeta del dataset de señales de la pulsera E3 wristband.

3.4. Exploración de los datos

A continuación se presenta una exploración detallada del dataset descrito en la Sección 3.3.1, en términos de las características de los pacientes, tipos de crisis identificadas (semiología), y número de crisis por paciente y por semiología.

3.4.1. Reporte de exploración de los datos

Las principales características de los once pacientes contenidos en el dataset se muestran en la Tabla 8. Vale la pena aclarar que en esta tabla no se lista el paciente con ID 8 ya que los datos de este paciente no estaban presentes en el dataset obtenido del MaD Lab. Los números de ID no fueron reasignados a partir del 8 en adelante para evitar confusiones en los registros de crisis epilépticas.

ID Paciente	Edad	Género	Brazo dominante	Número de crisis epilépticas
1	20	M	I	12
2	25	F	D	1
3	26	F	D	5
4	28	M	D	2
5	42	M	D	4
6	57	F	D	6
7	20	F	D	3
9	18	M	I	5
10	24	M	I	2

ID Paciente	Edad	Género	Brazo dominante	Número de crisis epilépticas
11	36	M	I	3
12	21	M	NA	3
TOTAL				46

Tabla 8. Principales características de los pacientes epilépticos contenidos en el dataset. Brazo dominante: brazo que presenta mayores amplitudes de la señal EDA. I: Izquierdo; D: Derecho; NA: No aplica.

La Tabla 9 muestra el número de crisis epilépticas identificadas por paciente en cada semiología. Por otro lado, la Tabla 10 presenta el número total de crisis epilépticas identificadas por semiología así como la agrupación de crisis según su presencia de movimiento, agrupación similar a la reportada por Heldberg et al. [10].

Semiología ID paciente	TO	HIPE	HIPO/A	A	AG	AUR	MCG	MC
1	12	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0
3	0	0	5	0	0	0	0	0
4	0	0	0	2	0	0	0	0
5	0	0	0	3	1	0	0	0
6	0	0	0	5	0	1	0	0
7	0	0	0	0	0	1	1	1
9	0	0	0	0	0	0	0	5
10	0	0	0	0	0	0	0	2
11	0	0	0	0	0	0	0	3
12	0	3	0	0	0	0	0	0

Tabla 9. Número de crisis epilépticas identificadas por cada paciente, discriminadas por semiología. TO: Tónica; HIPE: Hipermotora; HIPO/A: Hipomotora/Automotora; A: Automotora; AG: Automotora Generalizada; AUR: Aura; MCG: Motora Compleja Generalizada; MC: Motora Compleja.

Semiología	Tipo	Presencia de movimiento	Número de crisis
Tónica	Focal	Predominantemente motora	12
Hipermotora	Focal	Predominantemente motora	4
Hipomotora/Automotora	Focal	Predominantemente no motora	5
Automotora	Focal	Predominantemente no motora	10

Semiología	Tipo	Presencia de movimiento	Número de crisis
Automotora Generalizada	Generalizada	Predominantemente no motora	1
Aura	Focal	Predominantemente motora	2
Motora Compleja Generalizada	Generalizada	Predominantemente motora	1
Motora Compleja	Focal	Predominantemente motora	11
TOTAL			46

Tabla 10. Número total de crisis epilépticas por cada semiología.

3.5. Verificación de la calidad de los datos

Por medio de esta tarea se verificó la calidad de las señales contenidas en el dataset. En este sentido, esta tarea permitió responder preguntas como: ¿los datos están completos?, ¿los datos cubren todos los casos requeridos?, ¿los datos son correctos o contienen errores? Una vez realizada esta tarea, se identificaron las acciones necesarias para asegurar la calidad de los datos. De esta manera, los datos quedaron listos para ser preparados (pre-procesados), procedimiento descrito en la siguiente fase del proceso de minería de datos (Capítulo 4).

3.5.1. Reporte de la calidad de los datos

La verificación de la calidad de los datos recolectados se realizó considerando la completitud del dataset y cada uno de los problemas de calidad de datos identificados y descritos por Corrales et al. [46].

Con el objetivo de cumplir los criterios de éxito de la minería de datos, se requiere garantizar la completitud del dataset. De acuerdo a los criterios de éxito, el dataset debe incluir datos de crisis epilépticas motoras y de crisis epilépticas no motoras. Teniendo en cuenta que los datos recolectados incluyen crisis epilépticas motoras y no motoras, la completitud del dataset está asegurada (Tabla 10).

A continuación se describe brevemente cada uno de los problemas de calidad de datos identificados por Corrales et al. [46], así como los hallazgos encontrados en

el dataset y las acciones realizadas para asegurar la calidad de los datos (en los casos donde sea necesario).

Valores perdidos (*missing values*)

Se refiere al caso en el que una variable o atributo no contiene ningún valor. Este problema ocurre cuando la fuente de los datos presenta algún error o inconsistencia.

Las señales del dataset provienen de sensores que almacenan datos a una frecuencia de muestreo fija, la cual varía según el sensor. Esto permite que, en condiciones de un funcionamiento normal, no haya valores perdidos. No se identificaron comportamientos en las señales que pudieran dar indicios de fallas en los sensores, ni tampoco se identificaron valores perdidos en las señales.

Valores atípicos (*outliers*) y ruido (*noise*)

Un valor puede ser considerado como atípico cuando se desvía de forma considerable de otros valores, es decir, cuando parece ser inconsistente respecto a los demás. Por otro lado, el ruido representa datos irrelevantes o sin sentido.

El dataset puede contener ruido causado por los movimientos corporales de los pacientes. Estos artefactos de movimiento pueden afectar el análisis de señales como la EDA [44]. La Sección 4.2 describe las acciones realizadas para reducir artefactos de movimiento en la señal EDA.

Las señales pueden presentar valores atípicos debido a fallas en los sensores. Una falla podría hacer que los sensores registren valores fuera de su rango de operación. Sin embargo, ninguno de los valores provenientes de los sensores está por fuera del rango de operación (Tabla 6).

Alta dimensionalidad (*high dimensionality*)

La alta dimensionalidad se evidencia cuando el dataset contiene un número elevado de características.

El problema de la alta dimensionalidad no aplica para este dataset, dado que las señales no describen en sí mismas características ni instancias, sino que, por el contrario, necesitan de un procedimiento de caracterización para describirlas. Este procedimiento se describe en la Sección 4.3.1, en la cual no se identificó un problema de alta dimensionalidad después de la caracterización de señales.

Redundancia (*redundancy*)

La redundancia hace referencia a la existencia de instancias duplicadas en el dataset.

Existe un problema de redundancia en los datos a nivel de las señales recolectadas. Como se describió en la Sección 3.3.1, las señales EDA, ACC, PPG y TEMP fueron obtenidas de forma simultánea utilizando una pulsera E3 en cada brazo de los pacientes, por lo tanto, la información es redundante. Este problema fue resuelto seleccionando sólo las señales del brazo que mostrara la mayor amplitud en la señal EDA (Sección 4.1.1).

Capítulo 4

4.Preparación de los datos

La fase de preparación de los datos cubre todas las tareas necesarias para construir el dataset que será utilizado directamente para el entrenamiento de algoritmos de clasificación en la fase Modelado (Capítulo 5). Las tareas relacionadas con la preparación de los datos pueden ser ejecutadas de forma iterativa y en diferente orden, según las necesidades que se identifiquen en los datos. Estas tareas incluyen el pre-procesamiento de los datos, caracterización de señales, selección de características, y selección de un subconjunto representativo de los datos, las cuales serán documentadas a lo largo de este capítulo. Al final de esta fase se obtiene un segundo dataset, denominado dataset preparado.

Cabe mencionar que la fase Preparación de los datos es la que requiere más tiempo y esfuerzo en proyectos de minería de datos, se estima con frecuencia que puede requerir entre el 50% y el 70% del tiempo y esfuerzo total [18].

Esta fase está estrechamente relacionada con la fase de Modelado, ya que los datos deben ser preparados de una manera apropiada para las técnicas seleccionadas. Teniendo en cuenta que el tipo de problema de minería de datos que se está abordando es de clasificación, los datos fueron preparados de una forma apropiada para ser utilizados por los algoritmos de clasificación seleccionados (Tabla 18) en la fase de Modelado. La Tabla 11 muestra las tareas ejecutadas y resultados alcanzados durante esta fase.

Preparación de los datos	Selección de los datos
	- Datos seleccionados
	- Características seleccionadas

Limpieza y pre-procesamiento de los datos
Construcción de los datos
- Caracterización de señales
Integración y formateo de los datos
- Descripción del dataset preparado

Tabla 11. Tareas ejecutadas y resultados alcanzados en la fase Preparación de los datos.

4.1. Selección de los datos

Esta tarea tiene como fin seleccionar los datos a ser incluidos en el dataset preparado, y por lo tanto, en el análisis. La selección se realiza a partir del dataset descrito en la fase Comprensión de los datos, específicamente en las secciones 3.3.1 y 3.4.1. Comprende la selección de datos de señales, así como la selección de datos a nivel de variables o atributos (selección de características).

4.1.1. Datos seleccionados

Knuth [41] identificó que, cuando se recolectan señales EDA, cada paciente tiene un brazo en el que la amplitud promedio de la señal EDA es mucho mayor; a este brazo se le denomina brazo dominante. Por lo anterior, se seleccionaron para análisis sólo las señales procedentes de la pulsera E3 que estaban ubicadas en el brazo dominante de cada paciente (Tabla 8), excepto para el paciente con ID 12, del cual sólo se contaba con datos del brazo derecho. Después de seleccionar las carpetas de señales del dataset que pertenecen al brazo dominante de los pacientes, según lo registrado en la Tabla 8, el número de carpetas de señales que serán incluidas en el análisis se redujo de 38 (Sección 3.3.1), a 19 (alrededor de 500 horas de datos).

En la Tabla 10 se observa que de las 46 crisis epilépticas identificadas en el dataset, 44 de ellas son de tipo Focal, mientras que sólo 2 son de tipo Generalizada. Teniendo en cuenta que los tipos de crisis Focal y Generalizada tienen diferencias significativas desde el punto de vista médico, para el análisis de datos se excluyeron las dos crisis de tipo Generalizada. De esta manera, la Tabla 12 muestra la

distribución de crisis epilépticas seleccionadas por paciente, y la Tabla 13 muestra la distribución de crisis seleccionadas por semiología.

Semiología ID paciente	TO	HIPE	HIPO/A	A	AUR	MC
1	12	0	0	0	0	0
2	0	1	0	0	0	0
3	0	0	5	0	0	0
4	0	0	0	2	0	0
5	0	0	0	3	0	0
6	0	0	0	5	1	0
7	0	0	0	0	1	1
9	0	0	0	0	0	5
10	0	0	0	0	0	2
11	0	0	0	0	0	3
12	0	3	0	0	0	0

Tabla 12. Número de crisis epilépticas incluidas en el dataset preparado, discriminadas por semiología. TO: Tónica; HIPE: Hipermotora; HIPO/A: Hipomotora/Automotora; A: Automotora; AUR: Aura; MC: Motora Compleja.

Semiología	Tipo	Presencia de movimiento	Número de crisis
Tónica	Focal	Predominantemente motora	12
Hipermotora	Focal	Predominantemente motora	4
Hipomotora/Automotora	Focal	Predominantemente no motora	5
Automotora	Focal	Predominantemente no motora	10
Aura	Focal	Predominantemente motora	2
Motora Compleja	Focal	Predominantemente motora	11
TOTAL			44

Tabla 13. Número de crisis epilépticas incluidas en el dataset preparado, agrupadas por semiología.

Las señales recolectadas de los pacientes, las cuales se encuentran en cada una de las 19 carpetas seleccionadas son ACC, BVP (PPG), EDA, IBI y TEMP. De éstas, se seleccionaron las señales ACC, BVP y EDA para el análisis de datos. La señal IBI es extraída de forma automática por la pulsera E3 a partir de los datos de la

señal BVP, sin embargo, su extracción depende de la no presencia de artefactos de movimiento, por lo que a diferencia de las otras señales, la IBI no es registrada con una frecuencia de muestreo determinada [47]. Incluir esta señal en el dataset preparado introduciría el problema de calidad denominado valores perdidos (Sección 3.5.1), por lo que las señales IBI no fueron consideradas para el análisis de datos. La señal TEMP tampoco fue considerada para el análisis porque la temperatura de la piel no tiene una incidencia significativa en la detección de crisis epilépticas [41], incluirla introduciría en el dataset el problema de calidad denominado ruido (Sección 3.5.1).

4.1.2. Características seleccionadas

En la sección anterior se describieron las carpetas del dataset, los tipos de crisis epilépticas y las señales seleccionadas para su inclusión en el dataset preparado. El siguiente paso fue determinar de qué forma se representaría la información de las señales seleccionadas para el análisis. Teniendo en cuenta que las señales seleccionadas (ACC, BVP y EDA) no pueden ser utilizadas de forma directa en los algoritmos de clasificación seleccionados en la fase de Modelado, ya que no están expresadas como observaciones o instancias, es necesario realizar un proceso de caracterización de señales, el cual permitirá representar y/o describir matemáticamente segmentos de las señales.

La caracterización de una señal se realiza usualmente dividiendo la señal en segmentos (segmentación) y posteriormente, realizando cálculos matemáticos sobre cada segmento con el propósito de representarlo y/o describirlo. Cada variable calculada con el fin de representar un segmento es denominada característica (*feature*). Con frecuencia se extraen características en el dominio del tiempo, frecuencia o tiempo—frecuencia de las señales, utilizando conceptos estadísticos como la media, mediana y desviación estándar; o conceptos relacionados con la energía como por ejemplo la energía espectral.

Al caracterizar las señales es necesario definir una estrategia que permita seleccionar las características más relevantes. Las características a extraer pueden ser seleccionadas gracias a la recomendación de un experto, extrayendo a juicio del investigador diferentes características y seleccionando las mejores a través de la experimentación, identificando las más relevantes en la literatura, etc.

En esta disertación se realizó la caracterización de las señales con base en el trabajo de Knoth [41], quien realizó un análisis estadístico de características para la detección de crisis epilépticas como parte de su tesis de Maestría en Ingeniería Médica desarrollada en el MaD Lab. Este trabajo fue identificado por el estudiante de maestría durante la estancia de investigación en el laboratorio.

Knoth [41] analizó estadísticamente 52 características de segmentos pertenecientes a las fases pre-normal, pre-ictal, ictal, post-ictal y post-normal de 50 crisis epilépticas. Las fases pre-normal y post-normal representan el estado normal de las señales, es decir, señales con ausencia de crisis epilépticas. Estas dos fases tienen una duración de 30 minutos, mientras que las fases pre-ictal y post-ictal tienen una duración de 5 minutos. Las señales consideradas en el análisis fueron ACC, EDA, nivel de conductividad de la piel (SCL, por sus siglas en inglés), respuesta de la conductividad de la piel (SCR, por sus siglas en inglés), BVP (PPG), TEMP, ritmo cardiaco (estimado) y ritmo respiratorio (estimado).

El estudio de Knoth reporta el análisis de las 52 características en busca de diferencias significativas en los valores calculados de la característica para cada una de las fases. Si los valores calculados de una característica para las fases mencionadas anteriormente se diferencian significativamente, indica que la característica es apropiada para diferenciar las fases, por lo tanto, es apropiada también para utilizarse en la detección de crisis epilépticas. Para determinar si existe una diferencia significativa entre todas las fases para una característica, Knoth aplicó la prueba de Friedman con un nivel de significancia α de 0.05. Cuando se encontraron diferencias significativas entre todas las fases, aplicó la prueba Wilcoxon signed-rank para realizar una comparación por pares entre los valores de las características de la fase pre-normal y cada una de las demás fases (pre-ictal, ictal, post-ictal y post-normal). Esta comparación permitió establecer si había una diferencia significativa entre la fase pre-normal y las demás, dada una característica. Para la última prueba, el nivel adaptado de significancia α^* fue establecido en 0.0125.

Teniendo en cuenta el análisis de características realizado por Knoth, en esta disertación se establecieron tres criterios de inclusión para la selección de características de las señales ACC, EDA, SCL, SCR y BVP. Las señales SCL y SCR fueron consideradas porque pueden extraerse descomponiendo la señal EDA. Para su inclusión en el análisis de datos, cada característica debió:

1. Tener una probabilidad de significancia (*p-value*) menor que 0.05 para la prueba de Friedman.
2. Tener una probabilidad de significancia (*p-value*) menor que 0.0125 para la prueba Wilcoxon signed-rank tanto para la comparación pre-normal vs. ictal como para la comparación pre-normal vs. post-ictal.
3. Estar lo suficientemente descrita para ser calculada.

La Tabla 14 muestra las características seleccionadas, las cuales se extrajeron de las señales ACC, EDA, SCR y BVP, proceso descrito más adelante.

Signal	No.	Feature
ACC	1	STD of the X-axis
	2	STD of the Y-axis
	3	STD of the Z-axis
	4	STD of net acceleration
	5	VAR of the low-pass filter in X-axis
	6	VAR of the low-pass filter in Y-axis
	7	VAR of the low-pass filter in Z-axis
	8	DIFF min/max values of the low-pass filter in X-axis
	9	DIFF min/max values of the low-pass filter in Y-axis
	10	DIFF min/max values of the low-pass filter in Z-axis
	11	Mean of the accumulated first derivative of filtered signals
	12	STD of the accumulated first derivative of filtered signals
EDA	13	STD EDA
	14	FFT EDA energy
SCR	15	Integrated amplitude
	16	Maximum amplitude
	17	Number of peaks
BVP	18	FFT BVP energy

Tabla 14. Características seleccionadas con base en el trabajo de Knoth [41]. STD: Desviación estándar; VAR: Varianza; DIFF: Diferencia; FFT: Transformada rápida de Fourier.

4.2. Limpieza y pre-procesamiento de los datos

Para realizar la limpieza y pre-procesamiento de los datos seleccionados (Sección 4.1.1), se construyeron scripts para el lenguaje de programación Python. Dentro de éstos se utilizaron las librerías SciPy y Scikit-learn (Sección 2.4.2).

La señal EDA puede contener artefactos de movimiento que introducen ruido a la señal [48]. Con el fin de limpiar la señal de artefactos de movimiento y reducir componentes de alta frecuencia, se aplicó un filtro Butterworth paso bajo con una frecuencia de corte de 1.5 Hz [10].

Posteriormente, se descompuso la señal EDA en la señal SCR utilizando Ledapy [49]. Ledapy es una librería para Python que implementa una parte de la librería Ledalab, la cual es una librería para MATLAB para descomponer la señal EDA basada en deconvolución [50]. La señal SCR contiene cambios físicos en la conductividad eléctrica de la piel [51] y representa los periodos en los que la piel se vuelve una mejor conductora de la electricidad en el momento en que ocurre un estímulo fisiológico.

Teniendo en cuenta que los rangos de valores de las señales EDA, SCR y BVP pueden variar entre diferentes pacientes, se realizó una estandarización de estas señales en el rango 0 – 1 utilizando el método MinMaxScaler de la librería Scikit-learn.

Después de la estandarización de las señales se realizó la extracción de crisis epilépticas teniendo en cuenta la fase ictal y la post-ictal de cada crisis. En términos generales, es posible identificar tres momentos o fases relacionadas con la ocurrencia de una crisis epiléptica; el primero se conoce como pre-ictal, el cual representa el tiempo antes de la ocurrencia de la crisis epiléptica. El segundo momento se llama ictal, el cual representa el tiempo durante el cual ocurre una crisis. El tercer momento, post-ictal, representa el tiempo después de la ocurrencia de una crisis epiléptica, en el cual pueden observarse anomalías clínicas transitorias del sistema nervioso central [52].

La extracción de la fase ictal de cada crisis se realizó utilizando el archivo de anotaciones obtenido junto al dataset (Sección 3.3). El tiempo registrado de cada crisis en dicho archivo corresponde al periodo ictal (periodo en el que ocurre la crisis). La fase post-ictal de cada crisis se obtuvo extrayendo los 5 minutos siguientes a la fase ictal de la crisis, tal como se ha hecho en trabajos previos [10], [44]. Adicionalmente, se extrajeron las señales de cada paciente donde había ausencia de crisis epilépticas.

El archivo de anotaciones de crisis epilépticas contiene los tiempos de inicio y fin de las crisis epilépticas según anotaciones realizadas sobre los datos de la v-EEG, no sobre los datos de la pulsera. Para obtener los tiempos de las crisis equivalentes en

la pulsera, se calculó la diferencia de tiempo entre los eventos utilizados para la sincronización de los dos dispositivos, es decir, entre el artefacto en la señal de aceleración de las pulseras y el marcador de tiempo registrado en el software EEG (Sección 3.3.1).

Las señales extraídas fueron almacenadas en las carpetas *ictal*, *post_ictal* y *non_seizure* para facilitar el análisis de datos (Tabla 15); cada subcarpeta contenida dentro de las tres carpetas mencionadas contiene los siguientes archivos de señales: ACC.csv, BVP.csv, EDA.csv y SCR.csv. Cada subcarpeta de *ictal* y *post_ictal* contiene las señales de una crisis, identificadas de s1 – s44. Las subcarpetas de *non_seizure* representan las señales con ausencia de crisis de cada paciente.

ictal	post_ictal	non_seizure	
s1	s1	2014_07_25_d00367_p012_r	
s2	s2	2014_09_05_d00078_p009_l	
s3	s3	2014_09_06_d00037_p009_l	
...	
s44	s44	YYYY_MM_DD_d[deviceID]_p[patientID]_leftright	

Tabla 15. Estructura de carpetas de las señales extraídas.

4.3. Construcción de los datos

Esta tarea describe el proceso de construcción de los datos que serán utilizados por los algoritmos de clasificación. Este proceso se llevó a cabo mediante la caracterización de las señales.

4.3.1. Caracterización de señales.

Las características seleccionadas (Sección 4.1.2) fueron extraídas utilizando una ventana de 10 segundos en los segmentos *ictal*, *post-ictal* y *non_seizure* de las señales ACC, EDA, SCR y BVP. Las ventanas de los segmentos *ictal* y *post_ictal* se superpusieron en un 75% (*overlapping*). La superposición ha sido utilizada en estudios previos para caracterización de señales utilizando ventanas de tiempo [10], [27], [28]. Al conjunto de las 18 características calculadas sobre un segmento de 10

segundos de las señales se le denomina instancia. Después de la caracterización por segmento, cada instancia se etiquetó como crisis (*seizure*) o no crisis (*non_seizure*).

Las características 1 – 3 (STD of the X-axis, STD of the Y-axis, STD of the Z-axis) se extrajeron calculando la desviación estándar para cada eje del acelerómetro de acuerdo a la siguiente fórmula:

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (acc_j - \mu)^2} \quad (1)$$

Donde n representa la cantidad de datos del segmento, μ es la media aritmética de los datos del segmento y acc_j representa cada uno de los ejes del acelerómetro (acc_0, acc_1, acc_2).

La aceleración neta se calculó con la siguiente ecuación:

$$a_{net} = \sqrt{acc_0^2 + acc_1^2 + acc_2^2} \quad (2)$$

Por lo tanto, la característica 4 (STD of net acceleration) se calculó con:

$$\sigma_{anet} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (a_{net,i} - \mu_{anet})^2} \quad (3)$$

Donde N es el número de valores de aceleración neta y μ_{anet} es la media de los valores a_{net} dentro de un segmento [41].

Para detectar movimientos lentos en la señal de acelerometría, se aplicó un filtro paso bajo de Butterworth con una frecuencia de corte de 0.5 Hz a las señales de cada uno de los ejes del segmento de ACC. Las señales filtradas se utilizaron para el cálculo de las características 5 – 12. Las características 5 – 7 (VAR of the low-pass filter in X-axis, VAR of the low-pass filter in Y-axis, VAR of the low-pass filter in Z-axis) se extrajeron calculando la varianza de cada señal filtrada. La varianza de cada eje se calculó según la fórmula:

$$\sigma^2 = \sqrt{\frac{1}{n_2} \sum_{i=1}^{n_2} (x_i - \mu_2)^2} \quad (4)$$

Donde n_2 representa la cantidad de datos de las señales filtradas del segmento y μ_2 representa la media de los datos del segmento filtrado.

Las características 8 – 10 (DIFF min/max values of the low-pass filter in X-axis, DIFF min/max values of the low-pass filter in Y-axis, DIFF min/max values of the low-pass filter in Z-axis) se calcularon restando para cada eje filtrado el máximo valor del segmento menos el valor mínimo del segmento.

Adicionalmente se calculó la primera derivada de las señales filtradas y los tres ejes del segmento de ACC fueron combinados según la ecuación:

$$J[i] = \sqrt{\sum_{k=1}^3 \left(\frac{x_k[i] - x_k[i-1]}{\delta t} \right)^2} \quad (5)$$

Donde i es el índice para cada valor en el segmento filtrado de 10 segundos y x_k es la señal filtrada de eje k [41]. La media y la desviación estándar de este cálculo dieron como resultado las características 11 y 12 (Mean of the accumulated first derivative of filtered signals y STD of the accumulated first derivative of filtered signals).

La característica 13 (STD EDA) se obtuvo calculando la desviación estándar de los valores del segmento EDA de 10 segundos; el cálculo se realizó de forma análoga a la ecuación (1). Por otro lado, la característica 14 (FFT EDA energy) se calculó sobre el segmento EDA utilizando el módulo *fftpack* de la librería SciPy.

La característica 15 (Integrated amplitude) representa la suma de todos los valores pertenecientes al segmento de 10 segundos de la señal SCR; mientras que la característica 16 (Maximum amplitude) representa el valor máximo en dicho segmento. Por medio del módulo *signal* de la librería SciPy se calculó el número de picos presentes en el segmento de la señal SCR, este número representa la característica 17 (Number of peaks). Por último, la característica 18 (FFT BVP energy) se calculó por medio del módulo *fftpack* de la librería SciPy sobre el segmento de 10 segundos de la señal BVP.

4.4. Integración y formateo de los datos

Después de la caracterización de señales, se mezclaron los datos de todos los pacientes, conformando el dataset de características. El 96% del dataset de características quedó compuesto por instancias de la clase *non_seizure* (no crisis epiléptica), mientras que sólo el 4% de las instancias correspondía a la clase *seizure* (crisis epiléptica). Este desbalanceo de clases se debió a que los registros de señales estaban conformados en mayor parte por señales con ausencia de crisis epilépticas. Lo descrito representa el problema conocido como desbalanceo de clase binaria (*binary class imbalance*), en el cual la clase minoritaria es la clase de interés o clase positiva (*seizure*, en este caso). Este problema es encontrado de forma natural no sólo en el dominio del diagnóstico y predicción médica, sino también en bioinformática, mercadeo, detección de anomalías o fraudes, telecomunicaciones, biología, ecología, entre otros [53], [54].

Una gran cantidad de información de la clase negativa (*non-seizure*) en un dataset desbalanceado obstaculiza la capacidad de reconocimiento de las instancias de la clase positiva, además, las métricas generales de evaluación de la clasificación podrían proporcionar resultados engañosos [53], [55]. Por otro lado, el problema del desbalanceo de clases puede causar un deterioro significativo en métricas específicas de rendimiento de algoritmos de clasificación, debido al sesgo que produce la clase mayoritaria [56].

La tasa de desbalance (*imbalance ratio*) (IR, por sus siglas en inglés) es la tasa de instancias de la clase mayoritaria respecto a las instancias de la clase minoritaria [56]. Un dataset con un valor de IR superior a 1.5 es considerado desbalanceado, mientras que un valor de IR superior a 9 indica, con frecuencia, que un dataset está altamente desbalanceado [57], [58].

Se calculó el IR para el dataset de características, el cual dio el valor de 24, lo que lo cataloga como altamente desbalanceado. Para solucionar este problema, se aplicó una técnica de submuestro (*undersampling*).

El submuestreo elimina las instancias de la clase mayoritaria para reducir la tasa de desbalance [53]. La forma más sencilla de aplicar el submuestreo es eliminar instancias aleatorias de la clase mayoritaria; sin embargo, una eliminación aleatoria podría causar la pérdida de datos o conceptos potencialmente útiles de la clase

mayoritaria [59], [60]. Para evitar el problema mencionado, se han propuesto procedimientos alternativos de submuestreo basados en heurísticas [53]. En esta disertación se aplicó la técnica de submuestreo Instance Hardness Threshold (IHT), la cual se basa en el concepto de dureza de instancia propuesto por Smith et al. [61]. La técnica IHT se incluye en la librería para Python Imbalanced-learn [62].

Cuando se aplicó la técnica IHT, se eliminaron parte de las instancias de la clase mayoritaria, pero las instancias de la clase minoritaria permanecieron intactas. De esta manera, el dataset quedó conformado por el 25% de las instancias con clase *seizure* y el 75% de las instancias con clase *non_seizure*. Como resultado, la IR del dataset de características se redujo de 24 (altamente desbalanceado) a 3 (desbalanceado). Esta tasa de desbalance se redujo significativamente, pero no se redujo a 1 (igual cantidad de instancias para cada clase), ya que, en un contexto real de monitorización de pacientes epilépticos, los segmentos de la clase *non_seizure* son identificados la mayor parte del tiempo.

Posteriormente, el dataset fue dividido en dos particiones de forma aleatoria y estratificada (la tasa de desbalance de cada partición es la misma que la presente en los datos antes de la división). La primera partición fue conformada por el 75% de los datos, la cual se denomina dataset de entrenamiento. Esta se utilizó para el entrenamiento de clasificadores de *Machine Learning* (Sección 5.2). La última partición fue conformada por el 25% de los datos y es denominada dataset de prueba, con la cual se evaluó el mejor algoritmo de clasificación resultante de la fase de Modelado (Sección 5.3).

4.4.1. Descripción del dataset preparado

El dataset preparado está conformado por el dataset de entrenamiento (75% de los datos) y dataset de prueba (25% de los datos). Cada uno de los anteriores presenta una distribución interna de clases de la siguiente forma: el 75% de las instancias pertenecen a la clase *non_seizure*, mientras que el 25% restante pertenece a la clase *seizure*. Cada dataset se almacenó en dos archivos CSV diferentes (`dataset_e3_features_epilepsy_training.csv` y `dataset_e3_features_epilepsy_test.csv`), los cuales están almacenados según la estructura mostrada en la Tabla 16.

Nombre columna	Característica
acc_std_x	STD of the X-axis
acc_std_y	STD of the Y-axis
acc_std_z	STD of the Z-axis
acc_net_std	STD of net acceleration
acc_var_filtered_x	VAR of the low-pass filter in X-axis
acc_var_filtered_y	VAR of the low-pass filter in Y-axis
acc_var_filtered_z	VAR of the low-pass filter in Z-axis
acc_diff_min_max_x	DIFF min/max values of the low-pass filter in X-axis
acc_diff_min_max_y	DIFF min/max values of the low-pass filter in Y-axis
acc_diff_min_max_z	DIFF min/max values of the low-pass filter in Z-axis
acc_first_derivative_mean	Mean of the accumulated first derivative of filtered signals
acc_first_derivative_std	STD of the accumulated first derivative of filtered signals
eda_std	STD EDA
eda_fft_energy	FFT EDA energy
scr_integrated_amplitude	Integrated amplitude
scr_max_amplitude	Maximum amplitude
scr_number_peaks	Number of peaks
bvp_fft_energy	FFT BVP energy
class_1	Clase de la instancia (0: non_seizure, 1: seizure)

Tabla 16. Estructura de archivo de datasets de entrenamiento y prueba (formato CSV).

Capítulo 5

5. Modelado y Evaluación

En la fase de Modelado se seleccionan y aplican varias técnicas de clasificación de *Machine Learning* con el fin de encontrar los mejores algoritmos y los mejores parámetros que modelan el dataset de entrenamiento. Por otro lado, en la fase de Evaluación se evalúa con el dataset de prueba, el algoritmo que obtuvo el mejor rendimiento en la fase de Modelado. Las siguientes secciones describen este proceso, según las tareas mostradas en la Tabla 17.

Modelado	Selección de los algoritmos de clasificación
	Modelado y resultados
Evaluación	Evaluación del modelo
	- Métricas de rendimiento
	- Evaluación del rendimiento
	Evaluación de la calidad del proceso y del modelo desde la perspectiva clínica
	- Evaluación de acuerdo a estándares de validación clínica

Tabla 17. Tareas y resultados de las fases Modelado y Evaluación.

5.1. Selección de los algoritmos de clasificación

Se seleccionaron diez algoritmos de clasificación junto con sus parámetros por defecto para ser entrenados utilizando el dataset de entrenamiento. Los algoritmos fueron entrenados por medio de la librería Scikit-learn para Python, excepto el

algoritmo eXtreme Gradient Boosting, el cual se incluye en la librería para Python XGBoost [63]. Por razones de reproducibilidad, la Tabla 18 presenta los algoritmos seleccionados y los parámetros por defecto de cada uno de ellos, teniendo en cuenta que éstos pueden cambiar en diferentes versiones de la librería Scikit-learn.

Algoritmo / Configuración de parámetros por defecto
<p>Logistic Regression (LR) <code>LogisticRegression(penalty='l2', dual=False, tol=0.0001, C=1.0, fit_intercept=True, intercept_scaling=1, class_weight=None, random_state=None, solver='warn', max_iter=100, multi_class='warn', verbose=0, warm_start=False, n_jobs=None, l1_ratio=None)</code></p>
<p>Linear Discriminant Analysis (LDA) <code>LinearDiscriminantAnalysis(solver='svd', shrinkage=None, priors=None, n_components=None, store_covariance=False, tol=0.0001)</code></p>
<p>K-Neighbors Classifier (KNN) <code>KNeighborsClassifier(n_neighbors=5, weights='uniform', algorithm='auto', leaf_size=30, p=2, metric='minkowski', metric_params=None, n_jobs=None)</code></p>
<p>Decision Tree Classifier (CART) <code>DecisionTreeClassifier(criterion='gini', splitter='best', max_depth=None, min_samples_split=2, min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features=None, random_state=None, max_leaf_nodes=None, min_impurity_decrease=0.0, min_impurity_split=None, class_weight=None, presort=False)</code></p>
<p>Gaussian Naive Bayes (NB) <code>GaussianNB(priors=None, var_smoothing=1e-09)</code></p>
<p>Stochastic Gradient Descent (SGD) <code>SGDClassifier(loss='hinge', penalty='l2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=1000, tol=0.001, shuffle=True, verbose=0, epsilon=0.1, n_jobs=None, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5, early_stopping=False, validation_fraction=0.1, n_iter_no_change=5, class_weight=None, warm_start=False, average=False)</code></p>
<p>Support Vector Classification (SVM) <code>SVC(C=1.0, kernel='rbf', degree=3, gamma='auto_deprecated', coef0=0.0, shrinking=True, probability=False, tol=0.001, cache_size=200, class_weight=None, verbose=False, max_iter=-1, decision_function_shape='ovr', random_state=None)</code></p>
<p>Random Forest Classifier (RF) <code>RandomForestClassifier(n_estimators='warn', criterion='gini', max_depth=None, min_samples_split=2, min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', max_leaf_nodes=None, min_impurity_decrease=0.0, min_impurity_split=None, bootstrap=True, oob_score=False, n_jobs=None, random_state=None, verbose=0, warm_start=False, class_weight=None)</code></p>

Algoritmo / Configuración de parámetros por defecto	
Gradient Boosting Classifier (GBC)	<code>GradientBoostingClassifier(loss='deviance', learning_rate=0.1, n_estimators=100, subsample=1.0, criterion='friedman_mse', min_samples_split=2, min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_depth=3, min_impurity_decrease=0.0, min_impurity_split=None, init=None, random_state=None, max_features=None, verbose=0, max_leaf_nodes=None, warm_start=False, presort='auto', validation_fraction=0.1, n_iter_no_change=None, tol=0.0001)</code>
Extreme Gradient Boosting (XGBC)	<code>XGBClassifier(max_depth=3, learning_rate=0.1, n_estimators=100, verbosity=1, objective='binary:logistic', booster='gbtree', tree_method='auto', n_jobs=1, gpu_id=-1, gamma=0, min_child_weight=1, max_delta_step=0, subsample=1, colsample_bytree=1, colsample_bylevel=1, colsample_bynode=1, reg_alpha=0, reg_lambda=1, scale_pos_weight=1, base_score=0.5, random_state=0, missing=None)</code>

Tabla 18. Algoritmos de clasificación seleccionados y sus parámetros por defecto.

5.2. Modelado y resultados

Los diez algoritmos de clasificación seleccionados (Tabla 18) fueron entrenados y validados con el dataset de entrenamiento utilizando el método validación cruzada estratificada (*stratified cross-validation*) [64], el cual se aplicó con 10 pliegues (*folds*). Antes del entrenamiento, las características del dataset de entrenamiento fueron estandarizadas en el rango 0 – 1 utilizando el método *MinMaxScaler* de la librería *Scikit-learn*. Para la validación de los algoritmos se seleccionaron como métricas de entrenamiento las siguientes: exactitud (*accuracy*), precisión, sensibilidad (*recall*, *sensitivity*), y medida F (*F-measure*). En la Sección 5.3.1 se presenta una descripción de cada una de estas métricas.

La Tabla 19 presenta los resultados de las métricas de entrenamiento para cada algoritmo. Cada valor representa el valor medio de la métrica para los diez pliegues, además se presenta junto a cada valor el intervalo de confianza del 95%.

Algoritmo	Exactitud (%)	Precisión (%)	Sensibilidad (%)	Medida F (%)
LR	79 (+/- 2)	68 (+/- 6)	35 (+/- 6)	46 (+/- 7)
LDA	79 (+/- 2)	68 (+/- 6)	36 (+/- 4)	47 (+/- 4)
KNN	96 (+/- 1)	94 (+/- 3)	90 (+/- 3)	92 (+/- 3)
CART	98 (+/- 1)	95 (+/- 2)	96 (+/- 2)	96 (+/- 2)
NB	72 (+/- 2)	44 (+/- 4)	33 (+/- 6)	38 (+/- 5)
SGD	79 (+/- 2)	70 (+/- 4)	36 (+/- 7)	47 (+/- 7)

Algoritmo	Exactitud (%)	Precisión (%)	Sensibilidad (%)	Medida F (%)
SVM	79 (+/- 1)	73 (+/- 5)	30 (+/- 4)	43 (+/- 4)
RF	99 (+/- 1)	98 (+/- 2)	97 (+/- 1)	97 (+/- 1)
GBC	97 (+/- 1)	94 (+/- 2)	92 (+/- 2)	93 (+/- 2)
XGBC	97 (+/- 1)	94 (+/- 2)	94 (+/- 2)	94 (+/- 2)

Tabla 19. Resultados de métricas de entrenamiento de cada uno de los algoritmos seleccionados, utilizando el método validación cruzada estratificada con 10 pliegues.

Como puede observarse en la Tabla 19, el algoritmo Random Forest obtuvo el mejor rendimiento para todas las métricas, en comparación con los demás algoritmos. La Figura 5 muestra un diagrama de caja, el cual permite comparar la exactitud de los modelos en la etapa de entrenamiento.

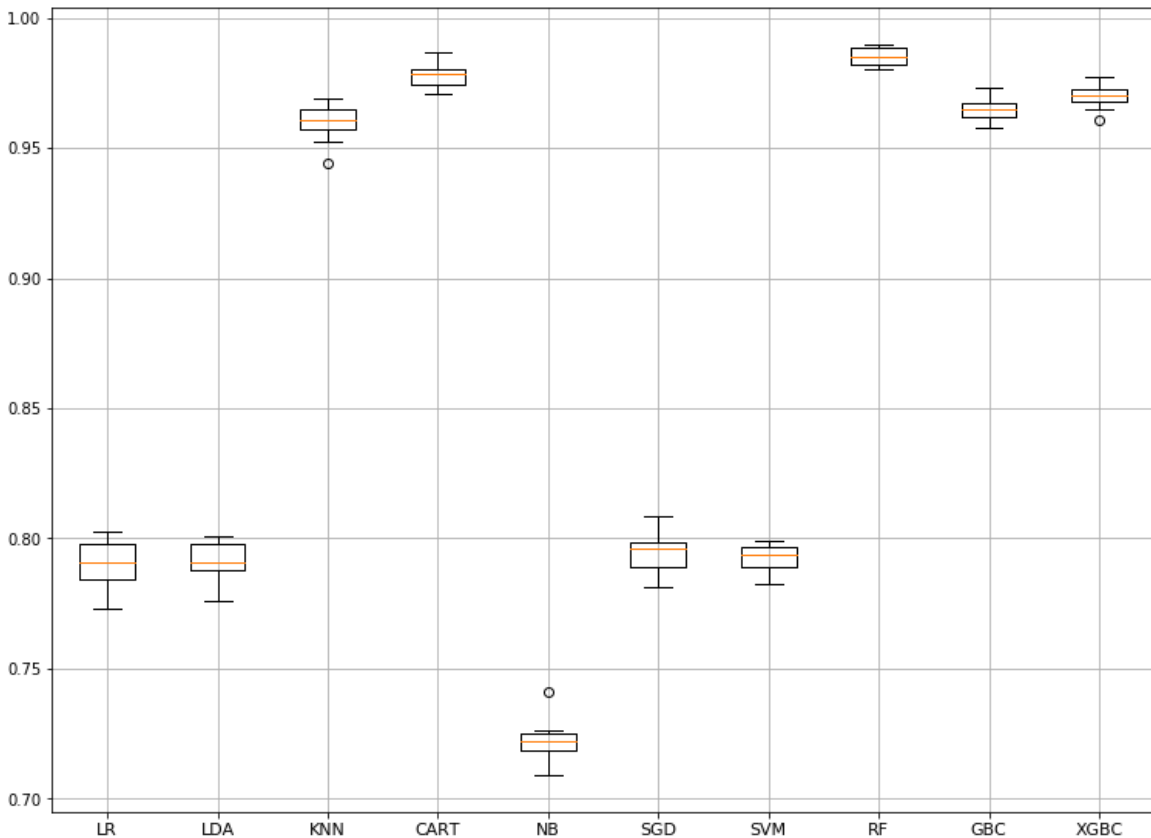


Figura 5. Exactitud alcanzada por cada algoritmo entrenado con parámetros por defecto en una configuración de validación cruzada estratificada de 10 pliegues.

Teniendo en cuenta los resultados del entrenamiento de los algoritmos, se selecciona el algoritmo Random Forest y sus parámetros por defecto, como el mejor algoritmo para la clasificación de instancias de clases *seizure* y *non_seizure*.

5.3. Evaluación del modelo

En esta sección se describen las métricas de rendimiento calculadas, el procedimiento de evaluación realizado sobre el algoritmo seleccionado en la Sección 5.2, y los resultados de la evaluación.

5.3.1. Métricas de rendimiento

Existen diferentes métricas que pueden calcularse con el fin de evaluar el rendimiento de un modelo de *Machine Learning*. La elección de éstas debe considerar el tipo de problema de minería de datos abordado, así como las características de los datos. En problemas de clasificación binaria, con frecuencia se calculan métricas derivadas de los valores de la matriz de confusión. Una matriz de confusión resume el rendimiento en la clasificación de un modelo, respecto a determinados datos de prueba, mediante la comparación de los valores predichos por el modelo vs. los valores reales [65]. La Tabla 20 muestra los valores de la matriz de confusión calculados en esta disertación.

		Valores de la predicción	
		Clase negativa (<i>non-seizure</i>)	Clase positiva (<i>seizure</i>)
Valores reales	Clase negativa (<i>non-seizure</i>)	Verdaderos Negativos (TN)	Falsos Positivos (FP)
	Clase positiva (<i>seizure</i>)	Falsos Negativos (FN)	Verdaderos Positivos (TP)

Tabla 20. Matriz de confusión y valores calculados.

Los verdaderos negativos (TN, por sus siglas en inglés) representan instancias que estaban etiquetadas como *non-seizure* en el dataset de prueba y que fueron

clasificadas como tal por el modelo. Los falsos positivos (FP, por sus siglas en inglés) representan instancias etiquetadas como *non-seizure* en el dataset de prueba pero que fueron clasificadas como *seizure* por el modelo. Los falsos negativos (FN, por sus siglas en inglés) representan instancias etiquetadas como *seizure* en el dataset de prueba pero que fueron etiquetadas como *non-seizure* por el modelo. Finalmente, los verdaderos positivos (TP, por sus siglas en inglés) representan instancias etiquetadas como *seizure* en el dataset de prueba y que fueron clasificadas como tal por el modelo.

A continuación se presentan las métricas que fueron calculadas en esta disertación, las cuales se derivan de la matriz de confusión. También se presenta la interpretación de cada métrica.

Exactitud (*accuracy*): Es la tasa de instancias correctamente clasificadas por el modelo del total de instancias que comprenden el dataset de prueba. Se calcula con la fórmula:

$$Exactitud = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

Precisión (*presicion*): También se denomina valor predictivo positivo (PPV). Representa la tasa de instancias *seizure* correctamente clasificadas entre todas las instancias clasificadas como *seizure* por el modelo. Se calcula con la fórmula:

$$Precisión = \frac{TP}{TP + FP} \quad (7)$$

Sensibilidad (*recall, sensitivity*): Es la tasa de instancias de clase *seizure* correctamente clasificadas por el modelo entre el número total de instancias *seizure* del dataset de prueba. Se calcula con la fórmula:

$$Sensibilidad = \frac{TP}{TP + FN} \quad (8)$$

Medida F (*F-measure*): Se define como la media armónica de la precisión y la sensibilidad. Valores altos de esta métrica representan valores altos y equilibrados de precisión y sensibilidad. Por el contrario, valores bajos representan valores bajos en estas últimas o que alguna de estas dos métricas sobresale a expensas de la otra. Se calcula con la fórmula:

$$Medida F = 2 \times \frac{Precisión \times Sensibilidad}{Precisión + Sensibilidad} \quad (9)$$

Tasa de falsos positivos (FPR, por sus siglas en inglés): Es la tasa de instancias de clase *non-seizure* incorrectamente clasificadas entre todas las instancias clasificadas como *non-seizure* por el modelo. Representa la probabilidad de que el modelo genere falsas alarmas (falsos positivos). Se calcula con la fórmula:

$$FPR = \frac{FP}{FP + TN} \quad (10)$$

Coeficiente de Correlación de Mateo (*Matthews Correlation Coefficient, MCC*): Es un coeficiente que permite evaluar modelos cuando contienen una distribución desbalanceada de clases; un coeficiente de +1 representa una predicción perfecta, 0 una predicción aleatoria promedio y -1 una predicción inversa [66]. Se calcula con la fórmula:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (11)$$

5.3.2. Evaluación del rendimiento

La evaluación del rendimiento fue realizada con el algoritmo seleccionado al final de la etapa de Modelado (Sección 5.2), es decir, con el algoritmo Random Forest con la configuración de sus parámetros por defecto.

Inicialmente se entrenó el algoritmo Random Forest con parámetros por defecto utilizando el dataset de entrenamiento para generar un modelo de *Machine Learning* capaz de diferenciar instancias pertenecientes a las clases *seizure* y *non_seizure*. Por último, el modelo de *Machine Learning* fue evaluado con el dataset de prueba, mediante la predicción de las instancias y el cálculo de la matriz de confusión.

Las métricas calculadas para la evaluación del modelo fueron exactitud, precisión, sensibilidad, medida F, FPR y MCC, las cuales fueron calculadas a partir de los valores de la matriz de confusión resultante.

La Tabla 21 muestra las métricas de evaluación del modelo de *Machine Learning*, calculadas a partir de la predicción realizada sobre el dataset de prueba. Por otro lado, también se muestra la matriz de confusión de la evaluación del modelo

Random Forest en el dataset de prueba (Figura 6.). Los valores de la matriz de confusión representan el porcentaje de instancias en el rango de 0 a 1.

Exactitud (%)	Precisión (%)	Sensibilidad (%)	Medida F(%)	FPR (%)	MCC
99	97	98	97%	1%	0.96

Tabla 21. Métricas de evaluación del modelo de *Machine Learning* (Random Forest).

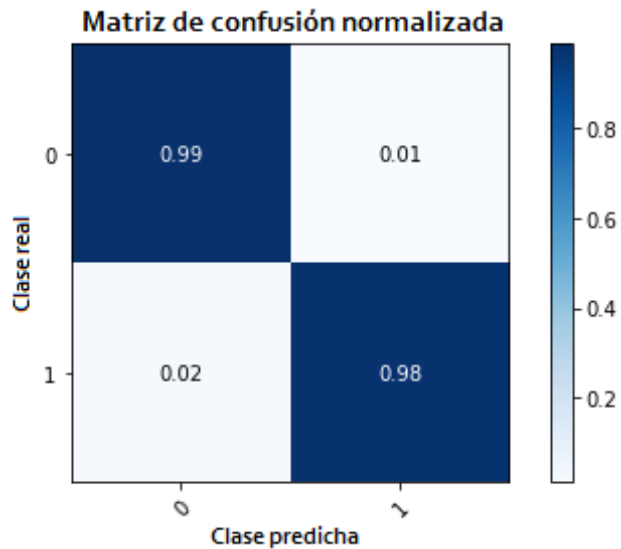


Figura 6. Matriz de confusión de la evaluación del modelo de *Machine Learning* (Random Forest). Los valores están dados en términos de porcentaje en el rango 0 – 1. Las etiquetas cero (0) representan la clase *non_seizure*; las etiquetas uno (1), la clase *seizure*.

Teniendo en cuenta los resultados de la evaluación, y el proceso de análisis de los datos realizado hasta el momento, se puede afirmar que el modelo de *Machine Learning* (Random Forest) entrenado con el dataset de entrenamiento cumple los criterios de éxito de la minería de datos (Sección 2.3.2) y los criterios de éxito del negocio (Sección 2.1.3). En este sentido, el modelo final satisface tanto el objetivo de la minería de datos, como el objetivo del negocio. Este modelo fue seleccionado para la fase de Despliegue (Capítulo 6).

5.4. Evaluación de la calidad del proceso y del modelo desde la perspectiva clínica

En esta sección se presenta una evaluación desde la perspectiva clínica de la calidad del modelo resultante de todo el proceso de minería de datos. Para este propósito se aplicó una herramienta de evaluación, cuya descripción y resultado se presenta a continuación.

5.4.1. Evaluación de acuerdo a estándares de validación clínica

El modelo final y el proceso de minería de datos (diseño del estudio) se evaluó de acuerdo a estándares para la validación clínica de dispositivos de detección de crisis epilépticas (*Seizure Detection Devices*, SDD) propuesto por Beniczky y Ryvlin [67]. De acuerdo a estos estándares, existen cuatro características clave a tener en cuenta en estudios de SDD: sujetos, registros, análisis de datos y alarmas, y estándar de referencia. Para cada una de estas cuatro características, existen diferentes aspectos que deben ser considerados por los estudios. Dependiendo del cumplimiento o no de éstos aspectos, un estudio puede ser clasificado en una de cinco fases, las cuales indican la robustez del estudio.

La Tabla 22 muestra las cuatro características clave para la validación clínica de SDD, los aspectos a considerar de cada una de ellas y las fases en las que un estudio puede ser clasificado.

La fase 0 corresponde a estudios iniciales relacionados con propuestas o desarrollos de métodos novedosos. La fase 1 corresponde a estudios de prueba de principio (*proof-of-principle studies*). La fase 2 corresponde a estudios que utilizan un dispositivo de detección dedicado. La fase 3 corresponde a estudios que consideran la confirmación final de la seguridad y la precisión de los SDD. La fase 4 consiste en estudios de campo de SDD utilizados en entornos cotidianos de los pacientes, los cuales abordan aspectos relacionados con la usabilidad.

Diseño del estudio	Fase				
	0	1	2	3	4
Sujetos					
Simulación / Sujetos sanos	+	±	-	-	-
Número de pacientes con crisis epilépticas	-	≥1	≥10	≥20	≥50
Número de crisis epilépticas	-	≥1	≥15	≥30	≥75
Registros					
Métodos convencionales	±	±	-	-	-
Dispositivo dedicado	±	±	+	+	+
Continuo	±	±	±	+	+
Multi-centro	±	±	±	+	na
Offline / Retrospectivo	+	+	+	-	-
Análisis de datos y alarmas					
Entrenamiento y prueba utilizando el dataset	+	+	+	-	-
Algoritmo y valores de corte predefinidos	-	-	-	+	+
Tiempo real	-	-	-	+	+
Ocultamiento	-	-	-	+	+
Estándar de referencia					
Registros de video o video-EEG	±	+	+	+	±
Información proveniente de pacientes y cuidadores	±	-	-	-	±

Tabla 22. Fases del desarrollo y validación clínica de SDD, según características clave. +, obligatorio; ±, opcional; -, excluido; na, no aplica.

Fuente: Beniczky y Ryvlin [67].

Teniendo en cuenta los estándares para la validación clínica de SDD, se realizó la evaluación de calidad que involucra al modelo final presentado en esta disertación y al proceso de minería de datos seguido (diseño del estudio). Los resultados se presentan en la Tabla 23.

Diseño del estudio	Resultado	Justificación
Sujetos		
Simulación / Sujetos sanos	-	Todos los sujetos incluidos en el estudio son pacientes reales con diagnóstico de epilepsia.
Número de pacientes con crisis epilépticas	11	Ver Tabla 12.
Número de crisis epilépticas	44	Ver Tabla 13.
Registros		

Diseño del estudio	Resultado	Justificación
Métodos convencionales	-	Se utilizó un dispositivo dedicado para la recolección de señales.
Dispositivo dedicado	+	Se utilizó la pulsera E3 wristband.
Continuo	+	Las señales fueron capturadas en periodos de tiempo continuos de 24 horas.
Multi-centro	-	La recolección se realizó sólo en una unidad de v-EEG.
Offline / Retrospectivo	+	Las señales fueron capturadas y almacenadas en un dataset. Posteriormente se realizó el análisis utilizando dicho dataset.
Análisis de datos y alarmas		
Entrenamiento y prueba utilizando el dataset	+	El dataset se dividió en dos partes, una para la fase de modelado y la otra para la fase de evaluación.
Algoritmo y valores de corte predefinidos	-	El estudio no partió de un algoritmo ya entrenado y con parámetros ajustados.
Tiempo real	-	El modelo final no realiza el análisis de forma inmediata a la recolección de señales, los datos son procesados y analizados cuando la recolección de señales finaliza.
Ocultamiento	-	En los trabajos que recolectaron los datos no se reporta una restricción de acceso a las señales de la pulsera al especialista que realizó el registro de las crisis epilépticas con el estándar de referencia.
Estándar de referencia		
Registros de video o video-EEG	+	El estándar de referencia utilizado para la identificación de crisis epilépticas fue la v-EEG.
Información proveniente de pacientes y cuidadores	-	La información utilizada en la identificación de crisis provino únicamente de los registros de v-EEG.

Tabla 23. Evaluación de la calidad del proceso y del modelo final desde la perspectiva clínica según los estándares propuestos por Beniczky y Ryvlin [67].

Los estándares para la validación clínica de SDD también proponen un conjunto de medidas de resultados (métricas) que deben ser reportadas con el fin de estimar el

rendimiento. La Tabla 24 muestra las medidas de resultados propuestas, y los valores calculados en esta disertación.

Medida de resultado	Valor
Tiempo total de registros	A rededor de 500 horas, con una duración promedio por registro de aproximadamente 24 horas.
Tiempo de deficiencia del dispositivo	En los trabajos de recolección del dataset no se reportaron fallas del dispositivo. Éstas tampoco se identificaron a partir de las señales analizadas.
Seguridad	El dispositivo E3 wristband aprueba estándares internacionales de seguridad. También ha sido aprobado por las juntas de revisión institucional en los principales hospitales de la UE y EE.UU [45].
Sensibilidad	98%
Tasa de falsas alarmas (FPR)	1%
Latencia de detección	No aplica. Los datos no se recolectaron en tiempo real.
Experiencia de usuario	No aplica. Esta valoración está por fuera del alcance de la investigación.

Tabla 24. Medidas de resultado correspondientes al proceso de minería de datos y al modelo final.

De acuerdo a los resultados de evaluación presentados en la Tabla 23 y las medidas de resultado presentadas en la Tabla 24, puede afirmarse que el modelo final clasifica como un detector de crisis epilépticas que se encuentra en la Fase 2 de calidad desde la perspectiva clínica, según los estándares para la validación clínica de SDD propuestos por Beniczky y Ryvlin [67].

Un estudio clasificado en la Fase 2 es aquel que utiliza un dispositivo de detección de crisis dedicado, incluye un número de pacientes epilépticos mayor o igual a 10, los cuales presenten por lo menos 15 crisis epilépticas. Además, considera la seguridad del dispositivo y utiliza un estándar de referencia para la identificación de crisis epilépticas. En esta fase se permite el análisis retrospectivo y el uso del mismo dataset para el entrenamiento y la evaluación.

La clasificación del proceso de minería de datos (diseño del estudio) y del modelo final como Fase 2, es un buen indicador de calidad, lo cual permite afirmar que el modelo final tiene un alto potencial para ser incorporado en sistemas para el

monitoreo continuo de crisis epilépticas, los cuales permitan llevar un registro automático y preciso de las crisis epilépticas de los pacientes. Si se quisiera alcanzar este propósito, el presente estudio debe ampliarse de tal manera que abarque las fases 3 y 4 de los estándares de validación clínica de SDD, lo cual se plantea como un trabajo futuro.

Capítulo 6

6.Despliegue

Esta es la última fase del proceso de minería de datos definido por CRISP-DM. El propósito de la fase de Despliegue es presentar el modelo final de una forma apropiada para su uso por parte de los usuarios finales.

En esta disertación se realizó el despliegue a través de la integración del modelo final con una aplicación Android y una pulsera E4 wristband. El propósito de la integración fue conformar un sistema móvil para la detección de eventos epileptiformes, el cual permita el conteo diario y automático de dichos eventos. Este sistema, el cual es alternativo a la EEG, tiene como usuarios objetivo los pacientes epilépticos y tiene el potencial de reemplazar a los auto-informes, teniendo en cuenta la problemática descrita en el Capítulo 1. La Tabla 25 presenta las tareas realizadas durante la fase de Despliegue.

Despliegue	Arquitectura de despliegue - Descripción del funcionamiento del sistema detector - Arquitectura de despliegue del sistema detector
	Aplicación móvil para la detección automática de eventos epileptiformes

Tabla 25. Tareas y resultados de la fase de despliegue.

6.1. Arquitectura de despliegue

El despliegue fue realizado mediante la integración del modelo de *Machine Learning* con una aplicación móvil y con una pulsera, generando así un sistema automático para la detección de eventos epileptiformes.

6.1.1. Descripción del funcionamiento del sistema detector

A continuación se describe el funcionamiento general del sistema detector. El propósito de esta sección no es describir de forma técnica la arquitectura del sistema, sino presentar una visión general de los componentes que interactúan, sus funciones principales, y el flujo de datos desde la recolección de señales hasta la clasificación por parte del sistema.

El sistema automático para la detección de eventos epileptiformes está conformado por una pulsera E4 wristband, una aplicación móvil, y los módulos de pre-procesamiento, caracterización de señales y *Machine Learning* (Figura 7).

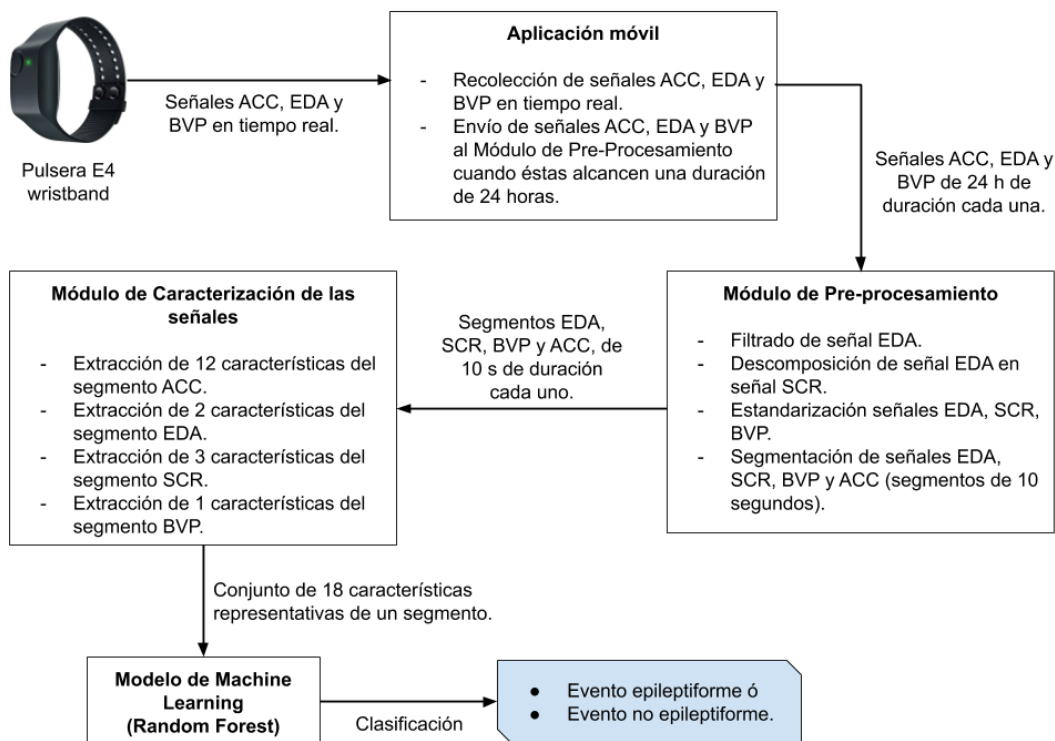


Figura 7. Componentes y funcionamiento general del sistema automático para la identificación de eventos epileptiformes.

Inicialmente, la pulsera conectada a un brazo del paciente envía señales de los sensores ACC, EDA y BVP en tiempo real a una aplicación móvil. La aplicación móvil recolecta estas señales y las almacena en su memoria interna, hasta que alcanza las 24 horas de grabación. Cuando esto sucede, la aplicación móvil envía las señales ACC, EDA y BVP (las cuales tienen una duración de 24 h), al módulo de pre-procesamiento, el cual realiza los pasos descritos en la Sección 4.2 (Limpieza y pre-procesamiento de los datos). Una vez las señales hayan sido pre-procesadas, el módulo de pre-procesamiento envía segmentos de las señales EDA, SCR, BVP y ACC (de 10 segundos de duración cada uno), al módulo de caracterización de señales. Este último módulo extrae las 18 características de los segmentos, según lo descrito en la Sección 4.3.1 (Caracterización de señales). Finalmente, el módulo de caracterización de señales envía las características extraídas de los segmentos al Modelo de *Machine Learning*, el cual realiza la clasificación del segmento e indica si es un evento epileptiforme o no lo es.

6.1.2. Arquitectura de despliegue del sistema detector.

Se desarrolló una aplicación móvil para Android la cual se conecta vía Bluetooth Low Energy (BLE) a una pulsera E4 wristband. La aplicación móvil recibe en tiempo real las señales ACC, EDA y BVP de la pulsera mientras está ubicada en el brazo del paciente. Cuando las señales completen 24 horas de grabación, la aplicación las envía a un servicio desplegado en la nube el cual integra el pre-procesamiento de las señales, la caracterización de señales y el modelo de *Machine Learning*. Una vez el servicio procesa las señales de 24 horas, envía a la aplicación móvil el número de eventos epileptiformes identificados. Este número representa un signo de alarma para el paciente y no un diagnóstico médico (Figura 8).

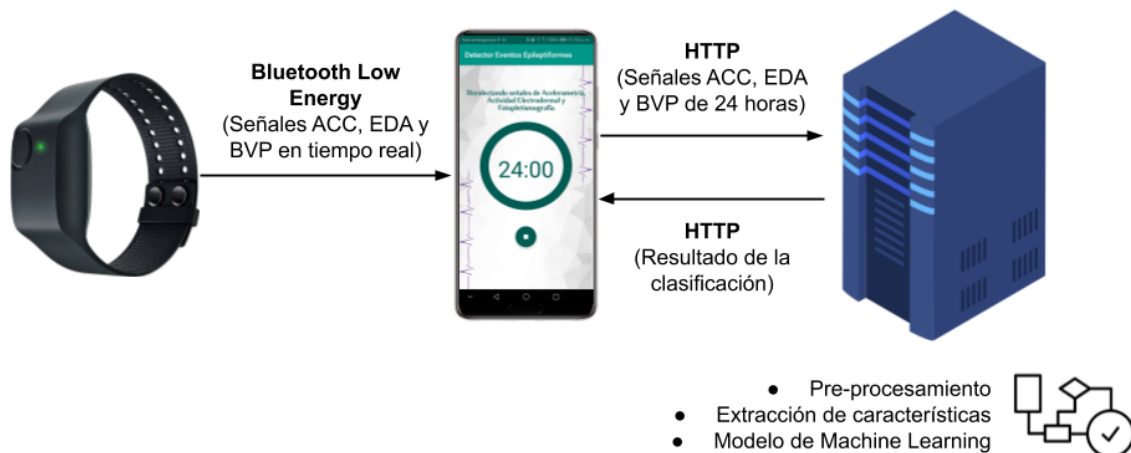


Figura 8. Arquitectura de despliegue del sistema detector.

6.2. Aplicación móvil para la detección automática de eventos epileptiformes

Tanto la aplicación móvil como el servicio en la nube descritos en la Sección 6.1.2 fueron desarrollados con el fin de desplegar el sistema de forma completa y funcional. A continuación se muestran capturas de pantalla de las principales interfaces de la aplicación móvil y su funcionamiento.

La Figura 9 muestra capturas de pantalla de las interfaces iniciales de la aplicación móvil. Al iniciar la aplicación por primera vez en un dispositivo Android (Smartphone o tableta), la aplicación solicitará la autorización de los permisos necesarios para su funcionamiento. Una vez el usuario enciende la pulsera E4 wristband y presiona el botón “Conectar Pulsera”, la aplicación realiza el emparejamiento con el dispositivo. En este momento la aplicación está lista para recolectar las señales provenientes de la pulsera (Figura 10a).

Cuando el usuario presiona el botón “Iniciar recolección de señales”, la aplicación inicia la recolección de señales y su almacenamiento en la memoria interna del dispositivo Android. Durante la recolección de señales la aplicación muestra un gráfico que indica al usuario cuánto tiempo restante queda de la grabación, teniendo en cuenta que la grabación es de 24 horas; el usuario puede detener la grabación cuando él lo desee (Figura 10b).

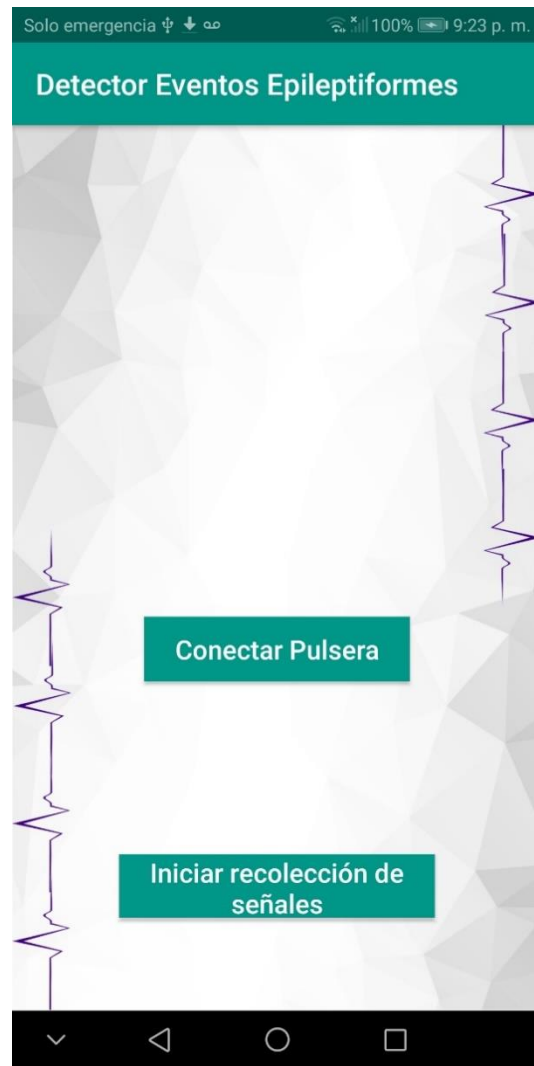
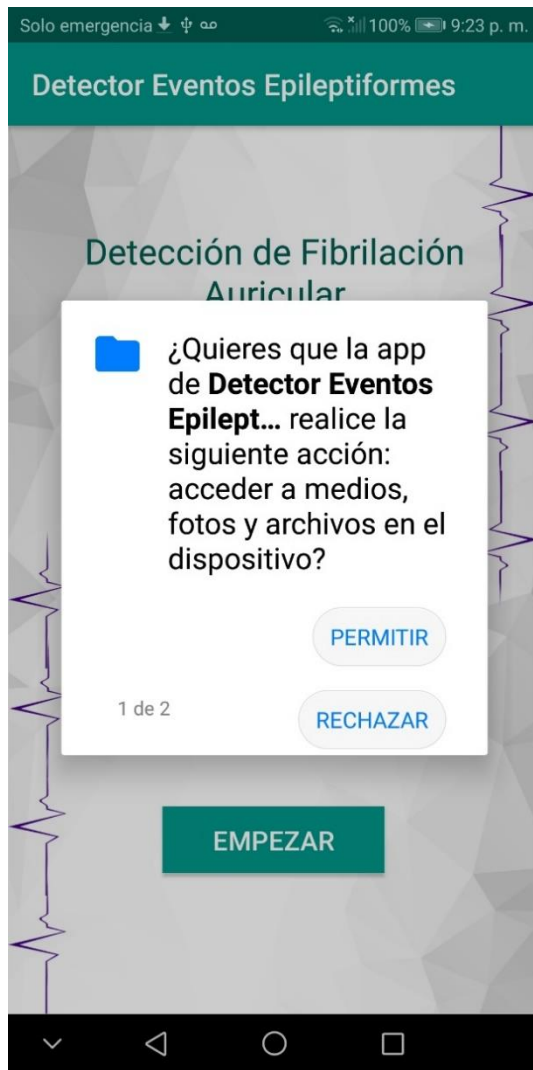


Figura 9. Capturas de pantalla de las interfaces iniciales de la aplicación móvil.



Figura 10. Capturas de pantalla de las interfaces para la recolección de señales ACC, EDA y PPG.

Cuando se completan las 24 horas de grabación o cuando el usuario detenga la recolección de señales, el sistema envía las señales al servicio en la nube, el cual procesa y clasifica las señales. Al finalizar el análisis, la aplicación móvil muestra al usuario la cantidad de eventos epileptiformes identificados (Figura 11).

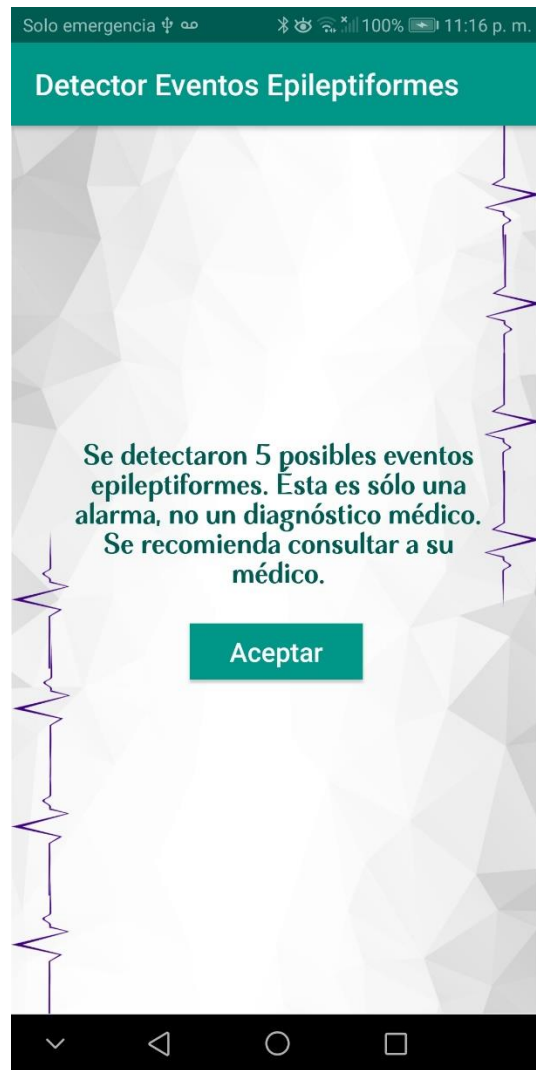


Figura 11. Capturas de pantalla de las interfaces de visualización de resultados de la detección.

Capítulo 7

7. Informe final

Al final del proceso de minería de datos, CRISP-DM propone un reporte con la presentación de los resultados alcanzados. La decisión del tipo de reporte a realizar se toma respecto a la audiencia del informe final. Teniendo en cuenta lo anterior, a continuación se presentan las contribuciones, conclusiones y los trabajos futuros.

7.1. Contribuciones

Las contribuciones de esta disertación son:

1. Un detector de crisis epilépticas motoras y no motoras basado en técnicas de *Machine Learning* para dispositivos wearable que incorporen sensores de ACC, EDA y PPG. El detector obtuvo en la evaluación 99% de exactitud, 97% de precisión, 98% de sensibilidad, 97% de medida F, 1% de FPR, y 0.96 de MCC. Con esta contribución se da cumplimiento al objetivo general de esta disertación.
2. Un dataset de características extraídas de las señales ACC, EDA, SCR y BVP de pacientes epilépticos, el cual contiene instancias etiquetadas según la presencia o no de crisis epilépticas. Las crisis epilépticas consideradas incluyen crisis de tipo motor y crisis de tipo no motor.

3. Un sistema móvil para la detección automática de eventos epileptiformes que integra el modelo de *Machine Learning* desplegado como servicio en la nube, una aplicación móvil para Android y una pulsera E4 wristband.
4. Un proceso de selección y descripción de características de señales ACC, EDA, SCR y BVP, enmarcadas en la detección de crisis epilépticas.
5. Publicación de un artículo de investigación:
Molina E., Salazar-Cabrera R., López D.M. (2018) *NeuroEHR: Open Source Telehealth System for the Management of Clinical Data, EEG and Remote Diagnosis of Epilepsy*. In: Figueroa-García J., López-Santana E., Rodríguez-Molano J. (eds) *Applied Computer Sciences in Engineering. WEA 2018. Communications in Computer and Information Science*, vol 915. Springer, Cham.
6. Sometimiento de artículo de investigación a revista *IEEE Journal of Translational Engineering in Health & Medicine*:
Título: *Intelligent Telehealth System to support Epilepsy Management*.
Autores: **E. Molina**, C. Sarmiento, R. Salazar, D. López, R. Vargas.
7. Elaboración de artículo de investigación, el cual será enviado a revisión a la revista *Sensors*. Este artículo contó con la participación de investigadores del MaD Lab.
Título: *Semiology-Independent Detector of Epileptic Seizures for Wearable Devices Based in Accelerometry, Electrodermal Activity, and Photoplethysmography*.
Autores: **Edward Molina**, Diego M. López, Rubiel Vargas-Canas, Ricardo Salazar-Cabrera, Bjoern Eskofier, Thomas Kautz y Burkhard Kasper.
8. Como contribución indirecta se realizó una propuesta doctoral para el programa de Doctorado en Ingeniería Telemática de la Universidad del Cauca, partiendo de los conocimientos sobre *Machine Learning* y e-Salud alcanzados durante el desarrollo de la investigación reportada en esta disertación. La propuesta y hoja de vida del investigador fue la mejor evaluada en el proceso de selección del Doctorado en Ingeniería Telemática realizado durante el primer semestre del año 2019.

Título: Modelo de Aprendizaje Automático para Predicción de Riesgo de Lesiones Gástricas Pre-Neoplásicas.

*Autor: **Edward Molina.***

Director: Diego M. López, PhD.

7.2. Conclusiones

1. Los detectores de crisis epilépticas alternativos a la EEG que han sido propuestos hasta el momento (Sección 2.1.1) presentan deficiencias como la no identificación de ciertos tipos de crisis epilépticas (en especial aquellas con ausencia de movimiento), el uso de dispositivos incómodos de portar y altas tasas de falsos positivos en la detección.
2. El detector propuesto en la presente disertación es capaz de diferenciar crisis epilépticas motoras y no motoras, con un buen nivel de precisión y con una baja tasa de falsos positivos. De esta manera, aporta satisfactoriamente a la problemática identificada.
3. La limpieza, pre-procesamiento de los datos, así como la caracterización de las señales ACC, EDA, SCR y BVP fueron pasos clave para la obtención de buenos resultados en la evaluación del rendimiento del modelo final.
4. La presente disertación realizó un aporte a la brecha de investigación número 3 identificada en la Sección 2.1.1, a pesar de que ésta no fue considerada inicialmente en los objetivos específicos (Sección 1.4.2). Esta brecha se identificó como "La mayoría de los trabajos presentan detectores o algoritmos aislados, es decir no integrados a ninguna plataforma o sistema para el monitoreo de crisis epilépticas". La contribución relacionada con esta brecha es el sistema móvil para la detección automática de eventos epileptiformes presentado en la Sección 6.2.
5. El sistema móvil para la detección automática de eventos epileptiformes tiene el potencial de aportar en el monitoreo de crisis epilépticas, como alternativa al método tradicional de auto-informe.

6. Desplegar el modelo en un sistema o prototipo real es fundamental para entender el funcionamiento de modelos de *Machine Learning* en etapas de producción, ya que esto implica la integración del modelo con diferentes tecnologías y flujos de datos.

7.3. Trabajos futuros

1. El dataset utilizado en el proceso de minería de datos y las anotaciones de las crisis epilépticas identificadas posibilitan investigaciones futuras relacionadas con la discriminación del tipo de crisis identificada. En este sentido, se podrían proponer clasificadores multi-clase que identifiquen el tipo de semiología de las crisis epilépticas.
2. Experimentar en la construcción de modelos personales a partir del dataset, es decir, modelos adaptados a las señales fisiológicas propias de cada paciente.
3. Realizar una evaluación experimental en un contexto real del sistema móvil para la detección automática de eventos epileptiformes.
4. Realizar las mejoras y ajustes necesarios para que el detector presentado en esta disertación se adapte a los aspectos relacionados con la fase 3 y 4 de los estándares para la validación clínica de SDD propuestos por Beniczky y Ryvlin [67].

Referencias

- [1] R. S. Fisher *et al.*, “ILAE Official Report: A practical clinical definition of epilepsy,” *Epilepsia*, vol. 55, no. 4, pp. 475–482, 2014.
- [2] R. S. Fisher *et al.*, “Epileptic seizures and epilepsy: Definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE),” *Epilepsia*, vol. 46, no. 4, pp. 470–472, 2005.
- [3] World Health Organization (WHO), “Epilepsy,” 2019. [Online]. Available: <http://www.who.int/news-room/fact-sheets/detail/epilepsy>. [Accessed: 30-Oct-2019].
- [4] A. Velez and J. Eslava-Cobos, “Epilepsy in Colombia: epidemiologic profile and classification of epileptic seizures and syndromes,” *Epilepsia*, vol. 47, no. 1, pp. 193–201, 2006.
- [5] M. de S. y P. Social, *Guía de Práctica Clínica (GPC) para la prevención, diagnóstico, tratamiento y rehabilitación de los Pacientes con Epilepsia en el Contexto Colombiano*, 23rd ed. Bogotá, Colombia, 2013.
- [6] C. E. Elger and C. Hoppe, “Diagnostic challenges in epilepsy: seizure under-reporting and seizure detection,” *Lancet Neurol.*, vol. 17, no. 3, pp. 211–212, 2018.
- [7] A. Aghaei-Lasboo and R. S. Fisher, “Methods for Measuring Seizure Frequency and Severity,” *Neurol. Clin.*, vol. 34, no. 2, pp. 383–394, 2016.
- [8] W. Van Paesschen, “The future of seizure detection,” *Lancet Neurol.*, vol. 17, no. 3, pp. 211–212, 2018.
- [9] J. Bidwell, T. Khuwatsamrit, B. Askew, J. A. Ehrenberg, and S. Helmers, “Seizure reporting technologies for epilepsy treatment: A review of clinical information needs and supporting technologies,” *Seizure*, vol. 32, pp. 109–117, 2015.
- [10] B. E. Heldberg, T. Kautz, H. Leutheuser, R. Hopfengartner, B. S. Kasper, and B. M. Eskofier, “Using wearable sensors for semiology-independent seizure detection - Towards ambulatory monitoring of epilepsy,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2015-Novem, pp. 5593–5596, 2015.
- [11] C. Jory, R. Shankar, D. Coker, B. Mclean, J. Hanna, and C. Newman, “Safe and sound? A systematic literature review of seizure detection methods for personal use,” *Seizure Eur. J. Epilepsy*, vol. 36, pp. 4–15, 2016.

- [12] A. Van de Vel *et al.*, “Non-EEG seizure detection systems and potential SUDEP prevention: State of the art: Review and update,” *Seizure*, vol. 41, pp. 141–153, 2016.
- [13] N. BH, “Non-EEG Seizure Detection Using Wearable Device,” *J. Korean Child Neurol. Soc.*, vol. 25, no. 4, pp. 209–214, Dec. 2017.
- [14] J. van Andel, R. D. Thijs, A. de Weerd, J. Arends, and F. Leijten, “Non-EEG based ambulatory seizure detection designed for home use: What is available and how will it influence epilepsy care?,” *Epilepsy Behav.*, vol. 57, pp. 82–89, 2016.
- [15] R. S. Fisher *et al.*, “Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology,” *Epilepsia*, vol. 58, no. 4, pp. 522–530, 2017.
- [16] A. Ulate-Campos, F. Coughlin, M. Gaínza-Lein, I. S. Fernández, P. L. Pearl, and T. Loddenkemper, “Automated seizure detection systems and their effectiveness for each type of seizure,” *Seizure*, vol. 40, pp. 88–101, 2016.
- [17] T. De Cooman *et al.*, “Adaptive nocturnal seizure detection using heart rate and low-complexity novelty detection,” *Seizure*, vol. 59, pp. 48–53, 2018.
- [18] P. Chapman *et al.*, “Crisp-Dm 1.0,” *Cris. Consort.*, p. 76, 2000.
- [19] A. Dalton *et al.*, “Development of a Body Sensor Network to Detect Motor Patterns of Epileptic Seizures,” *IEEE Trans. Biomed. Eng.*, vol. 59, no. 11, pp. 3204–3211, 2012.
- [20] J. Gubbi, S. Kusmakar, A. S. Rao, B. Yan, T. O’Brien, and M. Palaniswami, “Automatic Detection and Classification of Convulsive Psychogenic Nonepileptic Seizures Using a Wearable Device,” *IEEE J. Biomed. Heal. Informatics*, vol. 20, no. 4, pp. 1061–1072, 2016.
- [21] M. Velez, R. S. Fisher, V. Bartlett, and S. Le, “Tracking generalized tonic-clonic seizures with a wrist accelerometer linked to an online database,” *Seizure*, vol. 39, pp. 13–18, 2016.
- [22] E. Schulc *et al.*, “Measurement and quantification of generalized tonic-clonic seizures in epilepsy patients by means of accelerometry—An explorative study,” *Epilepsy Res.*, vol. 95, no. 1, pp. 173–183, 2011.
- [23] T. De Cooman, C. Varon, B. Hunyadi, W. Van Paesschen, L. Lagae, and S. Van Huffel, “Online Automated Seizure Detection in Temporal Lobe Epilepsy Patients Using Single-lead ECG,” *Int. J. Neural Syst.*, vol. 27, no. 07, p. 1750022, Feb. 2017.
- [24] C. Varon, A. Caicedo, K. Jansen, L. Lagae, and S. Van Huffel, “Detection of epileptic seizures from single lead ECG by means of phase rectified signal averaging,” in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 3789–3790.

- [25] W. J. C. van Elmpt, T. M. E. Nijsen, P. A. M. Griep, and J. B. A. M. Arends, "A model of heart rate changes to detect seizures in severe epilepsy," *Seizure*, vol. 15, no. 6, pp. 366–375, 2006.
- [26] D. Cogan, J. Birjandtalab, M. Nourani, J. Harvey, and V. Nagaraddi, "Multi-Biosignal Analysis for Epileptic Seizure Monitoring," *Int. J. Neural Syst.*, vol. 27, no. 01, pp. 1–18, Feb. 2017.
- [27] F. Onorati *et al.*, "Multicenter clinical assessment of improved wearable multimodal convulsive seizure detectors," *Epilepsia*, vol. 1, pp. 1–10, 2017.
- [28] M. Z. Poh *et al.*, "Convulsive seizure detection using a wrist-worn electrodermal activity and accelerometry biosensor," *Epilepsia*, vol. 53, no. 5, pp. 93–97, 2012.
- [29] MathWorks, "MATLAB," 2019. [Online]. Available: <https://la.mathworks.com/products/matlab.html>. [Accessed: 15-Nov-2019].
- [30] R Core Team, "R: A Language and Environment for Statistical Computing." Vienna, Austria, 2013.
- [31] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA Data Mining Software: An Update," *SIGKDD Explor. Newsl.*, vol. 11, no. 1, pp. 10–18, 2009.
- [32] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, p. 2825–2830, 2011.
- [33] P. Virtanen *et al.*, "SciPy 1.0--Fundamental Algorithms for Scientific Computing in Python," pp. 1–22, 2019.
- [34] I. A. W. Kotsopoulos, T. Van Merode, F. G. H. Kessels, M. C. T. F. M. De Krom, and J. A. Knottnerus, "Systematic Review and Meta-analysis of Incidence Studies of Epilepsy and Unprovoked Seizures," *Epilepsia*, vol. 43, no. 11, pp. 1402–1409, 2002.
- [35] M. J. Brodie, S. C. Schachter, and P. Kwan, *Fast facts: epilepsy*. Karger Medical and Scientific Publishers, 2012.
- [36] A. Scaramelli *et al.*, "Prodromal symptoms in epileptic patients: Clinical characterization of the pre-ictal phase," *Seizure*, vol. 18, no. 4, pp. 246–250, 2009.
- [37] W. T. Blume, H. O. Lüders, E. Mizrahi, C. Tassinari, W. Van Emde Boas, and J. Engel Jr. Ex-officio, "Glossary of Descriptive Terminology for Ictal Semiology: Report of the ILAE Task Force on Classification and Terminology," *Epilepsia*, vol. 42, no. 9, pp. 1212–1218, 2001.
- [38] H. Lüders *et al.*, "Semiological seizure classification," *Epilepsia*, vol. 39, no. 9, pp. 1006–1013, 1998.
- [39] K. Tufenkjian and H. O. Lüders, "Seizure semiology: Its value and limitations in localizing the epileptogenic zone," *J. Clin. Neurol.*, vol. 8, no. 4, pp. 243–

250, 2012.

- [40] M. Poh, S. Member, N. C. Swenson, and R. W. Picard, "A Wearable Sensor for Unobtrusive, Long-Term Assessment of Electrodermal Activity," vol. 57, no. 5, pp. 1243–1252, 2010.
- [41] T. Knoth, "Characterization and Detection of Epileptic Seizures Using Wearable Sensors," Lehrstuhl für Mustererkennung; Friedrich-Alexander-Universität Erlangen-Nürnberg, 2016.
- [42] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiol. Meas.*, vol. 28, no. 3, pp. R1--R39, Feb. 2007.
- [43] W. M. Association, "Declaration of Helsinki," vol. 79, no. October 1975, 2001.
- [44] B. E. Heldberg, "Multimodal Characterization of Epileptic Seizures," Friedrich-Alexander-Universität Erlangen-Nürnberg, 2014.
- [45] M. Garbarino, M. Lai, D. Bender, R. W. Picard, and S. Tognetti, "Empatica E3 - A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition," *Proc. 2014 4th Int. Conf. Wirel. Mob. Commun. Healthc. - "Transforming Healthc. Through Innov. Mob. Wirel. Technol. MOBIHEALTH 2014*, pp. 39–42, 2015.
- [46] D. C. Corrales, J. C. Corrales, and A. Ledezma, "How to address the data quality issues in regression models: A guided process for data cleaning," *Symmetry (Basel)*, vol. 10, no. 4, pp. 1–20, 2018.
- [47] Empatica, "E4 data - IBI expected signal," 2019. [Online]. Available: <https://support.empatica.com/hc/en-us/articles/360030058011-E4-data-IBI-expected-signal>. [Accessed: 17-Nov-2019].
- [48] J. Shukla, M. Barrera-Ángeles, J. Oliver, and D. Puig, "Efficient wavelet-based artifact removal for electrodermal activity in real-world applications," *Biomed. Signal Process. Control*, vol. 42, pp. 45–52, 2018.
- [49] M. Filetti, "Ledapy," *GitHub*, 2019. [Online]. Available: <https://github.com/HIIT/Ledapy>. [Accessed: 24-Aug-2019].
- [50] M. Benedek and C. Kaernbach, "Decomposition of skin conductance data by means of nonnegative deconvolution," *Psychophysiology*, vol. 47, no. 4, pp. 647–658, 2010.
- [51] M. E. Dawson, A. M. Schell, D. L. Filion, and G. G. Berntson, "The Electrodermal System," in *Handbook of Psychophysiology*, no. December 2013, J. T. Cacioppo, L. G. Tassinari, and G. Berntson, Eds. Cambridge: Cambridge University Press, 2000, pp. 157–181.
- [52] W. T. Blume, H. O. Lu, E. Mizrahi, C. Tassinari, W. V. E. Boas, and J. Engel, "Blume 2001 glossary of terminology for ictal semiology: Report of the ILAE Task Force on Classification and Terminology," *Epilepsia*, vol. 42, no. 9, pp. 1212–1218, 2001.

- [53] S. Vluymans, “Dealing with imbalanced and weakly labelled data in machine learning using fuzzy and rough set methods,” *Stud. Comput. Intell.*, vol. 807, pp. 1–249, 2019.
- [54] V. López, A. Fernández, S. García, V. Palade, and F. Herrera, “An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics,” *Inf. Sci. (Ny)*, vol. 250, pp. 113–141, 2013.
- [55] P. Branco, L. Torgo, and R. P. Ribeiro, “A Survey of Predictive Modeling on Imbalanced Domains,” *ACM Comput. Surv.*, vol. 49, no. 2, pp. 31:1--31:50, 2016.
- [56] V. García, J. S. Sánchez, and R. A. Mollineda, “On the effectiveness of preprocessing methods when dealing with different levels of class imbalance,” *Knowledge-Based Syst.*, vol. 25, pp. 13–21, 2012.
- [57] P. VILLAR, A. FERNÁNDEZ, R. A. CARRASCO, and F. HERRERA, “Feature selection and granularity learning in genetic fuzzy rule-based classification systems for highly imbalanced data-sets,” *Int. J. Uncertainty, Fuzziness Knowledge-Based Syst.*, vol. 20, no. 03, pp. 369–397, 2012.
- [58] V. López, A. Fernández, M. J. del Jesus, and F. Herrera, “A hierarchical genetic fuzzy system based on genetic programming for addressing classification with highly imbalanced and borderline data-sets,” *Knowledge-Based Syst.*, vol. 38, pp. 85–104, 2013.
- [59] H. He and E. A. García, “Learning from Imbalanced Data,” *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, Dec. 2009.
- [60] G. Weiss, K. McCarthy, and B. Zabar, “Cost-sensitive learning vs. sampling: Which is best for handling unbalanced classes with unequal error costs?,” *Dmin*, pp. 1–7, 2007.
- [61] M. R. Smith, T. Martinez, and C. Giraud-Carrier, “An instance level analysis of data complexity,” *Mach. Learn.*, vol. 95, no. 2, pp. 225–256, 2014.
- [62] G. Lemaître, F. Nogueira, and C. K. Aridas, “Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning,” *J. Mach. Learn. Res.*, vol. 18, pp. 1–5, 2017.
- [63] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*, 2016, vol. 19, no. 6, pp. 785–794.
- [64] R. Kohavi, “A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection,” in *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, 1995, pp. 1137–1143.
- [65] K. M. Ting, “Confusion Matrix,” in *Encyclopedia of Machine Learning and Data Mining*, C. Sammut and G. I. Webb, Eds. Boston, MA: Springer US, 2017, p. 260.

- [66] P. Baldi, S. Brunak, Y. Chauvin, C. A. F. Andersen, and H. Nielsen, "Assessing the accuracy of prediction algorithms for classification: An overview," *Bioinformatics*, vol. 16, no. 5, pp. 412–424, 2000.
- [67] S. Beniczky and P. Ryvlin, "Standards for testing and clinical validation of seizure detection devices," *Epilepsia*, vol. 59, no. October 2017, pp. 9–13, 2018.

Anexo A

Consentimiento Informado

NeuroMoTIC Sistema Móvil para el Apoyo Diagnostico de la Epilepsia

Director del proyecto: Diego M. López Gutiérrez

PROPOSITO DEL ESTUDIO:

El propósito de esta investigación es proponer y desarrollar un sistema móvil que permita la recolección, gestión y clasificación de información clínica para soportar el diagnóstico de la Epilepsia.

JUSTIFICACIÓN DEL ESTUDIO:

Se está solicitando su cooperación con esta investigación ya que si usted lo permite se le practicarán dos exámenes totalmente indoloros y de poca duración. Estos exámenes no constituyen riesgo para su salud, por el contrario permitirán hacer una valoración sobre la presencia de alteraciones neurológicas y / o mentales.

El estudio al que usted estará contribuyendo al firmar el presente consentimiento, aportará información a un equipo de investigación sobre cómo diagnosticar con mayor precisión y oportunidad sobre la presencia de Epilepsia. Si la investigación es exitosa, se espera que los médicos puedan disponer en un futuro, de mejores herramientas diagnósticas para enfrentar el diagnóstico y tratamiento de esta enfermedad.

METODOLOGÍA:

Si decido participar en el estudio, una vez haya firmado el consentimiento informado y cumpla con los criterios de inclusión, entiendo que estaré dando autorización para el manejo de los resultados sean utilizados en la presente y futuras investigaciones. Los resultados serán conocidos por usted en el momento que lo requiera; en los resultados se mantendrá siempre su confidencialidad.

BENEFICIOS DEL ESTUDIO:

Los beneficios del estudio serán la evaluación por médicos neurólogos durante el estudio. Además, se beneficiarán en un futuro todos los pacientes que padezcan de epilepsia o las secuelas derivadas de ellas, ya que podrían tener mejores tratamientos si se diagnostican a tiempo este tipo de patología.

Los datos obtenidos serán confidenciales y los que corresponden a mi participación me podrán ser revelados en caso de ser solicitados.

RIESGOS ASOCIADOS CON EL ESTUDIO:

Entiendo que mi participación en el estudio no presenta ningún tipo de riesgo.

El investigador encargado de la custodia de resultados: Dr. María Eugenia Miño Arango, Medica Neuróloga con sub-especialización en Neuropediatria. Tel. 3175138515.

RESPONSABILIDAD DE LOS PARTICIPANTES EN LAS PRUEBAS U OBSERVACIONES Y REGISTROS:

COMPENSACIÓN

Entiendo que no habrá ningún tipo de retribución económica por participar en la investigación y los resultados de la misma serán utilizados para el mejoramiento de las metodologías de los procesos enseñanza - aprendizaje

VOLUNTARIEDAD

Usted está siendo invitado a participar en este estudio de investigación. Antes de decidir si participa o no, debe conocer y comprender cada uno de los siguientes apartados. Este proceso se conoce como consentimiento informado, Siéntase con absoluta libertad para preguntar sobre cualquier aspecto que le ayude a aclarar sus dudas al respecto.

Una vez que haya comprendido el estudio y si usted desea participar, entonces se le pedirá que firme esta forma de consentimiento, de la cual se le entregará una copia firmada y fechada.

- No habrá ninguna consecuencia desfavorable usted, en caso de no aceptar la invitación.

La participación es libre y voluntaria; si decide participar en el estudio, puede retirarse en el momento que lo desee, -aun cuando el investigador responsable no se lo solicite-, informando las razones de su decisión, la cual será respetada en su integridad.

- En el transcurso del estudio usted podrá solicitar información actualizada sobre el mismo, al investigador responsable.
- Si considera que no hay dudas ni preguntas acerca de su participación, puede, si así lo desea, firmar la Carta de Consentimiento Informado que se anexa a este documento.

La información obtenida solamente será utilizada para la investigación mencionada en el presente documento y ante cualquier inquietud favor comunicarse con: (Diego M López, C.C. 76325018, Facultad de Ingeniería Electrónica y Telecomunicaciones, teléfono: 8209800 ext. 2175)

CARTA DE CONSENTIMIENTO INFORMADO

He leído y comprendido la información anterior y mis preguntas han sido respondidas de manera satisfactoria. He sido informado y entiendo que los datos obtenidos en el estudio pueden ser publicados o difundidos con fines científicos. Convengo en participar en este estudio de investigación.

Nombre:

Firma:

Documento de identificación:

Recibiré una copia firmada y fechada de esta forma de consentimiento.

Esta parte debe ser completada por el Investigador (o su representante):

He explicado al Sr(a). _____ la naturaleza y los propósitos de la investigación; le he explicado acerca de los riesgos y beneficios que implica su participación. He contestado a las preguntas en la medida de lo posible y he preguntado si tiene alguna duda. Acepto que he leído y conozco la normatividad correspondiente para realizar investigación con seres humanos y me apego a ella.

Una vez concluida la sesión de preguntas y respuestas, se procederá a firmar el presente documento.

Firma del investigador

Fecha.

Anexo B

Protocolo de recolección de señales EDA, ACM, HR y EEG

INFORMACIÓN GENERAL.

Título del protocolo: Recolección de señales de actividad electrodermal (EDA), Acelerometría (ACM), Ritmo Cardíaco (HR) y Electroencefalografía (EEG) para construcción de dataset.

Proyecto relacionado: NeuroMoTIC (Sistema móvil para el Apoyo Diagnóstico de la Epilepsia), financiado por Colciencias (ID 4427).

Investigador a cargo: Edward Alexis Molina Rivera, eamolina@unicauca.edu.co

Personal necesario: Técnico capacitado para toma de exámenes EEG.

Entidad involucrada: IPS AXON PED SAS.

OBJETIVO

Describir el procedimiento necesario para la recolección de señales EDA, ACM, HR y EEG mediante un dispositivo wearable y un electroencefalógrafo.

METODOLOGÍA

La recolección de datos se realizará en las instalaciones de la IPS AXON PED SAS sobre los pacientes que acuden a la toma de exámenes EEG y que acepten participar en el estudio, firmando el consentimiento informado para la captura de datos.

RECOLECCIÓN DE SEÑALES EEG.

La recolección de señales EEG se realizará por parte del personal técnico capacitado perteneciente a la IPS, utilizando el dispositivo BWII EEG. A continuación se describen un resumen del procedimiento de toma de exámenes EEG.

Inicialmente se registran los datos personales del paciente en el software BWAnalysis. Posteriormente se ubican 22 electrodos en el cuero cabelludo del paciente según el estándar de posicionamiento 10/20 y 2 electrodos en el pecho, verificando por medio del software que estén bien adheridos.

La grabación de las señales EEG se inicia cuando el investigador encargado lo indica (en el momento en que la pulsera inicie la grabación de señales, como se describirá más adelante).

Las señales son capturadas en tres montajes diferentes, éstos se configuran en el software con anterioridad. Cada montaje (de diez minutos de duración) determina qué electrodos se activan para la grabación de las señales. Durante el montaje 1 se realiza hiperventilación al paciente, durante el montaje 2 estimulación flash, y durante el montaje 3, apertura, cierre ocular y estimulación auditiva.

El técnico debe detener la grabación de señales EEG cuando el E4 sea apagado por el investigador encargado. Al finalizar el examen, el investigador encargado verifica que los datos sean almacenados por el software en un archivo con extensión EEG.

RECOLECCIÓN DE SEÑALES EDA, ACM Y HR.

Estas señales serán recolectadas por el investigador a cargo, utilizando el dispositivo wearable E4 Wristband.

Antes de utilizar el dispositivo por primera vez se necesita crear una cuenta de usuario en *Empatica Connect*. Además se debe instalar el software de escritorio *E4 Manager* en el computador en el que se encuentra conectado el dispositivo BWII EEG. Esto permitirá sincronizar la señal EEG con las señales provenientes del E4.

Creación de una cuenta en Empatica Connect.

Empatica Connect es una plataforma web en la cual se almacenan las sesiones de señales grabadas por el dispositivo, permitiendo la visualización y descarga de las mismas. Cada sesión se almacena con un tiempo de inicio (fecha y hora) y una duración; y contiene señales de actividad electrodermal (EDA), pulso de volumen sanguíneo (BVP), acelerometría (ACC), frecuencia cardíaca (HR) y temperatura (T).

Tanto para almacenar las señales como para sincronizar el tiempo de la pulsera es necesario crear una cuenta en *Empatica Connect*. Para hacerlo se debe acceder a la URL <https://www.empatica.com/connect/login.php#create> e ingresar los datos solicitados para la creación de una cuenta nueva. Al crear la cuenta se enviará un correo electrónico de confirmación en la dirección ingresada. Debe hacer click en el link enviado para completar el proceso de registro.

Instalación de E4 Manager.

Esta aplicación de escritorio permite cargar los datos almacenados en el E4 a *Empatica Connect*. La pulsera se conecta al *E4 Manager* utilizando la base de carga y una conexión USB.

1. Descargue el instalador de: <https://www.empatica.com/empatica-manager-download>
2. Antes de iniciar la instalación, asegúrese de que la pulsera no está conectada al equipo.
3. Al finalizar la instalación abra el *E4 Manager*. En seguida conecte el E4 al computador (ver Figura 1) y espere a que se instalen los drivers del dispositivo (ver Figura 2).
4. Ingrese los datos de inicio de sesión y haga click en *Login*. El LED de la pulsera deberá parpadear en verde y deberá ser reconocida por el *E4 Manager*. Si el software no reconoce la pulsera, desconéctela y conéctela de nuevo.

Para mayor información sobre la instalación ingrese a: <https://support.empatica.com/hc/en-us/articles/206373545-Download-and-install-the-Empatica-Manager-on-your-Windows-computer>

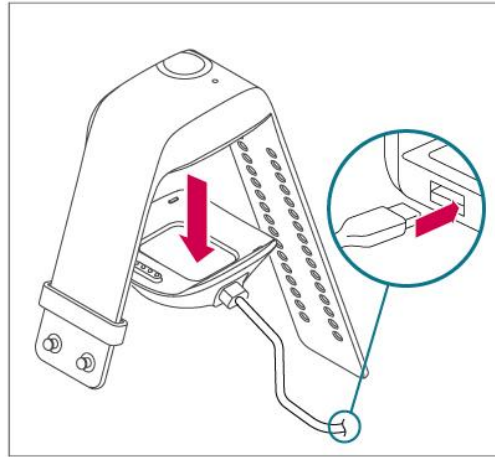


Figura 1. Conexión del E4 al computador.

Fuente: <https://support.empatica.com/hc/en-us/articles/206373545-Download-and-install-the-Empatica-Manager-on-your-Windows-computer>

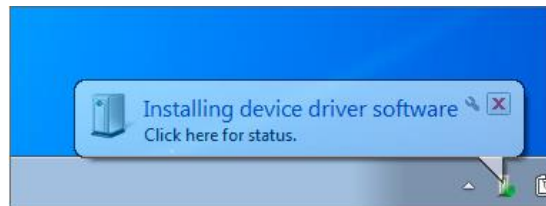


Figura 2. Instalación automática de drivers del E4.

Fuente: <https://support.empatica.com/hc/en-us/articles/206373545-Download-and-install-the-Empatica-Manager-on-your-Windows-computer>

Recolección de datos.

Antes de iniciar con los pasos siguientes se debe verificar que el dispositivo esté cargado. La única forma de observar el nivel de batería del E4 es en el modo *streaming* y por medio de la aplicación móvil *E4 Realtime*. Para acceder al modo *streaming* y observar el nivel de batería debe ubicar la pulsera en la muñeca y seguir las instrucciones en: <https://www.empatica.com/get-started-e4#Stream-your-signals-using-the-BLE-connection>.

Para cargar el dispositivo puede conectarlo a un puerto USB como se indica en la Figura 1. Cuando el dispositivo se conecta, el LED se encenderá en verde para

indicar que está conectado, después cambiará a amarillo para indicar que está cargando y una vez la carga esté completa el LED se apagará.

Sincronizar tiempo de la pulsera (*timestamp*).

Inicialmente se debe sincronizar el tiempo de la pulsera con el tiempo del computador en el que se encuentra conectado el dispositivo BWII EEG. Esto permitirá sincronizar la señal EEG con las señales provenientes del E4.

Para sincronizar la pulsera abra el *E4 Manager* e ingrese sus datos de acceso. Después conecte el E4 al computador (ver Figura 1), el cual será reconocido por el *E4 Manager*. En este momento el software sincronizará el reloj interno del E4 con el reloj de la computadora. Por último desconecte el E4 del computador.

Este procedimiento se debe hacer una vez por día, antes de comenzar la grabación de señales de pacientes.

Ubicación de la pulsera E4.

La pulsera debe ser ubicada en la muñeca del paciente justo después de que el técnico de la IPS termine de colocar todos los electrodos en el cuero cabelludo del paciente. Es sumamente importante que la pulsera haya sido sincronizada con el computador que grabará las señales EEG. En caso de que se haya conectado a un celular en modo *streaming*, estará sincronizada con el celular, lo cual es incorrecto; si este es el caso, se debe conectar de nuevo al computador para una correcta sincronización. Una vez se tenga seguridad de que el E4 está sincronizado con el computador se procede a ubicar la pulsera en la muñeca del paciente.

Coloque el sujetador de la manilla en el extremo que indica el recuadro 1 de la Figura 3. Después ubique el extremo de los electrodos de la pulsera, en la muñeca del brazo no dominante¹ del paciente, tal como se muestra en el recuadro 2 de la Figura 3.

¹ Aunque algunos eventos neurológicos como crisis epilépticas pueden ocurrir solo en un lado. Para más información ver: Picard, R. W., Fedor, S., & Ayzenberg Y., "Multiple Arousal Theory and Daily Life Electrodermal Activity Asymmetry" *Emotion Review*, March 2015.

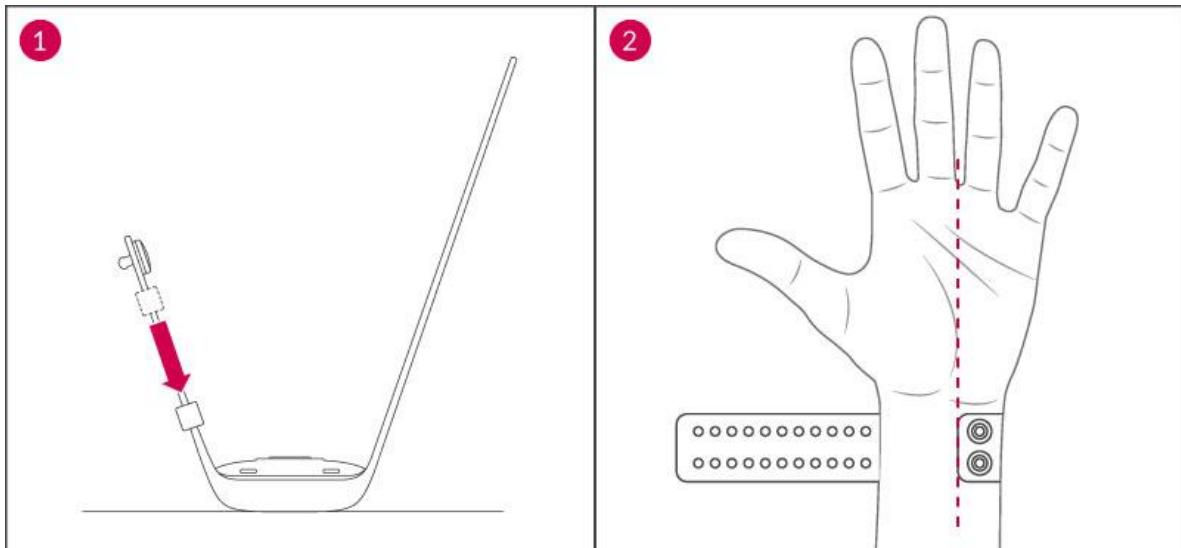


Figura 3. Ubicación de la pulsera E4 en la muñeca del paciente (1).

Fuente: <https://www.empatica.com/get-started-e4>

Por último asegure la pulsera a la muñeca como se indica en la Figura 4, de tal forma que los electrodos no cambien de posición a causa del movimiento corporal. Se debe tener cuidado de no ajustar la manilla tanto como para causar molestia en el paciente o restringir el flujo sanguíneo.

La pulsera está bien ubicada en la muñeca si no se puede observar ninguna luz que se escape del sensor PPG, el cual está ubicado debajo de la pulsera.

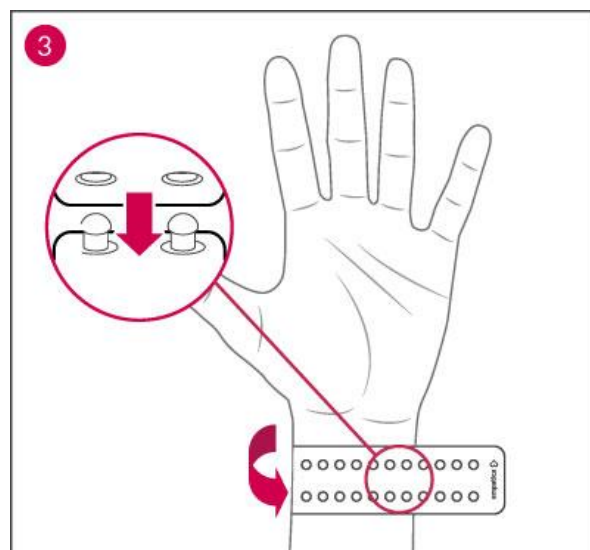


Figura 4. Ubicación de la pulsera E4 en la muñeca del paciente (2).

Fuente: <https://www.empatica.com/get-started-e4>

Al cambiar de usuario se deben limpiar todas las partes del dispositivo que estén en contacto con la piel. La desinfección se puede realizar mediante exposición a lámpara UV o mediante limpieza con algodón humedecido con una pequeña cantidad de alcohol etílico o isopropílico.

Grabación de señales.

Las señales serán almacenadas en la memoria interna del dispositivo. Ésta permite almacenar hasta 60 horas de datos. Los datos permanecen en la memoria hasta que el dispositivo es conectado al software *E4 Manager*, el cual transfiere las sesiones almacenadas a *Empatica Connect*, limpia la memoria interna y sincroniza el reloj.

Una vez el E4 esté ubicado correctamente en la muñeca del paciente, encienda la pulsera presionando el botón de la misma durante dos segundos. El LED comenzará a parpadear en verde durante 60 segundos, después de los cuales pasará a parpadear en rojo durante algunos segundos, indicando que la grabación de señales ha iniciado (ver Figura 5). Finalmente el LED se apagará para ahorrar energía.

En el momento que inicie la grabación de señales en el E4 (inicio del parpadeo rojo en el LED) el investigador encargado pedirá al técnico de la IPS que inicie la grabación de señales EEG en el software conectado al dispositivo BWII EEG. Esto permitirá que todas las señales se registren de forma paralela. Las señales del BWII EEG y del E4 estarán sincronizadas ya que los dos dispositivos utilizan como marca de tiempo (*timestamp*) el reloj del mismo computador.



Figura 5. Ubicación de la pulsera E4 en la muñeca del paciente (2).

Fuente: <https://www.empatica.com/get-started-e4>

Para marcar un evento en cualquier momento de la grabación puede presionar el botón una vez, al hacerlo el LED se encenderá en rojo durante un segundo. Tenga cuidado de presionar rápidamente el botón, para evitar que el dispositivo se apague o se reinicie.

Para finalizar la sesión presione durante dos segundos el botón. Cuando esto pasa, el dispositivo guarda la sesión en un archivo y se apaga. Como indicador, el LED se encenderá en rojo durante un segundo. Puede verificar si la sesión ha finalizado observando con cuidado el sensor PPG debajo de la pulsera, el cual deberá estar apagado y sin emitir luz. Cuando se apague el dispositivo, el investigador a cargo pide al técnico de la IPS que termine la grabación de señales EEG.

Almacenamiento de sesiones.

Al finalizar la toma de señales del último paciente, abra el *E4 Manager* e ingrese sus datos de acceso. Después conecte el E4 al computador (ver Figura 1), el cual será reconocido por el *E4 Manager*. En este momento el software transfiere las sesiones almacenadas a *Empatica Connect* y limpia la memoria interna del E4.

Este procedimiento se debe hacer una vez por día, al finalizar todas las grabaciones de pacientes.

Visualización y descarga de sesiones almacenadas.

Para listar las sesiones almacenadas inicie sesión en *Empatica Connect* (<https://www.empatica.com/connect/login.php>) y haga click en *Sessions*. Podrá observar la interfaz mostrada en la Figura 6.

Desde esta interfaz podrá visualizar o descargar los datos de las sesiones. Como se mencionó anteriormente, cada sesión se almacena con un tiempo de inicio (fecha y hora) y una duración; y contiene señales de actividad electrodermal (EDA), pulso de volumen sanguíneo (BVP), acelerometría (ACC), frecuencia cardíaca (HR) y temperatura (T). La Figura 7 muestra la visualización de las tres primeras señales de una sesión.

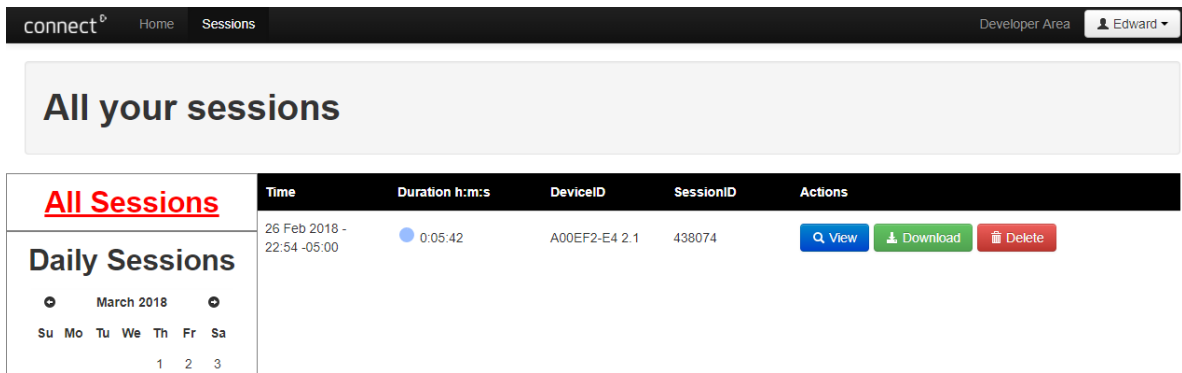


Figura 6. Listado de sesiones almacenadas.
Fuente propia.

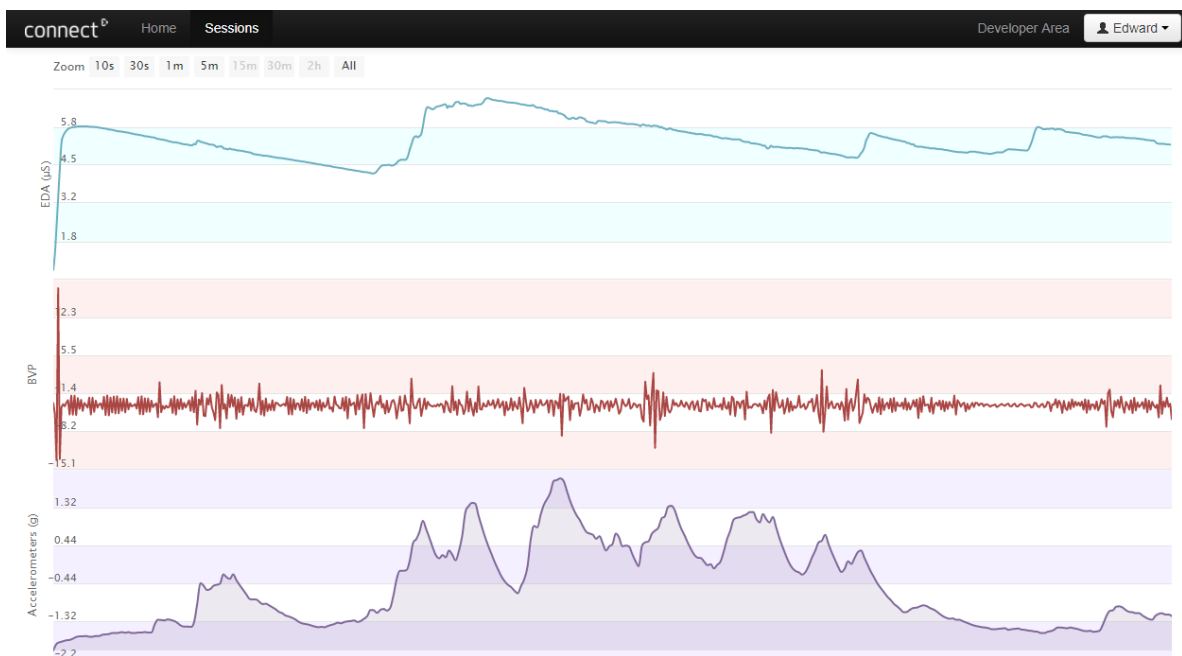


Figura 7. Visualización de una sesión en Empatica Connect.
Fuente propia.

CONSIDERACIONES DE SEGURIDAD²

1. Fugas de corriente a través de los electrodos EDA: el circuito EDA impulsa una pequeña cantidad de corriente iónica a través de la piel para operar: Máx. 100 μ A según IEC 60601-1:2005. Nivel de riesgo muy bajo, sin daños, el E4

² Tomado del manual del fabricante: <https://empatica.app.box.com/v/E4-User-Manual>

cumple con la ISO 60601-1:2005, lo que significa que la cantidad de corriente producida es segura para el usuario.

2. Exposición de los ojos a la luz LED: los LED del sensor PPG son extremadamente brillantes, normalmente están totalmente oscurecidos por el contacto con la piel. Sin embargo, el usuario puede quitar el E4 de la muñeca, exponiendo la fuente de luz. Nivel de riesgo muy bajo, sin daños, la cantidad de luz es comparable con otras fuentes de luz a las que los usuarios pueden estar expuestos en su entorno diario. Se aconseja al usuario que no mire los LED del sensor PPG durante un tiempo prolongado.
3. Compatibilidad electromagnética: Nivel de riesgo muy bajo, sin daños. El E4 es seguro para su uso con otros dispositivos electrónicos. Cumple con los estándares IEC 60601-1-2:2007 (tercera edición), FCCC FR47Part15bedIC (Industry Canada) y ETSIEN301489-1e3V. Equipos de comunicaciones RF portátiles y móviles pueden afectar el comportamiento del E4. La pulsera necesita precauciones especiales con respecto a EMC (compatibilidad electromagnética) y debe instalarse de acuerdo con la información de EMC.

RESULTADOS ESPERADOS

1. Cien (100) archivos de grabaciones EEG realizadas a 100 pacientes con el dispositivo BWII EEG.
2. Cien (100) sesiones realizadas a 100 pacientes con el dispositivo E4 Wristband. Cada sesión con las siguientes señales capturadas: actividad electrodermal (EDA), pulso de volumen sanguíneo (BVP), acelerometría (ACC), frecuencia cardíaca (HR) y temperatura (T).

CONSENTIMIENTO INFORMADO

El Consentimiento informado se encuentra anexo al presente documento.