



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

**CLASIFICACIÓN AUTOMATIZADA DE ACTIVIDAD CEREBRAL
ANORMAL EN PACIENTES NEUROCRÍTICOS PARA MEJORAR
CAPACIDAD DIAGNÓSTICA**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

JAVIER IGNACIO VALDÉS ESPINOZA

PROFESOR GUÍA:
JUAN VELÁSQUEZ SILVA

MIEMBROS DE LA COMISIÓN:
ENZO BRUNETTI FONSECA
ROCÍO RUIZ MORENO

SANTIAGO DE CHILE
2022

CLASIFICACIÓN AUTOMATIZADA DE ACTIVIDAD CEREBRAL ANORMAL EN PACIENTES NEUROCRÍTICOS PARA MEJORAR CAPACIDAD DIAGNÓSTICA

El Estado Epiléptico (EE) es una condición médica grave que implica riesgo vital y deterioro neurológico funcional, por lo que su pronto reconocimiento y tratamiento es indispensable. Tiene una incidencia de alrededor de 56 casos por cada 100.000 adultos al año, y puede presentarse tanto en pacientes con antecedentes de epilepsia como sin ella. El EE puede dividirse en formas convulsivas y no convulsivas. Se estima que esta última puede llegar a incidir en la muerte de hasta 3.000 personas cada año en Chile, y que solo en pérdida de productividad y costos de atención médica directa, implica un perjuicio económico de 14.400 millones de pesos al año, asumidos por las personas o por el estado, considerando las prestaciones que deben efectuarse tanto en salud pública como privada.

El pronóstico del EE y la eficacia del tratamiento dependen directamente de la precocidad de su inicio. Por esta razón, la detección temprana resulta particularmente crítica para maximizar las oportunidades de intervenciones exitosas, que deriven en aumento de la sobrevivencia y disminución de las secuelas permanentes. Por tanto, en el presente trabajo se desarrolló un modelo clasificador de actividad epileptiforme mediante análisis de la actividad cerebral utilizando algoritmos de Machine Learning, con la finalidad de mejorar la capacidad de diagnóstico y así propiciar tratamiento temprano y mejora en el pronóstico de los pacientes del Instituto de Neurocirugía Asenjo (INCA). Para ello, se evaluó distintos modelos de clasificación de actividad epileptiforme.

El modelo que obtuvo los mejores resultados con los registros disponibles es capaz de detectar crisis epilépticas en hasta un 87 % de los casos. Dicho resultado se presenta al nivel de las alternativas comerciales estudiadas. A partir de ello, se realiza un prototipo de aplicación, capaz de hacer visibles las potencialidades del uso de un sistema de clasificación automática.

A través de una evaluación de impacto social y económico se establece que el desarrollo de una aplicación con las características descritas sería beneficiosa tanto social como económicamente para los pacientes, para los equipos médicos, las instituciones de salud y gobierno. La implementación de un sistema capaz de realizar análisis automático para clasificar o detectar actividad cerebral anormal, y segmentación del registro continuo en ventanas de interés, puede permitir reducir sustancialmente los tiempos de estudio y diagnóstico de registro electroencefalográfico, lo que en trabajos futuros permitirá ampliar la cobertura de seguimiento a un mayor número de camas a nivel local, así como la posibilidad de realizar un monitoreo neurológico remoto y continuo en UCI a nivel nacional.

Agradecimientos

A mis papás Alejandra y Ricardo que me lo han dado todo. Su amor, su apoyo incondicional, su esfuerzo y su dedicación han sido una inspiración para mí y una fuente de motivación para mi vida y mi curso universitario.

A mi hermana Paulina por su compañía, nuestras risas, su amistad y por permitirme ver el mundo desde otros ojos.

A mis abuelos Gregoria y Miguel por siempre creer en mí y abrazarme con cada acción.

A quienes aportaron a esta investigación desde el WIC. Agradezco a Juan por su plena confianza y su guía. Especiales gracias también a Rocío por sus consejos, su compañía y ayuda durante el proceso. A Enzo por el conocimiento médico y técnico aportado al trabajo.

A quienes hicieron de la Universidad un espacio de disfrute, crecimiento y aprendizajes más allá de lo académico. Especialmente a mis “personas favoritas de Santiago”: Javi Cembrano, Barbi y Javi Serrano. A los que me regalaron su compañía, su cariño y sus conversas, como la Feña, el Nico, el Jose, los chiquillos de la fanfarria, del preu y el CEIN.

A quienes después de tantos años siguen presentes en cada momento de mi vida: Bastián Díaz, Bastián Venegas, Benjamín, Ignacio, David, Pablo, Álvaro, el resto de la Filial y los #lolein.

A todos en quienes encontré un hogar lejos de casa, especialmente a Camilo, mi tía Teresa, Victor, Emma, Seba y mis chiquilles de Porto.

A mi familia, tías, tíos, primas y primos; compañeras, compañeros, amigas y amigos; a todos quienes sin ser nombrados han estado ahí para mí, tanto en este trabajo como a lo largo de estos años.

Gracias de todo corazón.

Tabla de Contenido

1. Introducción	1
1.1. Antecedentes Generales	1
1.2. Características institucionales	4
1.2.1. Web Intelligence Centre (WIC)	4
1.2.2. Instituto de Neurocirugía Dr. Alfonso Asenjo (INCA)	5
1.3. Descripción del problema	5
1.4. Hipótesis de investigación	7
1.5. Objetivos	8
1.5.1. Objetivo general	8
1.5.2. Objetivos específicos	8
1.6. Marco conceptual	8
1.6.1. Electroencefalografía	8
1.6.2. Método KDD - Descubrimiento de conocimiento en bases de datos . .	10
1.6.3. Minería de datos	12
1.6.4. Aprendizaje automático	12
1.6.5. Support Vector Machines (SVM)	13
1.6.6. Redes neuronales artificiales (ANN)	14
1.6.7. Redes neuronales convolucionales (CNN)	15
1.6.8. Evaluación de modelos	17
1.7. Metodología	18
1.8. Resultados esperados	19
1.8.1. Alcances y riesgos	20
2. Estado del arte	21
2.1. Temko, Thomas, et al. (2011)	21
2.2. Ullah, Hussain, et al. (2018)	23
2.3. Golmohammadi, et al. (2019)	24
2.4. Zhang, et al (2016)	26
2.5. Song, et al. (2012)	27
2.6. Gómez, et al. (2020)	29
2.7. Struck, et al. (2017)	31
2.8. Soluciones comerciales	32
2.9. Resumen y conclusiones	34
3. Análisis de datos	37
3.1. Set de datos	37
3.2. Exploración de registros	38

3.3. Modelos de clasificación	39
3.3.1. Support Vector Machine (SVM)	39
3.3.2. Red Neuronal Convolutacional de una dimensión (CNN-1D)	40
3.3.3. Red Neuronal Recurrente (RNN)	40
3.3.4. Red Neuronal Convolutacional de dos dimensiones (CNN-2D)	41
3.4. Resultados y discusión	42
4. Prototipo expositivo	44
4.1. Esquema propuesto	44
4.2. Mockup propuesto	45
4.3. Prototipo desarrollado	47
4.4. Impacto económico y social	48
5. Conclusiones	51
5.1. Conclusiones generales	51
5.2. Trabajos futuros	53
Bibliografía	54
Anexo A. Capítulo 1	58
A.1. Criterios para crisis no convulsivas	58

Índice de Tablas

1.1.	Posibles combinaciones de resultados de una clasificación binaria.	17
2.1.	Comparativa puntuación de riesgo optimizada.	32
2.2.	Comparativa de alternativas comerciales. Con datos del trabajo de Koren et al.[26]	33
2.3.	Comparativa de los principales resultados de los trabajos estudiados.	35
3.1.	Estructura de la red neuronal utilizada.	41
3.2.	Comparativa de los resultados de los modelos aplicados. Remuestreo hecho a través de interpolación lineal.	42
3.3.	Comparativa de los resultados de los modelos aplicados. Remuestreo hecho a través de Fourier.	43
3.4.	Comparativa de los resultados de los modelos aplicados. La señal está procesada usando la transformada rápida de Fourier.	43
4.1.	Comparación de costos estimados por país.	49
4.2.	Comparación de PIB per cápita al año 2020, según datos de las cuentas nacionales del Banco Mundial y archivos de datos de las cuentas nacionales de la OCDE.	49
4.3.	Comparación de efectividad en tratamiento de Estado Epiléptico. Fuente: Treiman & Walker [47]	50

Índice de Ilustraciones

1.1.	Etapas del EE según la definición de estado epiléptico de la ILAE. Basado en el diagrama de LaRoche & Haider [27]	2
1.2.	Electroencefalografía de EE no convulsivo. Actividad epileptiforme continua tipo espiga-onda ampliamente distribuidas sobre el hemisferio izquierdo. Fuente: Lawson & Quintana [28]	3
1.3.	Resumen de causas y consecuencias del problema identificado en un árbol de problemas.	6
1.4.	Sistema Internacional 10-20 para la colocación de los electrodos extracraneales. Fuente: Rojas et al. [38].	9
1.5.	Descargas generalizadas de tipo espiga-onda de 3 Hz en un electroencefalograma.	10
1.6.	El proceso de Knowledge Discovery in Databases. Basado en esquema de Maimon & Rockach [29].	11
1.7.	Representación de un hiperplano separando categorías a través de SVM.	13
1.8.	Esquema de una neurona artificial. Basado en esquema de Dongare, Kharde & Kachare [11].	14
1.9.	Esquema de una red neuronal artificial. Basado en esquema de Dongare, Kharde & Kachare [11].	15
1.10.	Efectos de diferentes filtros o matrices de convolución sobre una imagen. Fuente: Albawi, Mohammed & Al-Zawi. [1].	16
1.11.	Transformación de una imagen a través de max-pooling de 2x2. Fuente: Albawi, Mohammed & Al-Zawi. [1].	17
2.1.	Ejemplos de la representación visual de la señal tomando la amplitud como atributo del píxel para cuatro pacientes del conjunto de datos del CHB-MIT. En vertical se codifica los canales en configuración bipolar, y en horizontal la longitud de la ventana (4s) con la relación de aspecto ajustada para su visualización. La barra de color muestra el rango de amplitud de las señales en μV . Fila superior: ejemplos ictales, fila inferior: ejemplos interictales. Fuente: Gómez et al.[17].	30
2.2.	Software comparado por Koren et al. [26]	33
3.1.	Comparativa entre variables derivadas de cada ventana temporal. Elaboración propia.	39
3.2.	Esquema de las capas convolucionales del modelo CNN. Elaboración propia.	41
4.1.	Esquema propuesto para el prototipo a realizar.	44
4.2.	Mockup propuesto - Vista de diagnóstico antes de cargar registro EEG.	45
4.3.	Mockup propuesto - Vista de diagnóstico luego de cargar registro EEG.	46
4.4.	Esquema propuesto para el prototipo a realizar.	47
4.5.	Ejemplo del funcionamiento de la plataforma.	48

Capítulo 1

Introducción

1.1. Antecedentes Generales

El Estado Epiléptico (EE)¹ es una condición médica grave que implica riesgo vital y deterioro neurológico funcional para quien la sufre [13]. Ocurre cuando el paciente presenta crisis epilépticas² lo suficientemente prolongadas, o repetidas crisis en un intervalo breve de tiempo [20]. Una crisis epiléptica es un episodio transitorio de disfunción cerebral, durante el que se producen descargas eléctricas anormales y/o caóticas en la corteza cerebral, debidas a actividad neuronal excesiva o síncrona anormal, que afectan los mecanismos de homeostasis (o auto-regulación) en el cuerpo. Las crisis son generalmente breves y, por definición son auto-limitadas [37].

Inicialmente, la Liga Internacional contra la Epilepsia (ILAE) y la Epilepsy Foundation of America definieron esta condición como “crisis que persisten por más de 30 minutos, o una serie de crisis repetidas sin recuperación de la conciencia entre ellas en un periodo de 30 minutos”[27]. Esto se basa en estudios que demuestran un elevado riesgo de daño neuronal luego de dicho período de tiempo, cuando los mecanismos de auto-regulación del cuerpo fallan [34]. Lo anterior implica que la necesidad de un diagnóstico rápido de la condición es indispensable para su tratamiento y una posterior evolución favorable del paciente. Es por ello que para su manejo se han definido criterios más acotados. Se considera que un paciente se encuentra en Estado Epiléptico si presenta una crisis de más de 5 minutos de duración, o si tiene más de una crisis sin recuperación neurológica entre ellas, período luego del cual se aconseja iniciar tratamiento farmacológico. Más específicamente, el año 2015 la ILAE propone una definición del Estado Epiléptico, esquematizada en la Figura 1.1, que señala que el EE “es una condición que resulta de la falla de los mecanismos responsables de la terminación de una crisis, o del inicio de mecanismos que conducen a crisis anormalmente largas (después del punto de tiempo t_1). La afección puede tener consecuencias a largo plazo (después del punto t_2) según el tipo y duración de la o las crisis” [48].

¹ También llamado Status Epilepticus (SE) o Status Epiléptico.

² En las que se observa actividad epileptiforme en el cerebro.

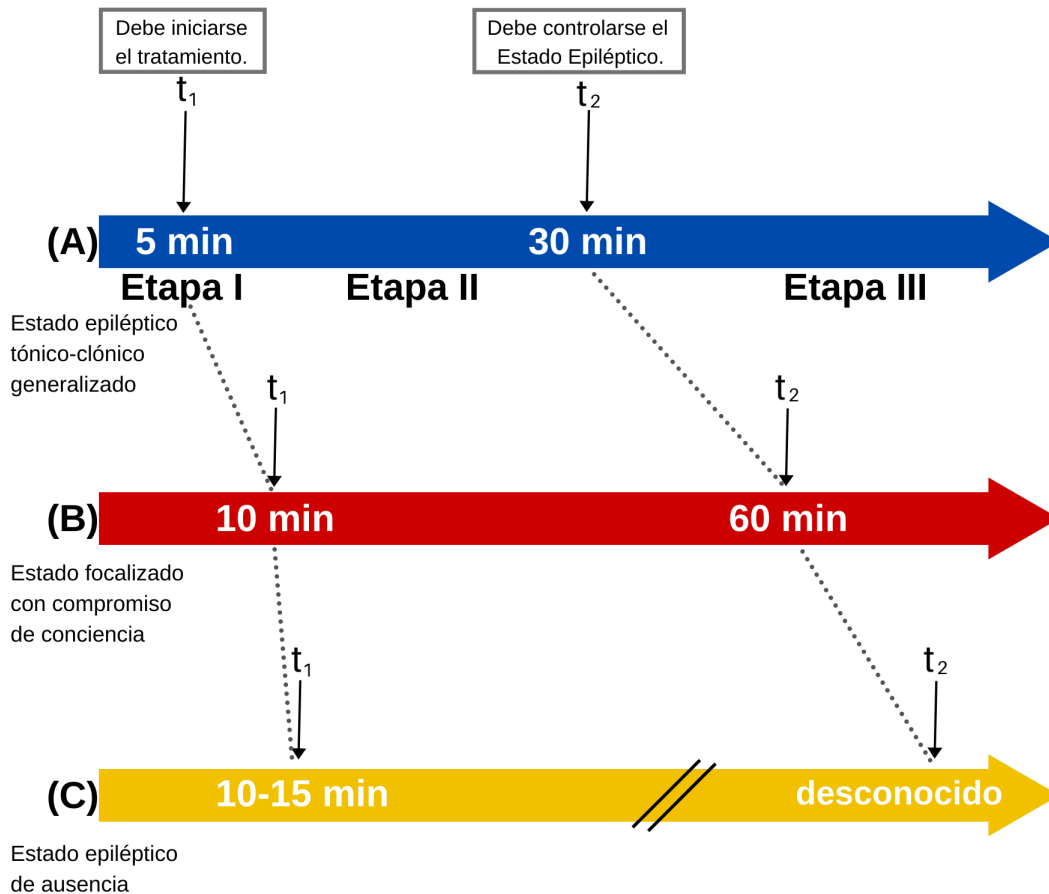


Figura 1.1: Etapas del EE según la definición de estado epiléptico de la ILAE. Basado en el diagrama de LaRoche & Haider [27]

En la Figura 1.1 se describen las etapas de la condición según el tipo y duración de la crisis:

- El momento t_1 se refiere al punto en el que una crisis debe ser considerada prolongada y es poco probable que termine sin intervención.
- El momento t_2 se refiere al punto a partir del cual existe riesgo de consecuencias a largo plazo.

El cuadro puede presentarse tanto en pacientes que cuenten con historial clínico de epilepsia, como en los que no, siendo sintomático de una amplia gama de causas subyacentes, entre las que se incluyen traumatismos, infecciones o accidentes cerebrovasculares, entre otros. El EE puede además presentarse en formas diversas, agrupadas en dos tipos principales: las crisis con síntomas motores prominentes y las crisis sin síntomas motores prominentes. El primero de ellos considera el Estado Epiléptico Convulsivo (EEC)³, el que puede ser diagnosticado por su presentación clínica, sin requerir necesariamente de electroencefalografía (EEG). El segundo tipo recibe la denominación de Estado Epiléptico No Convulsivo (EENC)⁴

³ En inglés llamado Convulsive Status Epilepticus (CSE).

⁴ En inglés llamado Nonconvulsive Status Epilepticus (NCSE).



Figura 1.2: Electroencefalografía de EE no convulsivo. Actividad epileptiforme continua tipo espiga-onda ampliamente distribuidas sobre el hemisferio izquierdo. Fuente: Lawson & Quintana [28]

En las unidades de pacientes críticos, el EENC se presenta a menudo como una disminución del nivel de conciencia, que puede llegar hasta el coma. Dado que otras afecciones que se observan comúnmente en unidades de pacientes críticos también pueden presentarse con una disminución del nivel de conciencia, se requiere de electroencefalografía continua (cEEG), como se muestra en la Figura 1.2, para un diagnóstico definitivo.

Uno de los estudios epidemiológicos prospectivos clásicos realizado en Estados Unidos, que incluye adultos y niños, describe una incidencia de 40 a 80 casos por cada 100.000 personas al año [7]. A nivel latinoamericano, según un estudio realizado en Argentina, la incidencia⁵ del Estado Epiléptico es de 56,6 casos por cada 100.000 adultos, en donde cerca de la mitad representan casos no convulsivos [39]. Esta última condición muestra una mortalidad de entre un 33 % a un 57 %, y las variables significativas asociadas a un aumento de la mortalidad son la duración de la crisis y la demora en el diagnóstico [52].

Considerando que la población estimada de Chile por el Instituto Nacional de Estadísticas (INE) es de 19.458.310 personas⁶, se pueden estimar 10.994 casos de Estado Epiléptico al año. Con eso, y considerando las proporciones planteadas en el estudio de Romano [39], los casos de EENC en Chile serían alrededor de 5.277 al año. Todo lo anterior implica que esta condición podría llegar a incidir en la muerte de alrededor de 3.008 personas cada año en Chile.

⁵ Cabe decir que hay estudios que indican que la prevalencia del EENC está subestimada por su complejidad diagnóstica [6].

⁶ Proyección al 30 de junio de 2020 [24].

Por otra parte, el año 2013 [2] se estimaba que el costo de tratamiento en una unidad de cuidados intensivos para pacientes neurológicos era en promedio \$3,4 millones por paciente. Dicho valor considera equipamiento, medicamentos, equipos médicos involucrados, duración promedio de estadía en UCI neurológica, entre otros. Considerando la incidencia estimada en Chile anualmente, solo los costos directos de la condición en estudio ascienden a por lo menos \$18 mil millones al año. Por otra parte, deben ser considerados los costos indirectos asociados a mayores tiempos de recuperación de los pacientes, la pérdida de capacidad productiva por invalidez, los costos económicos y sociales propios de la muerte, y a los potenciales pacientes no tratados por el uso de los distintos recursos médicos.

1.2. Características institucionales

El presente trabajo se enmarca en un proyecto de colaboración entre 2 instituciones: el Web Intelligence Centre (WIC) y el Instituto de Neurocirugía Dr. Alfonso Asenjo (INCA).

1.2.1. Web Intelligence Centre (WIC)

El Web Intelligence Centre es un centro de investigación perteneciente al Departamento de Ingeniería Industrial (DII) de la Facultad de Ciencias Físicas y Matemáticas (FCFM) de la Universidad de Chile. Desarrolla investigación en tecnologías de la información y ofrece servicios en áreas tales como analítica y ciencias de los datos, inteligencia de negocios, desarrollo de software y consultoría de inteligencia artificial, y arquitectura e ingeniería de datos [4].

El WIC cuenta con un equipo de 21 personas y se encuentra liderado por su directora ejecutiva Rocío Ruiz, el fundador y director académico Juan Velásquez, y el Director de Tecnología Felipe Vera. Entre los profesionales que pertenecen al centro se encuentran ingenieros de proyectos, diseñadores, data scientists, entre otros. Además, cuenta con la participación de memoristas de ingeniería en sus proyectos.

Declaraba su misión, visión y objetivos en su página web [3] del siguiente modo:

Misión: Desarrollar investigación de frontera en el campo de Tecnologías de Información creando nuevas soluciones para abordar problemas complejos de ingeniería utilizando herramientas basadas en la Web de las Cosas.

Visión: Ser un líder a nivel internacional en la investigación de tecnologías de información y comunicaciones aplicadas a la resolución de problemas del mundo real.

Objetivos:

- Publicar en las principales revistas, conferencias y editoriales relacionadas con Web Intelligence.
- Proveer a servicio profesional, excelente y rápido para todos nuestros clientes.
- Dictar cursos de orientación práctica acerca de las Tecnologías de Información y su aplicación en los negocios.

Parte importante de sus proyectos se desarrolla en el ámbito médico, siendo este uno de sus focos de investigación, tal como se declara en su sitio web: *“Queremos ser un actor relevante en el área del data science y sus aplicaciones en la salud. Para ello contaremos al 2022 con al menos 3 proyectos transferidos en distintos centros de salud”* [4]. Para ese objetivo, cuenta dentro del equipo con la participación de la doctora Flavia Guñazú, quien asesora y lidera el desarrollo de proyectos desde la perspectiva médica.

1.2.2. Instituto de Neurocirugía Dr. Alfonso Asenjo (INCA)

El Instituto de Neurocirugía Dr. Alfonso Asenjo (INCA) es una institución pública, centro de referencia de patología neuroquirúrgica a nivel país. Es dependiente del Servicio de Salud Metropolitano Oriente de Santiago y presta servicios programados y de urgencia a la Región Metropolitana y a las regiones del resto del país, en neurocirugía de mediana y alta complejidad [8].

El servicio de pacientes críticos de la institución (SPC), cuenta con una UCI de 10 camas, y una UTI de 8 camas. El INCA posee tecnología de vanguardia en términos de equipamiento para la evaluación diagnóstica y terapéutica de patología neurológica y neuroquirúrgica, incluyendo resonancia magnética de alto teslaje, scanner y angiografo digital. Su SPC se encuentra en importante proceso de expansión y crecimiento, logrando actualmente estar dotado por un equipo de residentes neurólogos 24/7 para el manejo agudo de patología intensiva.

La relevancia del INCA para el presente trabajo es su capacidad de monitoreo en el servicio de pacientes críticos, cuyos datos serán la base del estudio. Se cuenta con la participación del doctor Enzo Brunetti, neurólogo, especialista en epilepsia y trastornos asociados.

1.3. Descripción del problema

El pronóstico del EE y la eficacia del tratamiento dependen directamente de la precocidad de su inicio. Por esta razón, la detección temprana resulta particularmente crítica para maximizar las oportunidades de intervenciones exitosas, que deriven en aumento de la sobrevida y disminución de las secuelas permanentes.

En la práctica clínica, el diagnóstico del EENC en pacientes críticos requiere del análisis visual de los registros electroencefalográficos (EEG), realizados por neurofisiólogos experimentados, quienes identifican patrones característicos de las actividades anormales. Dada la naturaleza espontánea de las crisis epilépticas, se necesita hacer uso de electroencefalografía continua (cEEG)[19] a largo plazo de los pacientes, con duraciones que van desde horas a días, ya que su objetivo principal es registrar las convulsiones que ocurren sin previo aviso. La cEEG produce volúmenes muy grandes de registro electroencefalográfico y, en muchos casos, video. Debido a esto, las señales de EEG resultantes pueden ser extremadamente extensas, lo que hace que su análisis sea un proceso muy costoso, ineficiente y que requiere mucho tiempo. Idealmente, estos datos se revisarían en tiempo real para proporcionar retroalimentación inmediata al equipo médico que atiende al paciente. Sin embargo, esto generalmente no es factible, ya que en la mayoría de las situaciones, la interpretación adecuada requiere revisión de un neurofisiólogo capacitado. Es por ello, que generalmente los datos de cEEG se suelen interpretar de forma intermitente.

Por ejemplo, según señala Ebersole [12], en su institución los datos del cEEG se revisan en el plazo de una hora desde el inicio del estudio. Después de esto, se revisa con un mínimo de intervalos de 6 horas durante el día y de 8 horas por la noche. Si alguna de las revisiones revela crisis, actividad epileptiforme o cambios significativos de revisiones anteriores, el EEG se revisa con más frecuencia, generalmente en intervalos de 1 a 2 horas. Debido a que muchas crisis son no convulsivas o sutiles, el registro completo debe examinarse revisando el EEG en bruto completo o los segmentos seleccionados según el algún software de detección. La revisión del EEG bruto completo puede ser un proceso que requiere mucho tiempo, requiriendo un mínimo de 30 minutos para revisar 24 horas de datos (10 segundos de registro por página, a una frecuencia de 5 páginas por minuto [12]), suponiendo que no hay eventos o anomalías significativas que requieran una revisión más detallada. Un examen más detenido del EEG o del video durante las convulsiones u otros eventos de interés aumentará, por tanto, significativamente el tiempo requerido para la revisión.

Debido a la escasez del recurso humano para la interpretación del cEEG, y a la presión impuesta por el registro muchas veces simultáneo de varios pacientes, el trazado es habitualmente analizado de manera visual a una alta velocidad. En este contexto, la sensibilidad y especificidad del reporte son altamente variables, particularmente entre neurofisiólogos no entrenados. A ello se suma el hecho que muchos de los eventos de actividad anormal son sutiles y de difícil interpretación visual.

Es por los motivos antes mencionados que el problema identificado es que *Existe una capacidad de diagnóstico del EENC limitada*.

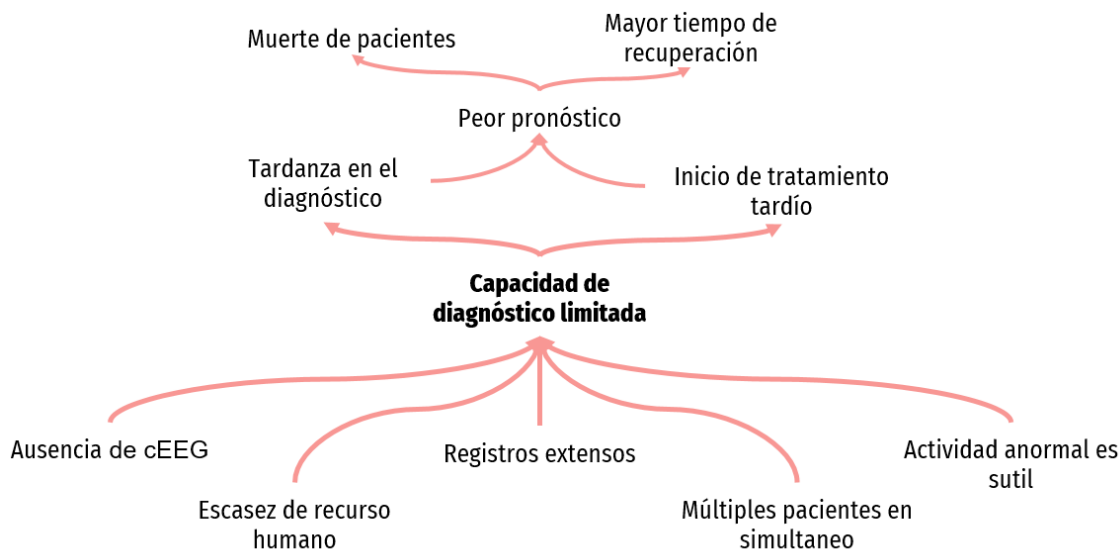


Figura 1.3: Resumen de causas y consecuencias del problema identificado en un árbol de problemas.

Independiente de cual sea en específico el cuello de botella que genera esta limitación en la capacidad diagnóstica, tiene como consecuencia una tardanza en la realización de un diagnóstico que pueda dar pie al inicio de un tratamiento oportuno. La demora en la identificación de una crisis epiléptica, y su consecuente tratamiento tardío afecta gravemente el pronóstico del paciente, considerando que esas son precisamente las variables más significativas asociadas a un aumento de la mortalidad [52]. En caso de que el daño producido por el evento no sea fatal, las expectativas de recuperación y el tiempo que tomará este proceso se ven afectadas negativamente. Una tardanza en la recuperación a su vez implica un mayor uso de recursos médicos, tales como el uso de camas de unidades de cuidados intensivos, personal médico, entre otros. Esto conduce a un aumento del costo del tratamiento del paciente y a externalidades negativas, asociadas a la escasez de recursos médicos, que pueden terminar afectando a terceros.

Ante lo expuesto, el análisis automatizado de los registros de EEG se presenta como una alternativa conveniente, ya sea seleccionando ventanas de interés de la señal para su revisión detallada, o ejecutando directamente una clasificación que permita identificar cuándo se está ante un episodio de crisis. A lo largo de los años se han explorado diversos métodos con ese enfoque[16]. Entre ellos se incluyen técnicas de procesamiento de señales digitales en los dominios tiempo-frecuencia[35] [18], análisis dinámico no lineal del EEG[43], sistemas expertos que intentan imitar a un observador humano[10], entre otros. Por otra parte, además, hoy en día existe alternativas en el mercado que buscan resolver la problemática planteada. Entre ellas se pueden encontrar los desarrollos de Besa⁷, que utiliza técnicas analíticas sobre el espectro de potencia de la señal; Encevis⁸, que hace análisis en los dominios de tiempo y frecuencia; y Persyst⁹, que utiliza un enfoque basado en redes neuronales. A pesar del desarrollo de este campo, el traspaso de estas metodologías a entornos clínicos como las UCI, ha sido limitado, principalmente debido a las altas tasas de falsas detecciones[22].

1.4. Hipótesis de investigación

Dados los antecedentes presentados, las características institucionales y los fundamentos del problema expuesto, se plantea la siguiente hipótesis de investigación:

Es posible optimizar la gestión diagnóstica de actividad cerebral epileptiforme en pacientes críticos del Instituto de Neurocirugía Asenjo (INCA) mediante análisis de la señal EEG, utilizando algoritmos de clasificación de Machine Learning.

Corroborar esta hipótesis, es decir, ser capaces de identificar crisis epilépticas mediante un sistema automático podría significar un impacto mayor en la capacidad diagnóstica y de tratamiento de los pacientes del INCA. Además podría modificar el modo en el que se ejecuta el estudio de los registros electroencefalográficos, llegando a impactar también las rutinas de trabajo de los médicos del centro.

⁷ BESA GmbH, Gräfelfing, Alemania

⁸ AIT Austrian Institute of Technology GmbH, Viena, Austria

⁹ Persyst Development Corporation, California, EE. UU.

1.5. Objetivos

Dada la problemática descrita previamente, se definen para este trabajo de título los objetivos descritos a continuación:

1.5.1. Objetivo general

Desarrollar un modelo clasificador de actividad epileptiforme en pacientes críticos del Instituto de Neurocirugía Asenjo (INCA), mediante análisis de la actividad cerebral utilizando algoritmos de Machine Learning, para mejorar la capacidad de diagnóstico y así propiciar tratamiento temprano y mejora en el pronóstico del paciente.

1.5.2. Objetivos específicos

1. Establecer el estado del arte en cuanto a detección, clasificación y predicción de episodios de actividad epileptiforme en el mundo.
2. Construir un set de datos con muestras de pacientes reales para el entrenamiento de modelos.
3. Seleccionar, implementar y evaluar modelos de clasificación de actividad epileptiforme.
4. Desarrollar un prototipo expositivo del mejor modelo.
5. Evaluar impacto económico y social de la implementación.

1.6. Marco conceptual

Se aborda el problema desde la perspectiva de los sistemas de información y ciencias de los datos. Los modelos estadísticos estudiados en el curso formativo de la carrera de ingeniería civil industrial son la base de las herramientas de minería de datos que se utilizan. Para la detección y clasificación de actividad epileptiforme se hace necesario el desarrollo e integración de algoritmos de minería de datos y aprendizaje de máquinas, junto a conceptos de registro y análisis de electroencefalografía.

1.6.1. Electroencefalografía

Un electroencefalograma (EEG) es una exploración del registro de actividad eléctrica propia del cerebro (en particular de la corteza cerebral) mediante un equipo de electroencefalografía, un mecanismo que es no invasivo. El registro se hace a través de múltiples electrodos colocados en el cuero cabelludo, donde cada uno de ellos corresponde a un canal del registro.

El EEG mide los cambios de voltaje que resultan de corrientes iónicas entre las neuronas del cerebro [32], cuyos efectos pueden ser apreciados de forma local o generalizada. Para determinar esto, es que los electrodos se reparten a lo largo de la cabeza en diferentes configuraciones.

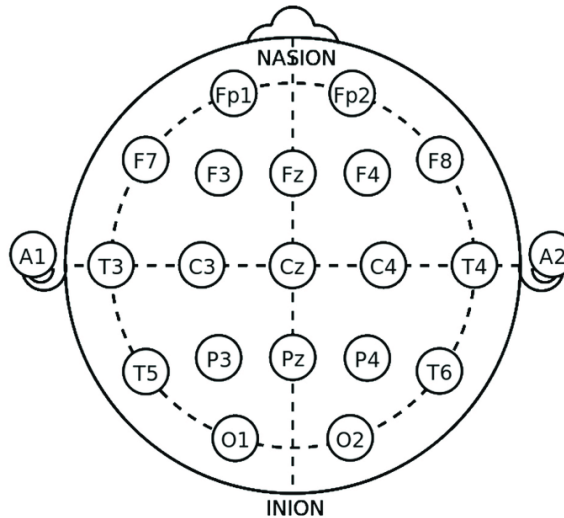


Figura 1.4: Sistema Internacional 10-20 para la colocación de los electrodos extracraneales. Fuente: Rojas et al. [38].

El montaje estándar 10-20 o sistema internacional 10-20 es una configuración reconocida internacionalmente para ubicar los electrodos en el cuero cabelludo. Este montaje estandarizado busca garantizar que los resultados del estudio de un sujeto puedan reproducirse, analizarse y compararse eficazmente. En esta configuración, las distancias reales entre los electrodos adyacentes son el 10% o el 20% de la distancia total de adelante hacia atrás (desde el nasión¹⁰ hasta el inión¹¹) o de derecha a izquierda del cráneo (desde un oído a otro). Tal como se aprecia en la Figura 1.4, cada uno de los 21 electrodos se nombra según el lóbulo o área del cerebro sobre la que se ubica y de la que está recibiendo señales: prefrontal (Fp), frontal (F), temporal (T), parietal (P), occipital (O) y central (C)¹². Los electrodos del lado derecho de la cabeza se etiquetan con números pares (2, 4, 6, 8), mientras que los de la izquierda con números impares (1, 3, 5, 7). A los electrodos ubicados en el plano sagital¹³ se les asigna la letra Z¹⁴ (FpZ, Fz, Cz, Oz) y se usan principalmente como puntos de referencia en las mediciones.

El EEG se usa con mayor frecuencia para diagnosticar epilepsia, que se manifiesta como anomalías en los registros electroencefalográficos [45]. Entre otras aplicaciones diagnósticas, se encuentran trastornos del sueño, profundidad anestésica, coma, encefalopatías y muerte cerebral. El electroencefalograma solía ser uno de los principales métodos de diagnóstico para tumores, accidentes cerebro-vasculares y otros trastornos cerebrales focales [5], pero este uso ha disminuido con la aparición de técnicas de imágenes anatómicas de alta resolución como la resonancia magnética (MRI) y la tomografía computarizada (TC). A pesar de su resolución espacial limitada, del orden de un par de centímetros, el EEG se mantiene como una

¹⁰ El nasión es el punto de intersección del hueso frontal y de dos huesos nasales del cráneo humano. En el rostro se aprecia como una zona notoriamente deprimida que se halla entre los ojos, justo por arriba del puente de la nariz.

¹¹ El inión es la proyección más prominente del hueso occipital en la parte posterior-inferior del cráneo humano.

¹² No existe un *lóbulo central*. Debido a su ubicación, y dependiendo del individuo, los electrodos C pueden representar la actividad EEG más típica de la actividad frontal, temporal y un poco de parietal-occipital.

¹³ En anatomía el plano sagital es el que divide al cuerpo en mitades izquierda y derecha.

¹⁴ Por cero en inglés (zero).

herramienta valiosa para la investigación y el diagnóstico. Presenta como grandes ventajas, el hecho de ser una de las pocas técnicas móviles disponibles y que su resolución temporal de rango de milisegundos no es posible de alcanzar con TC o MRI.



Figura 1.5: Descargas generalizadas de tipo espiga-onda de 3 Hz en un electroencefalograma.

Existe ciertos tipos de patrones asociados al desarrollo de una crisis¹⁵ que constituya un episodio de Estado Epiléptico no Convulsivo, por ejemplo las descargas tipo espiga-onda, como las registradas en la Figura 1.5.

1.6.2. Método KDD - Descubrimiento de conocimiento en bases de datos

En el contexto de procesos que requieren del análisis de grandes cantidades de datos, que pueden ser imposibles de analizar por una persona en un tiempo acotado, surge el Knowledge Discovery from Databases, también conocido como KDD, o descubrimiento de conocimiento en bases de datos.

Definido por Fayyad en 1996 como “el proceso no trivial de identificar patrones válidos, novedosos, potencialmente útiles y, en última instancia, comprensibles a partir de los datos”, es un proceso iterativo, del que existen diferentes versiones. Según describen Maimon & Rockach, consta de nueve pasos [29]:

¹⁵ Ver tabla de *Criterios para crisis no convulsivas* en Anexos.

1. **Comprensión del dominio del estudio:** Es un paso preparatorio en el que se definen los límites y objetivos del análisis a realizar. En este paso se identifican las fuentes de información más importantes para el desarrollo del estudio.
2. **Selección de datos y creación de dataset:** Se define la información útil para la investigación. Además debe determinarse qué parte de esa información ya está disponible y cual debe ser obtenida, para la creación de un único set de datos con los atributos que serán necesarios para el estudio.
3. **Pre-procesamiento de datos:** Se debe verificar que los datos seleccionados estén correctamente estructurados. Para ello, se debe eliminar el ruido y los datos aislados. Usando el conocimiento previo se eliminan las inconsistencias y los duplicados. También debe definirse estrategias para manejar la información faltante en los datasets.
4. **Transformación de los datos:** Se procesa los datos para dejarlos en un formato adecuado para poder aplicar las técnicas de minería de datos.
5. **Selección de la tarea de minería de datos:** Dependiendo del objetivo se pueden ejecutar distintos tipos de tarea. Puede buscarse resultados estadísticos (como mediana o media por ejemplo), o bien, alguna una predicción, una asociación, o identificación de secuencias de datos.
6. **Selección de el o los algoritmos de minería de datos:** Cada algoritmo tiene diferentes parámetros y técnicas de aprendizaje, por lo que una vez definido el objetivo, debe elegirse el algoritmo que más se adecúe a este.
7. **Implementación del algoritmo de minería de datos:** El algoritmo busca patrones y modelos relevantes, siguiendo sus reglas preestablecidas. Pueden ser árboles de decisión, modelos de regresión, modelos de clusterización, modelos mixtos, entre otros.
8. **Interpretación y evaluación:** Esta etapa es en la cual se procede a transformar el trabajo previo en conocimiento para poder concluir sobre algún fenómeno en específico o tomar decisiones.
9. **Uso del conocimiento descubierto:** Este paso determina la efectividad del proceso completo de KDD. El conocimiento descubierto se vuelve activo y puede generar cambios en un sistema, con efectos medibles.

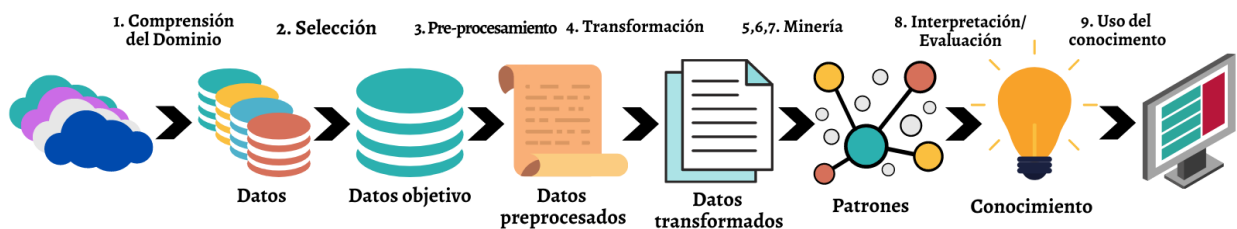


Figura 1.6: El proceso de Knowledge Discovery in Databases. Basado en esquema de Maimon & Rockach [29].

También se considera a veces la visualización para entender mejor los resultados obtenidos.

1.6.3. Minería de datos

La minería de datos¹⁶ es la etapa del proceso KDD en la que busca descubrir patrones en grandes volúmenes de datos [29]. Para ello, hace uso de herramientas de inteligencia artificial, aprendizaje automático, sistemas de bases de datos y estadísticas. El uso de la minería de datos puede usarse para el análisis automático o semi-automático de grandes cantidades de datos en busca de patrones relevantes que sean desconocidos, como los grupos de datos, análisis de clúster, detección de anomalías o relaciones de dependencias entre los datos. Dichos patrones pueden ser usados en el aprendizaje automático, para problemas de clasificación y análisis predictivo.

1.6.4. Aprendizaje automático

El aprendizaje automático, o machine learning, es una rama de la inteligencia artificial que busca dotar a las máquinas de capacidad de aprendizaje, es decir, la generalización de un conocimiento a partir de un conjunto de experiencias. En concreto, se busca el uso de algoritmos y heurísticas para convertir muestras de datos en programas de computador, sin tener que escribirlos explícitamente. Los modelos o programas resultantes deben ser capaces de generalizar comportamientos e inferencias para un conjunto más amplio (potencialmente infinito) de datos.

Según el objetivo del análisis de los datos, los algoritmos utilizados se clasifican en supervisados y no supervisados [51]:

Algoritmos supervisados

Este tipo de algoritmos requieren de intervención humana para etiquetar o clasificar los datos que se introducen a sus modelos. Los algoritmos supervisados entregan como datos de salida alguna variable esperable, según las etiquetas o clasificaciones hechas previamente por alguien en los datos que alimentan el entrenamiento del modelo.

Es posible separar este tipo de algoritmos en dos clases:

- **Clasificación:** A partir de las variables de los datos, son capaces de separarlos según clases predefinidas. Por ejemplo la identificación de si un correo corresponde o no a spam.
- **Regresión:** Entregan como resultado un valor numérico, que es el que mejor se ajusta en función de las variables asociadas. Un ejemplo es la predicción de temperatura o rendimiento de un vehículo.

Algoritmos no supervisados

En este caso la intervención humana de los datos no es necesaria. En este caso los algoritmos buscan patrones o relaciones en la estructura misma de los datos. Los datos de entrada no deben estar etiquetados previamente.

¹⁶ Referida también por su nombre en inglés *data mining*.

A su vez, también es posible separar este tipo de algoritmos en dos clases:

- Clustering: Clasifican los datos de salida en grupos. Un ejemplo de aplicación es la segmentación de clientes de algún negocio.
- Asociación: Son capaces de descubrir reglas dentro de los conjuntos de datos. Por ejemplo, compras relacionadas en un supermercado, tales como leche y cereales.

Las tareas de aprendizaje automático se ejecutan por lo general siguiendo una estructura estándar. En primer lugar se separan los datos en al menos 2 subconjuntos: uno de entrenamiento y uno de testeo. El primero sirve como base para que la máquina “aprenda” a detectar estructuras o patrones en la data, con el fin de predecir valores o clasificar en categorías los datos del subconjunto de testeo. A partir de dicha predicción o clasificación es posible comparar el resultado de los modelos con los valores reales y estimar un error. Los algoritmos de aprendizaje de máquinas buscan minimizar este error mediante distintos mecanismos de optimización.

1.6.5. Support Vector Machines (SVM)

Esta técnica de clasificación considera un conjunto de datos, cada uno marcado como perteneciente a una de dos categorías, y sus respectivos vectores de características. Consiste en aumentar la dimensionalidad de dicho set de características, posicionando cada elemento del conjunto de datos en un punto dado del hiperespacio, y luego buscando ajustar un hiperplano capaz de dividir dicho espacio entre las categorías descritas. Así, como se muestra en la Figura 1.7, un modelo de SVM intenta maximizar el ancho de la brecha entre las dos categorías. El algoritmo SVM es usado ampliamente en las ciencias biológicas y médicas por su fácil interpretabilidad.

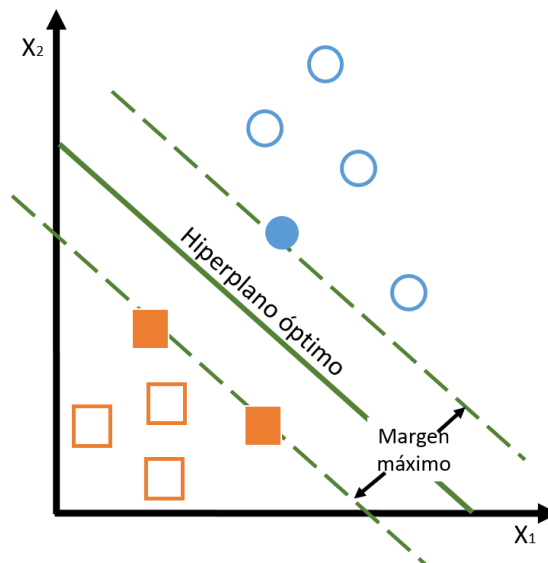


Figura 1.7: Representación de un hiperplano separando categorías a través de SVM.

1.6.6. Redes neuronales artificiales (ANN)

Las ANN¹⁷ son sistemas computacionales inspirados en las redes neuronales presentes en los cerebros de los seres vivos. Están compuestas por conjuntos de neuronas artificiales, a su vez inspiradas en las neuronas biológicas.

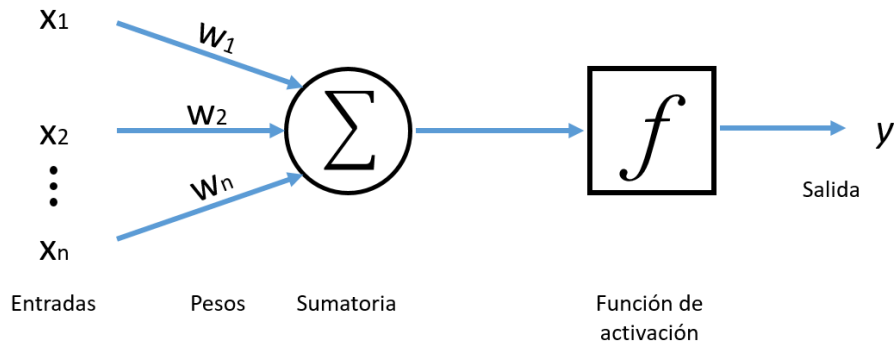


Figura 1.8: Esquema de una neurona artificial. Basado en esquema de Dongare, Kharde & Kachare [11].

Las neuronas artificiales reciben una señal o vector de entrada X . Cada uno de los elementos de la señal, o variables del vector x_i , se combinan en una sumatoria. Como se muestra en la Figura 1.8, esta suma no es directa, sino que cada entrada se pondera por un peso w_i , lo que determina cuánto influye cada variable en el resultado final. En el interior de la neurona se suman las entradas ponderadas para dar paso a una función de activación. Esta función recibe la información de la neurona, como la combinación lineal entre los valores de las variables y sus pesos, y entrega como resultado un valor único y , lo que corresponde a la salida de la neurona artificial. Existen diversas funciones de activación, y pueden variar según el tipo de resultado que se espere. Por ejemplo, para intentar predecir si un cliente comprará o no un producto se usarán funciones de activación con salidas binarias, mientras que para predecir el número total de ventas se puede usar funciones de activación lineales.

Para formar una red neuronal las neuronas normalmente se agrupan en capas, como se muestra en la Figura 1.9. Las diferentes capas pueden realizar diferentes transformaciones a la información que reciben, según la estructura o sus distintas funciones de activación. Las señales pasan desde la primera capa (la capa de entrada) hasta la última capa (la capa de salida), después de eventualmente atravesar las distintas capas ocultas de la red, donde la salida de cada capa se transforma en las variables de entrada de la siguiente.

¹⁷ Por su nombre en inglés *Artificial Neural Networks*.

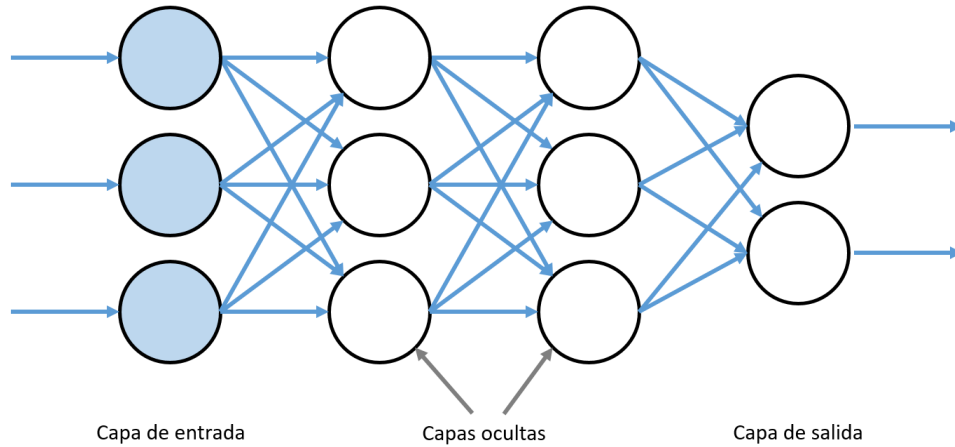


Figura 1.9: Esquema de una red neuronal artificial. Basado en esquema de Dongare, Kharde & Kachare [11].

Como se menciona anteriormente, la forma en la que estos modelos “aprenden” es a través de la estimación del error entre un valor real y el valor predicho por la red. Estos algoritmos modifican los valores de los pesos w_i de las neuronas que las componen, de modo de disminuir el error. A esto se le llama ajuste del modelo y es un proceso iterativo, en donde se busca llevar el error al mínimo posible.

1.6.7. Redes neuronales convolucionales (CNN)

Un tipo particular de redes neuronales artificiales son las CNN¹⁸. Estas redes de aprendizaje supervisado son generalmente usadas para tareas que involucren señales en dos dimensiones, tales como imágenes o fotogramas de video[25], buscando imitar el procesamiento visual humano para identificar elementos distintivos en dichas señales y así poder reconocer patrones en ellas. Las capas de una red convolucional se especializan en el reconocimiento de un tipo específico de estructura, las que se hacen más complejas a medida que la red es más profunda. Así, las primeras capas de una CNN son capaces de detectar líneas o curvas, llegando a reconocer estructuras como objetos concretos, siluetas o rostros en capas más profundas.

Para lograr lo anterior es que se utilizan las convoluciones. En el caso del procesamiento de imágenes, la red considera como entrada los píxeles de la misma, por lo que para una imagen de $m \times n$ píxeles, la capa de entrada de la CNN deberá tener $m \times n$ neuronas. Lo anterior, en el caso de estar trabajando con una imagen monocromática. En el caso de las imágenes a color, es necesario utilizar 3 canales, donde cada uno tendrá la representación RGB¹⁹ que le corresponda. Las convoluciones consisten en operar mediante producto punto grupos de píxeles cercanos de la imagen original contra una matriz llamada kernel o filtro. Así, se recorre toda la imagen de entrada y se genera una matriz por cada convolución, sobre la que se aplica una función de activación para obtener la matriz de salida. El resultado final dependerá del filtro utilizado, como se observa en la Figura 1.10. Es necesario hacer notar que en este paso la cantidad de datos a procesar aumenta considerablemente. Suponiendo

¹⁸ Por su nombre en inglés *Convolutional Neural Networks*.

¹⁹ Es un modelo de color basado en la síntesis aditiva, con el que es posible representar un color mediante la mezcla por adición de rojo, verde y azul.

que después de recorrer toda la imagen original con el kernel se generan z nuevas imágenes filtradas, entonces luego de este proceso habrá $m \times n \times z$ imágenes.

Operación	Filtro	Imagen resultante
Identidad	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
Detección de bordes	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
Mejora de nitidez	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Desenfoque de caja	$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	
Desenfoque gaussiano	$\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	

Figura 1.10: Efectos de diferentes filtros o matrices de convolución sobre una imagen. Fuente: Albawi, Mohammed & Al-Zawi. [1].

El proceso anteriormente descrito corresponde a una sola capa. Antes de ejecutar una nueva convolución es necesario reducir dimensionalmente la cantidad de neuronas para evitar un crecimiento exponencial de la data que dificulte su procesamiento. Para ello se ejecuta un proceso de subsampling, reduciendo el tamaño de las imágenes filtradas. Una de los tipos de subsampling más usado es max-pooling. Como se muestra en la Figura 1.11, este método consiste en agrupar los píxeles cercanos en una imagen, preservando solo el mayor valor del grupo seleccionado. Lo anterior permite reducir de forma importante la cantidad de data a procesar, manteniendo la información más relevante para detectar las características de la imagen inicial.

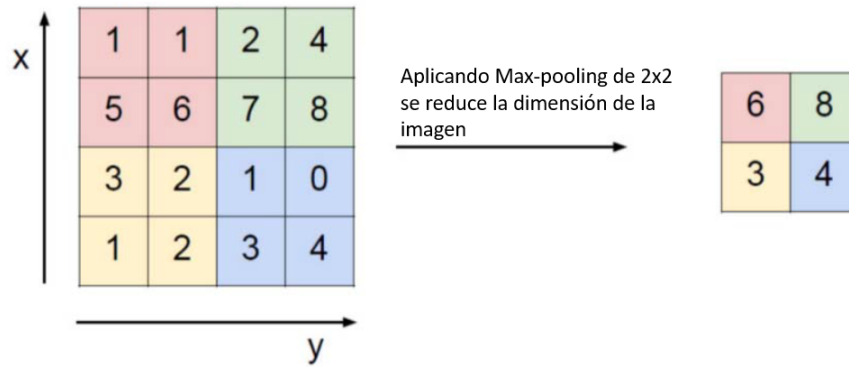


Figura 1.11: Transformación de una imagen a través de max-pooling de 2x2.
Fuente: Albawi, Mohammed & Al-Zawi. [1].

Este proceso se repite según la cantidad de capas de convolución que se definan según la arquitectura, para luego conectar la capa de subsampling con una capa tradicional, que estará encargada de ejecutar la tarea de clasificación. El proceso de aprendizaje es similar al descrito anteriormente, en donde se modifican los pesos asociados a cada neurona, con la diferencia de que ahora también es posible modificar los filtros usados para procesar las imágenes.

Para el análisis de señales de una dimensión se han desarrollado modelos de CNN en 1D, con buenos resultados en clasificación de datos biomédicos y diagnóstico temprano, monitoreo de salud estructural y detección e identificación de anomalías en electrónica de potencia, entre otros [25]. El método de trabajo de una CNN en 1D es análogo al de las CNN en 2D, con la diferencia de que las operaciones de convolución corresponden a multiplicaciones escalares y sumas en este caso.

1.6.8. Evaluación de modelos

A la hora de entrenar y testear un modelo es necesario considerar las métricas con la que se evaluará al contrastar los resultados y las observaciones en el mundo real. Estas son necesarias también para hacer comparaciones entre distintos modelos.

En tareas de clasificación las métricas derivan de los aciertos o errores del modelo a la hora de identificar la pertenencia de una muestra a cierta clase. La pertenencia a una clase está definida por la observación de una condición discriminante a ojos de un experto. El problema estudiado corresponde a clasificación binaria, por tanto existen dos clases. La condición discriminante en este caso es la existencia de crisis epilépticas, por lo que esa será la clase positiva.

	Condición positiva	Condición negativa
Resultado de la clasificación positivo	Verdadero positivo	Falso positivo
Resultado de la clasificación negativo	Falso negativo	Verdadero negativo

Tabla 1.1: Posibles combinaciones de resultados de una clasificación binaria.

Tal como se resume en la Tabla 1.1, un verdadero positivo (TP)²⁰ es un resultado en el que el modelo predice correctamente la clase positiva. De manera similar, un verdadero negativo (TN)²¹ es un resultado en el que el modelo predice correctamente la clase negativa. Un falso positivo (FP) es un resultado en el que el modelo predice incorrectamente la clase positiva. Y un falso negativo (FN) es un resultado en el que el modelo predice incorrectamente la clase negativa.

Para evaluar los modelos de clasificación se utiliza métricas derivadas de estos cuatro resultados:

- Accuracy: $(TP+TN)/(TP+TN+FP+FN)$, es decir, cuantos elementos son clasificados correctamente en la categoría a la que pertenecen.
- Recall: $TP/(TP + FN)$, es decir, cuantos elementos relevantes se está seleccionando (ie. Cuantas crisis es capaz de identificar el modelo del total de crisis que realmente ocurren).
- Precision: $TP/(TP + FP)$, es decir, cuantos de los eventos seleccionados son relevantes (ie. Cuantos de los eventos clasificados como crisis son efectivamente crisis).
- Specificity: $TN/(TN + FP)$, es decir, la proporción de elementos que no tienen la condición que no son seleccionados. (ie. Cuantos de los eventos clasificados como normales son efectivamente de ese modo).

1.7. Metodología

Las etapas de este trabajo siguen los pasos establecidos por el Knowledge Discovery in Databases. Para entender el contexto en el que se trabaja, se realiza una investigación del estado del arte tanto en relación a caracterizaciones de la condición clínica, como en utilización de algoritmos de aprendizaje de máquinas en diagnóstico médico y aplicaciones previas de clasificación automática en condiciones neurológicas. A partir de ello, se define qué información es efectivamente útil o relevante para la investigación, además de pertinente al contexto en el que esta se realiza²².

Los datos serán obtenidos de pacientes del Instituto Dr. Asenjo (INCA) y deberán ser procesados de modo que su identificación no sea posible. Es esencial que las muestras incluyan pacientes que hayan sufrido una o más crisis epilépticas, y que éstas estén debidamente identificadas en los registros a utilizar, para el posterior entrenamiento de los modelos.

Además, se realizará la verificación de que los datos seleccionados estén correctamente estructurados. Se transformará los datos a un formato adecuado para poder aplicar las técnicas de minería, capaces de analizar los datos y poder detectar patrones en estos.

²⁰ Del inglés true positive.

²¹ Del inglés true negative.

²² Ver Alcances.

En principio, se pretende parcelar los registros en pequeñas ventanas temporales, a partir de las cuales sea posible extraer un determinado set de características, también llamado vector de características. A partir de cada segmento, se deberá transformar el registro de cada canal a dicho vector de características, agrupados en dominios de tiempo y frecuencia. Determinar el conjunto de variables a extraer es uno de los objetivos del presente trabajo, y dependerá en gran parte de la primera etapa de trabajo planteada. Se deberá testear además la relevancia de dichas variables.

Luego de procesar y transformar los datos surgidos exclusivamente de los registros de EEG se determinará la posibilidad de incluir en los modelos variables personales de los pacientes, tales como edad, sexo, patologías relacionadas, entre otros. Para ello se deberá realizar un cruce entre dichos datos y las variables procesadas previamente.

Una vez determinadas las variables, se procederá a entrenar los modelos para la detección de actividad epileptiforme. Cómo dicha actividad puede presentarse de forma local, en uno o pocos electrodos, o generalizada, cuya manifestación se hace presente en la mayoría de los canales a estudiar, será necesario definir, dependiendo del modelo a usar, si se etiquetará los segmentos identificados como crisis de actividad epileptiforme para su determinado segmento temporal de forma local (por canal) o global (considerando todos los canales).

Una vez realizados dichos modelos, se evaluará cuál de ellos presenta mejores resultados en términos de *precision* y *recall*. Por las características del problema estudiado, se otorgará mayor relevancia a un mejor resultado en términos de *recall*, de modo de identificar la mayor cantidad de crisis posibles. Un nivel de aceptación mínimo de *precision* (o falsos negativos) se definirá en función de los resultados previos, estudiados del estado del arte. El modelo que presente el mejor rendimiento será utilizado para diseñar un prototipo expositivo. Este prototipo usará como entrada el registro EEG de algún paciente, y segmentará y etiquetará ventanas de tiempo relevantes en las que podría estar ocurriendo una crisis epileptiforme. El principal foco de su desarrollo es entregar un apoyo visual a los resultados del presente estudio.

La etapa final de este trabajo corresponde a la estimación del impacto, tanto económico como social del uso del conocimiento obtenido. Dicho uso contempla los eventuales costos y beneficios de su implementación, con el objetivo de determinar su factibilidad y posibilidades de sostenimiento a lo largo del tiempo, además de entregarse un conjunto de recomendaciones y elementos a considerar para trabajos futuros. También esta etapa contempla la determinación de los efectos de la aplicación para los pacientes y equipos médicos involucrados tradicionalmente en las tareas de registro e interpretación de electroencefalografía.

1.8. Resultados esperados

Los resultados esperados del presente trabajo de título, que a la vez incluye los entregables son:

- Una revisión actualizada del estado del arte.
- Un set de datos con muestras de pacientes reales para el entrenamiento de modelos.
- Las variables relevantes para la clasificación de la actividad cerebral y sus mecanismos

de extracción de la señal.

- Una evaluación de modelos de clasificación de actividad epileptiforme.
- El mejor modelo obtenido, junto a su prototipo expositivo, para la visualización de los resultados del trabajo.
- Una evaluación del impacto económico y social de una posible implementación del mecanismo propuesto.
- Un conjunto de recomendaciones para futuros trabajos a partir del aprendizaje obtenido.

1.8.1. Alcances y riesgos

El presente trabajo de título se realiza en un escenario acotado, cuyas principales limitaciones y alcances se detallan a continuación.

Se trabajará con a una fuente de datos primaria, en este caso con la colaboración del Dr. Brunetti del INCA, de pacientes internados con neuro-monitorización. Estos datos incluirán registros electroencefalográficos, caracterización demográfica de los pacientes, patologías asociadas, entre otros. Para su análisis y estudio se tendrá especial cuidado en cumplir con las garantías de confidencialidad mediante la anonimización de datos personales y el uso de consentimientos informados.

Las muestras que serán utilizadas para entrenar y evaluar los modelos corresponden a pacientes adultos. Este aspecto es relevante, considerando las diferencias morfológicas y fisiológicas del cerebro entre neonatos, niños y adultos. Por lo anterior, los resultados que se puedan obtener no son necesariamente generalizables a un grupo etario diferente del señalado.

Se considera para el desarrollo de este trabajo el uso de pacientes del servicio de pacientes críticos, por lo que las muestras corresponderán a pacientes neurocríticos. La actividad cerebral en el escenario de un cerebro dañado, bajo tratamiento farmacológico o anestésico, es distinta a una actividad cerebral normal, con patrones diferentes y típicamente menos organizados. Lo anterior hace que la actividad, tanto normal como anormal, de pacientes monitorizados en UCI pueda variar ampliamente de un sujeto a otro.

Se identifica como riesgo la posible limitación al acceso de datos, al tratarse de datos médicos, protegidos por ley, incluso a nivel constitucional. En caso de no ser posible utilizar datos clínicos primarios, se podría recurrir al uso de registros de repositorios o de terceros.

Finalmente, es importante señalar que los objetivos de este trabajo apuntan a identificar si es posible o no detectar y clasificar actividad anormal de tipo epileptiforme, y a evaluar su posible implementación, mas no contempla la implementación de mecanismos de ningún tipo in situ.

Capítulo 2

Estado del arte

En el presente capítulo se revisan algunos trabajos previos que abordan la problemática propuesta desde distintas perspectivas. Se mencionan así diversos modelos que buscan categorizar la actividad neurológica durante crisis epilépticas. Los trabajos incluidos en esta revisión se consideran pertinentes por el alcance de sus resultados, la similitud de los datos utilizados con los estudiados en este trabajo o por la utilidad de sus métodos para el análisis.

2.1. Temko, Thomas, et al. (2011)

El estudio[46] busca detectar la presencia de crisis epilépticas en EEG neonatal usando un clasificador de Support Vector Machine (SVM). El clasificador usado es de tipo binario, capaz de distinguir entre períodos de crisis y períodos sin crisis. Además el sistema es independiente del paciente, es decir, el modelo entrenado no requiere de ver al paciente estudiado de antemano.

Set de datos

Se utilizó registros de 55 recién nacidos de la unidad de cuidados intensivos neonatal del Cork University Maternity Hospital, de los que 17 presentaron crisis epilépticas. El montaje del EEG fue del tipo 10-20, modificado para neonatos. En particular se estudian ocho canales bipolares: F4-C4, C4-O2, F3-C3, C3-O1, T4-C4, C4-Cz y C3-T3. El tiempo total de registro es de 267,9 horas, durante las que se registraron 705 crisis.

Se señala que los registros no fueron modificados para reducir artefactos propios de la UCI.

Variables utilizadas

Las variables elegidas para el estudio se basan principalmente en las que muestran ser relevantes en la detección de crisis en neonatos en diversos trabajos anteriores[46]. Las variables extraídas se pueden agrupar en 3 grupos principales: dominio de la frecuencia, dominio del tiempo y teoría de la información.

Modelo

Para el desarrollo del modelo se preprocesa cada canal del EEG por separado y se segmenta en épocas. Cada una de estas épocas es de 8 segundos y hay un traslape de 50 % entre épocas. Se extraen las variables de cada época y el vector de características se introduce al clasificador SVM, que entrega una probabilidad de crisis para cada época del registro. Dichas probabilidades se suavizan con un filtro de media móvil, transformándolo en decisiones binarias. Las decisiones binarias de cada canal posteriormente se combinan en una decisión binaria multicanal. Finalmente se procesa el registro para expandir la crisis (decisiones positivas del clasificador) a los límites propios del evento¹.

El clasificador SVM utiliza el método LOOCV de validación. Durante el entrenamiento se etiquetan las épocas de crisis y épocas sin crisis para cada canal. Se selecciona 2 minutos de crisis por paciente para cada canal involucrado en la crisis. Para la clase sin actividad epileptiforme se utilizan 10.000 épocas seleccionadas aleatoriamente de todos los canales, de todos los pacientes del set de entrenamiento. La data de entrenamiento se normaliza anisotrópicamente para ser utilizada.

Desempeño

Se definen 2 métricas de desempeño: tasa de buenas detecciones (GDR)² y el número de falsas detecciones por hora (FD/h). Para el primero, si una crisis era detectada en un período etiquetado como crisis por un experto en EEG neonatal, se consideró una buena detección. Para el segundo se definen tres clases: actividad de fondo libre de artefactos, EEG contaminado con artefactos y actividad similar a crisis.

El sistema mostró ser capaz de detectar (GDR) alrededor del 89 % de crisis correctamente a 1FD/h, $\sim 96\%$ a 2FD/h y $\sim 100\%$ a 4FD/h. A su vez, fue capaz de detectar más del 90 % de las crisis de duraciones mayores a 2 minutos, y alrededor del 70 % de las crisis de duración menor a 1 minuto. Este menor rendimiento se atribuye al efecto del filtro de media móvil, que puede hacer que crisis de corta duración se confundan con la señal de fondo.

Conclusiones

El método de procesamiento por canales implica que el sistema es aplicable a registros con cualquier cantidad de canales, lo que significa un beneficio en su uso clínico. El sistema basado en SVM permite controlar la decisión final al elegir distintos niveles de confianza que pueden ser flexibles acorde a las necesidades clínicas.

¹ Mediante el uso de una técnica “collar”, propia del procesamiento del habla, en donde se usa para evitar recortar el inicio y final de las palabras.

² Del inglés good detection rate.

2.2. Ullah, Hussain, et al. (2018)

Se presenta un sistema que busca clasificar la actividad de pacientes usando un conjunto de modelos de redes neuronales convolucionales unidimensionales[49] (1D-CNN). El clasificador propuesto es de tipo ternario, distinguiendo estado normal, estado ictal y estado interictal de los pacientes. Se utilizan técnicas de aumento de datos para sobrellevar la baja cantidad de datos y reducción de parámetros para un mejor uso de los recursos computacionales.

Set de datos

Los datos utilizados corresponden al set obtenido por el equipo de investigación de la Universidad de Bonn. Dichas señales EEG fueron registradas con un montaje 10-20 estándar. Los registros corresponden a instancias monocanales de cinco voluntarios sanos y cinco pacientes epilépticos en distintos estados (ictal-interictal).

El numero de instancias de dicho set de datos no es suficiente para entrenar un modelo de aprendizaje profundo. Por las dificultades prácticas de reunir un mayor numero de registros de EEG, se plantea un esquema de aumento de datos, hasta una cantidad suficiente para el entrenamiento del modelo CNN.

Variables utilizadas

Al ser un modelo CNN, se tiene un análisis end-to-end, por lo que no hubo selección manual de las variables involucradas.

Modelo

El modelo propuesto consiste en 3 módulos: (i) el modulo de entrada, que recibe la señal EEG y la divide en subseñales usando ventanas traslapadas de tamaño fijo y las entrega a los modelos P-1D-CNN, (ii) el arreglo en el que las subseñales son clasificadas por los modelos P-1D-CNN, y (iii) la fusión y decisión, en la que las decisiones locales se integran usando voto mayoritario para tomar la decisión final.

Una vez que el modelo es entrenado, se generan copias para armar el arreglo piramidal, en donde cada modelo cumple el rol de un experto examinando una parte diferente de la señal de entrada. Cuando una señal es entregada al arreglo para su clasificación, se divide en las ventanas traslapadas, donde cada una se entrega a distintos expertos (modelos) para su análisis local.

El modelo P-1D-CNN es un modelo profundo con tres tipos de capas: convolución (Conv), normalización por lotes (BN), y completamente conectada (FC). Para las capas Conv y FC se utiliza ReLU como función de activación.

Se explica que a través de la profundidad del modelo se combinan distintas capas convolucionales con un numero de kernels decrecientes, con la idea de extraer un mayor número de microestructuras en las capas de bajo nivel, que al combinarse en capas de alto nivel generan un menor número de características, pero discriminantes. Por tanto el modelo realiza una

selección de variables implícita.

Desempeño

Luego de testear en diversas configuraciones los datos existentes, se llega a un accuracy de $99.1 \pm 0.9\%$ en la detección de epilepsia.

Conclusiones

El modelo propuesto obtiene buenos resultados considerando la cantidad de datos, los esquemas de aumento de datos, la rapidez y bajo uso de recursos computacionales. El modelo, al ser entrenado en un conjunto de datos de referencia, debe ser validado en una configuración clínica real.

2.3. Golmohammadi, et al. (2019)

En este estudio[16] se propone un sistema híbrido basado principios de machine learning y big data. La arquitectura híbrida integra modelos ocultos de Markov para el descifrado del EEG, con modelos de deep learning. El sistema es capaz de clasificar tres patrones de interés clínico en la actividad cerebral: (1) espigas y/o ondas agudas, (2) descargas epileptiformes periódicas generalizadas, (3) descargas epileptiformes lateralizadas periódicas. También clasifica tres patrones utilizados para modelar la actividad del EEG de fondo: (1) movimiento de los ojos, (2) artefactos y (3) fondo.

Set de datos

Se utiliza el TUH EEG Corpus[33], que incluye registros desde el año 2002 al 2015. En suma, contiene datos de más de 13.500 pacientes, correspondiendo a sobre 1,8 años de señal multicanal. Este conjunto de datos posee registros de pacientes ambulatorios, en unidades de cuidados intensivos, salas de emergencia, entre otros. Al ser registros clínicos, contienen diversos artefactos presentes en el mundo real. Cada sesión incluye un archivo EDF y un reporte fisiológico, generado por neurólogos certificados.

Los registros varían entre 20 y 128 canales, pero para el estudio fueron usados los registros de 22 canales, asociados al montaje 10-20 estándar.

Para los propósitos de este trabajo, se etiquetó los datos manualmente, por parte de un equipo de anotadores de pregrado, para los seis eventos descritos anteriormente. El set de entrenamiento se diseñó para incluir segmentos de 359 sesiones, mientras que el de evaluación contiene segmentos de 159 sesiones. Cada paciente aparece solo una vez en el conjunto de datos. Las anotaciones hechas por el equipo son a nivel de canal.

Variables utilizadas

En el estudio se usa una metodología basada en en coeficientes cepstrales³, en donde se extraen flujos de características independientes de la señal de EEG multicanal. En particular, se usa coeficientes cepstrales de frecuencia lineal.

A través de la experimentación, usando la metodología anterior, se obtiene que el mejor desempeño corresponde a un arreglo de 26 características. Este vector de características incluye nueve variables absolutas, consistente en siete coeficientes cepstrales, un termino en el dominio de la frecuencia y un termino de diferencial de energía.

Se identifica como un tradeoff razonable entre complejidad computacional y performance, el dividir las ventanas de 10s en épocas de 1s, y luego dividir las en cuadros de 0.1s.

Modelo

El output de la extracción de variables son 22 canales de data, donde cada canal contiene un vector de dimensión 26 correspondiente a cada 0.1s.

Los modelos se entrenaron utilizando todos los eventos en todos los canales, resultando en modelos independientes del canal. Se señala que no se ha demostrado que los modelos dependientes del canal proporcionen un impulso en el rendimiento y agreguen una complejidad considerable al sistema.

Se considera entonces un modelo de múltiples pasos: (i) la decodificación secuencial utilizando modelos ocultos de Markov, (ii) el análisis de contexto espacial y temporal basado en el aprendizaje profundo , (iii) modelado de lenguaje estadístico.

El primer paso permite comprimir la señal. Posteriormente se aplica un tipo de red profunda conocida como *Stacked denoising Autoencoder*⁴ (SdA). Desde un punto de vista estructural, los SdA se componen de múltiples capas de autocodificadores de eliminación de ruido de manera que la entrada a cada capa es la representación latente del autocodificador de eliminación de ruido que se encuentra en la capa de abajo. La característica más importante de los codificadores automáticos de eliminación de ruido que los hace apropiados para el análisis automático de EEG es su capacidad para reconstruir una entrada reparada a partir de una versión dañada de la misma[16].

Desempeño

El rendimiento general del sistema de clasificación de aprendizaje profundo con modelos ocultos de Markov, una vez completados todos los pasos es de del 90.10% de sensibilidad y 4.88% de especificidad.

³ Coeficientes Cepstrales en las Frecuencias de Mel (MFCC), usada en reconocimiento del habla.

⁴ En español: Autoencoder de eliminación de ruido apilado.

Conclusiones

Se plantea en el estudio que “los médicos han indicado que una sensibilidad del 95 % con una especificidad por debajo del 5 % era el requisito mínimo para la aceptación clínica”, resultado que no logra alcanzar el modelo propuesto, incluso en sus mejores casos. Pese a ello, se considera que dichos resultados hacen al sistema de detección automática una opción viable de análisis clínico

Es relevante para el estado del arte el haber generado un modelo que funcione sobre una base de datos amplia, que incluye data de ambiente clínico. Lo anterior por la existencia de artefactos propios de las salas de hospitales y clínicas, que no se presentan en las bases de datos recolectadas en condiciones de estudio.

2.4. Zhang, et al (2016)

Se presenta un modelo[53] que pretende hacer clasificación ternaria (normal, ictal, interictal). Para ello se utiliza un método de fusión entre VMD⁵, y autorregresión para la extracción de variables, y un clasificador de random forest.

Set de datos

El set de datos utilizado es el de la Universidad de Bonn, descrito anteriormente.

Variables utilizadas

Para la extracción de las características relevantes se utiliza el método de VMD. Este descompone la señal en un número fijo de distintas subseñales. En este caso, se determina de forma empírica que 15 subseñales entregan mejores resultados. Una vez separadas, se normaliza su amplitud y se procede a determinar el orden óptimo para un modelo autorregresivo (AR). Para la extracción de parámetros del modelo AR, se construye usando el método de Burg⁶. Aunque el orden óptimo del modelo AR puede estimarse y sus correspondientes modelos son capaces de ajustarse a la secuencia de la señal procesada con la mayor precisión posible, la dimensionalidad de los coeficientes AR de cada dato es diversa entre sí. Dado que, bajo esta circunstancia, es muy difícil clasificar varias clases de señales EEG utilizando los coeficientes AR cuyas dimensionalidades son inconsistentes, se pone en marcha un segundo procedimiento de extracción de características, en donde se calculan ocho características estadísticas: energía, longitud, máximo, mínimo, media, varianza, asimetría y curtosis de los coeficientes del mejor modelo AR.

⁵ La descomposición en modo variacional (VMD) es un método de descomposición de señales moderno utilizado recientemente para muchos diagnósticos y controles de ingeniería, que reemplazó al método tradicional de descomposición en modo empírico (EMD)

⁶ detallar

Modelo

El modelo utilizado corresponde a un random forest, que es la agregación de múltiples árboles de decisión. El modelo recibe como input las ocho características descritas anteriormente, y se entrenan los modelos ajustando por 4 métricas: sensibilidad, especificidad, precisión y poder predictivo. El mejor resultado se obtiene fijando el número de árboles de decisión en 250. Cada árbol ejecuta la clasificación independientemente y luego se usa el voto de mayoría para decidir sobre el segmento estudiado.

Desempeño

El mejor resultado sobre el set de datos de Bonn es un accuracy de 97,35 % para los factores estudiados.

Conclusiones

El trabajo presenta una perspectiva distinta respecto a los demás revisados, utilizando técnicas de análisis de señales para descomponer el registro propio de EEG, junto a un clasificador de random forest, lo que amplía el espectro de posibilidades para abordar el problema de clasificación. Sin embargo, sus resultados no resultan significativamente distintos a los demás clasificadores, para el mismo set de datos. Es necesario destacar su buena capacidad de distinguir una crisis, pero a la vez mencionar nuevamente que el trabajo es realizado en condiciones de estudio que difieren a las condiciones clínicas en las que se enfoca el presente trabajo.

2.5. Song, et al. (2012)

El trabajo[42] propone un método para la detección automática de crisis epilépticas basado en un algoritmo de Sample Entropy Optimizado (O-SampEn), combinado con Extreme Learning Machine (ELM). Es un clasificador binario (crisis - no crisis).

Set de datos

Se utiliza el dataset de la Universidad de Bonn. Es importante señalar que este set de datos incluye registros intracraneales, utilizados en el desarrollo de este modelo. La investigación detalla los resultados con y sin el uso de las señales intracraneales, y el presente trabajo se enfoca solo en los que no lo hacen.

Variables utilizadas

Las variables utilizadas corresponden a métricas de tipo energético, descritas como entropía de la señal. Los parámetros utilizados se describen en el proceso de desarrollo de los múltiples pasos de los algoritmos presentados.

Modelo

Se utiliza un modelo de extreme learning machines. Según sus creadores, estos modelos pueden producir un buen rendimiento de generalización y aprender miles de veces más rápido que las redes entrenadas mediante retropropagación. En la literatura, también muestra que estos modelos pueden superar a las máquinas de vectores de soporte en aplicaciones de clasificación y regresión.

El método de Extreme Learning Machine simplifica el trabajo de una red neuronal de una única capa oculta, al fijar los pesos de las variables de entrada y sesgos a la capa oculta de modo aleatorio. Luego se busca el óptimo resolviendo una ecuación de orden lineal, respecto a los outputs de la capa oculta.

Desempeño

Se compara el rendimiento del modelo propuesto con modelos de SVM y BPNN.

La suma de los tiempos de ejecución de la extracción de características O-SampEn con las tres mejores combinaciones de parámetros y el tiempo de prueba del modelo ELM para un segmento de EEG con 4.096 puntos de muestreo es de aproximadamente 0,2122 s, que es mucho más corto que la duración del segmento de EEG (23,6 s).

En términos de accuracy, el modelo de ELM presenta un 99 %, el de SVM un 98 % y el de BPNN un 97.5 %.

Conclusiones

Se compara los tiempos de ejecución y entrenamiento de el modelo propuesto y modelos de SVM y BPNN, siendo el modelo propuesto considerablemente más rápido.

La alta precisión de identificación (recall de 99,00 %) y la baja carga computacional hacen posible construir un sistema de detección de ataques epilépticos en tiempo real sobre la base del enfoque propuesto.

Finalmente es relevante considerar nuevamente que en este caso fueron usados registros intracraneales, que muestran ser mucho más limpios de artefactos y ruidos causados por el ambiente o la aislación propia de los tejidos, que los registros a nivel de cuero cabelludo.

2.6. Gómez, et al. (2020)

El estudio más reciente a la fecha⁷. Se plantea un clasificador binario[17] usando datos de dos datasets distintos: el CHB-MIT Scalp EEG y el EPILEPSIAE, que incluye registros de EEG a nivel de cuero cabelludo y EEG profundo⁸. Este análisis solo se centrará en los registros a nivel de cuero cabelludo. La novedad de este estudio es el uso de redes convolucionales para analizar la representación visual del EEG, simulando la inspección clínica de un experto.

Set de datos

La base de datos de *CHB-MIT Scalp EEG* contiene registros de EEG a nivel de cuero cabelludo de 23 pacientes pediátricos, admitidos en el Boston Children's Hospital. El montaje usado fue el 10-20 estándar, a 256 Hz y 16 bits. Los registros fueron etiquetados por expertos, indicando el comienzo y el fin de las crisis. De los 24 casos, se obtienen 941,6 h de actividad inter-ictal y 3 h de actividad ictal, de 181 crisis registradas.

La base de datos europea de epilepsia⁹ *EPILEPSIAE* contiene registros de 275 pacientes. Los datos incluyen registros superficiales y profundos. Los registros superficiales se encuentran en las configuraciones estándar 10-20 y 10-10. Para este estudio se seleccionan los registros de pacientes que cuenten simultáneamente con EEG intracaneal y EEG a nivel de cuero cabelludo, con la misma configuración que los datos de los pacientes del CHB-MIT. Dicho filtro resulta en el análisis de 10 pacientes, con 1302,2 h de actividad inter-ictal y 2,5 h de actividad ictal, de 109 crisis registradas.

Variables utilizadas

Al ser un modelo de Fully Convolutional Networks (FCN), se tiene un análisis end-to-end, por lo que no hubo selección manual de las variables involucradas.

Modelo

Para poder aplicar el modelo, se representó las señales de una forma visual, creando una imagen 2D, en donde las filas corresponden a los distintos canales del registro EEG y las columnas a la dimensión temporal. Se separó el registro en ventanas de tiempo de 4 segundos cada una y luego se codificó cada píxel en escala de grises según la amplitud del registro.

⁷ Publicado en diciembre del 2020.

⁸ Intracraneal.

⁹ European Epilepsy Database en inglés.

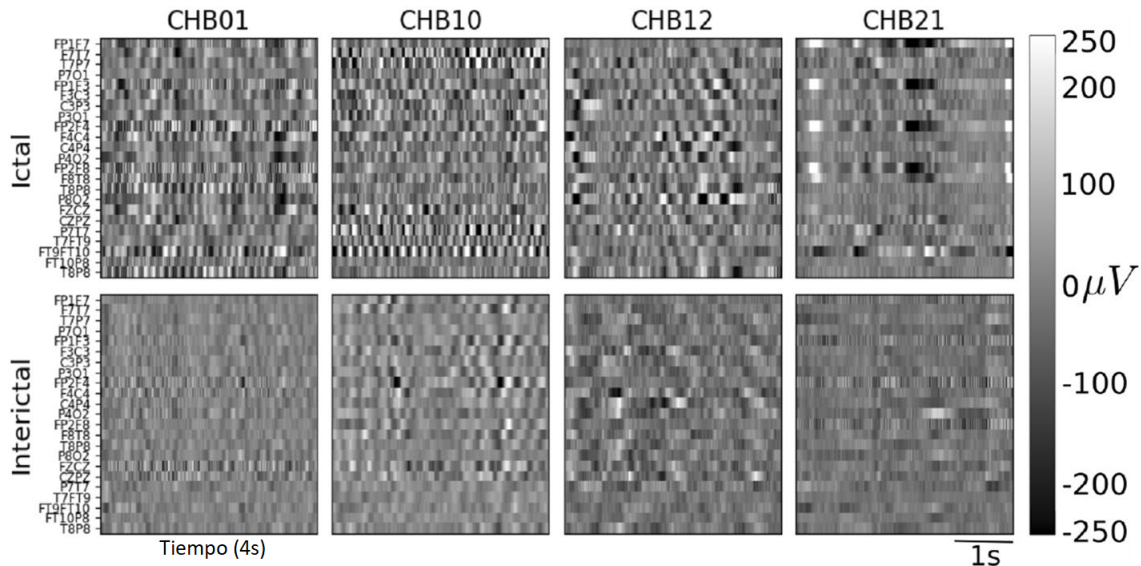


Figura 2.1: Ejemplos de la representación visual de la señal tomando la amplitud como atributo del píxel para cuatro pacientes del conjunto de datos del CHB-MIT. En vertical se codifica los canales en configuración bipolar, y en horizontal la longitud de la ventana (4s) con la relación de aspecto ajustada para su visualización. La barra de color muestra el rango de amplitud de las señales en μV . Fila superior: ejemplos ictales, fila inferior: ejemplos interictales. Fuente: Gómez et al.[17].

Se exploran varias configuraciones de FCN usando validación cruzada triple (3FCV¹⁰), y con la mejor configuración, se entrenó y evaluó el modelo Leave-One-Patient-Out (LOO¹¹) para evaluar la capacidad del modelo para detectar convulsiones en datos de pacientes no vistos.

Desempeño

La mejor configuración del modelo es una FCN con 3 bloques, 128 filtros y capas de abandono con probabilidades de 0,1 y 0,3 en las primeras y últimas capas convolucionales, respectivamente.

Para todos los pacientes, se obtienen niveles globales de precisión y especificidad de $92,9 \pm 21,8$ y $93,1 \pm 21,9$ respectivamente. Además, se alcanzan valores globales de precisión de $51,4 \pm 34,1$, de recall de $53,1 \pm 25,5$ y de F-score de $46,6 \pm 31,0$.

El mejor modelo alcanza un promedio de precisión y especificidad de 99,3% y 99,6% respectivamente.

Conclusiones

Las ventajas de usar un enfoque de aprendizaje profundo al clasificar crisis epilépticas, radica en que no hay necesidad de extraer ni estimar variables, requiriendo poco preprocesa-

¹⁰ Separando los pacientes en 3 sets, usando 2 de ellos para entrenar y el restante para evaluar, cubriendo las 3 combinaciones posibles.

¹¹ Se entrenan 24 modelos, donde cada uno usa datos de $k - 1$ pacientes, con k número total de pacientes.

miento de la señal EEG pura. En este caso, se presenta un enfoque novedoso, con rendimiento igual o superior a trabajos previos, lo que abre una nueva arista para abordar el problema de clasificación.

Finalmente, la inclusión de registros intracraneales además de los superficiales mejora todas las métricas para algunos pacientes, lo que sugiere que ambos tipos de grabaciones llevan información complementaria que es necesaria para una mejor detección de los ataques epilépticos.

2.7. Struck, et al. (2017)

Se presenta un sistema de estimación de riesgo de crisis epilépticas a partir de elementos del registro EEG y la historia clínica de los pacientes. Al modelo se le llama 2HELPS2B[44].

Set de datos

Se utilizó datos de varios centros (Emory University Hospital, Brigham and Women's Hospital y Yale University Hospital) con variables clínicas y electrográficas de 5427 sesiones de registro de pacientes con EEG continuo por indicaciones clínicas, excluyendo los ingresos por epilepsia.

Modelo y variables utilizadas

El modelo consiste en un sistema de puntuación para estimar el riesgo de convulsiones en pacientes agudos sometidos a EEG continuo. Este modelo se construye utilizando un nuevo método de aprendizaje automático (RiskSLIM¹²[50]) que está diseñado para producir sistemas de puntuación precisos y calibrados en función del riesgo con un número limitado de variables.

Se evalúa un set de 24 variables candidatas a ser incluidas, seleccionadas a partir de su prevalencia en la base de datos y su relación con las crisis epilépticas según estudios previos. Se evaluó la precisión y la calibración del riesgo de los modelos con distintas combinaciones y números de variables, contrastando los resultados de RiskSLIM con modelos de regresión logística penalizada (PLR). Se valida el rendimiento de todos los modelos mediante una validación cruzada 5 iteraciones (5-CV).

El modelo resultante incluye 6 variables, en donde 5 de ellas son derivables del estudio del EEG, y una de ellas es referente al historial clínico del paciente (historial de crisis previas).

Se define un sistema de puntajes, en el que cierto puntaje corresponde a la probabilidad de una crisis epiléptica durante las siguientes 72 horas.

¹² Del inglés Risk-Calibrated Supersparse Linear Integer Model.

Resultados

Variable	Puntaje Total						
	0	1	2	3	4	5	>6
Riesgo probable de crisis, %	5	12	27	50	73	88	>95
Prevalencia real de crisis, % (IC 95%)	3 (2-3)	12 (10-13)	34 (31-37)	52 (46-57)	71 (63-78)	84 (71-99)	92 (77-100)

Tabla 2.1: Comparativa puntuación de riesgo optimizada.

Los riesgos de crisis para cada posible puntuación de 2HELPS2B son: 5 % para una puntuación de 0, 12 % para una puntuación de 1, 27 % para una puntuación de 2, 50 % para una puntuación de 3, 73 % para una puntuación de 4, 88 % para una puntuación de 5 y más del 95 % para una puntuación de 6 o 7. En la Tabla 2.1 se comparan con la prevalencia real de las crisis, con un intervalo de confianza al 95 %.

Conclusiones

El sistema de 2HELPS2B se presenta como una herramienta de uso sencillo para ayudar a los equipos médicos a evaluar el riesgo de los pacientes de tener crisis epilépticas. Sin embargo, en su intento de simplificar omite ciertos factores que pueden ser de relevancia a la hora de predecir o identificar una crisis, como lo es la duración de las mismas. El propósito de este estudio es facilitar la determinación y comunicación del riesgo entre equipo clínico, con un enfoque predictivo, y no la clasificación o detección de las crisis en sí. Además, las variables clínicas que se requieren para obtener el puntaje requieren de la intervención de un especialista capaz de identificarlas. Estas consideraciones hacen que no sea comparable con los demás resultados presentados en esta revisión, sin embargo, al usar historial clínico tener un enfoque predictivo, aporta una nueva perspectiva para abordar el problema.

2.8. Soluciones comerciales

Además de los trabajos recién expuestos, es posible encontrar en el mercado diferentes alternativas que se ofrecen como solución al problema de detección de crisis epilépticas. Existen pocos trabajos en los que se evalúe sistemáticamente y se compare los rendimientos del software disponible comercialmente. Del trabajo de Koren et al.[26] se puede tener una referencia en cuanto al nivel de detección de crisis, la cantidad de falsas alarmas por hora y los tiempos de respuesta de algunos softwares.

Para la comparación se utilizó un set de datos compuesto por 81 pacientes seleccionados aleatoriamente, cuyos registros de EEG están acompañados además por registro de video-monitoreo.



Figura 2.2: Software comparado por Koren et al. [26]

Los paquetes comparados son:

- Besa 2.0 (*BESA GmbH, Gräfelfing, Alemania*), usa técnicas de análisis de poder espectral. Se presenta en paquete con el software de detección de picos de Besa. La licencia para el conjunto tiene un precio de aproximadamente \$11.200.000[31] (CHF 13.130)
- Encevis 1.7 (*AIT Instituto Austriaco de Tecnología GmbH, Viena, Austria*), que hace un análisis en los dominios de tiempo y frecuencia.
- Persyst 13 (*Persyst Development Corporation, California*), usa redes neuronales avanzadas.

Software	Sensibilidad	Ratio de Falsas Alarmas	Retraso en detección	Precio
Besa 2.0	67.6 %	0.7/h	30s	\$11.200.000
Encevis 1.7	77.8 %	0.2/h	25s	-
Persyst 13	81 %	0.9/h	20s	-

Tabla 2.2: Comparativa de alternativas comerciales. Con datos del trabajo de Koren et al.[26]

Si bien todos muestran rendimientos similares, sin ser estadísticamente distintos, con sensibilidades relativamente altas para el mismo dataset, difieren en el ratio de falsas alarmas por hora y el tiempo de retraso en la detección. Persyst presenta (Figura 2.2) la mejor tasa de detección de crisis, pero también la mayor cantidad de falsas alarmas por hora, mientras que Encevis tiene una sensibilidad levemente menor, pero el menor ratio de falsas alarmas de la muestra.

Es relevante conocer las capacidades de productos presentes en el mercado para conocer con mayor precisión el efecto de incorporar una alternativa, en particular para el INCA, tanto en capacidad diagnóstica como en costos o ahorros eventuales.

2.9. Resumen y conclusiones

De la revisión del estado del arte, se reconoce que el área de estudio presenta un desarrollo avanzado, existiendo amplia variedad de modelos y sistemas de análisis enfocados en la clasificación de crisis epilépticas. Estos se distinguen principalmente según el tipo de clasificación, el tipo o fuente de los datos utilizados y en el propio método de clasificación.

Sobre el tipo de clasificación, los modelos se diferencian entre los que buscan hacer clasificación binaria (crisis epiléptica, no crisis epiléptica), los que hacen clasificación ternaria (estado normal, ictal, interictal) y los que hacen clasificaciones más amplias, como el de Golmohammadi et al., en el que se identifican tres patrones de interés clínico y otra actividad de fondo.

Sobre el tipo de dato, se distinguen entre los que utilizan set de datos de referencia y los que utilizan datos de su propia experiencia clínica. De los modelos estudiados, son varios los que utilizan el set de la Universidad de Bonn. Dicho dataset contiene muestras de voluntarios y pacientes epilépticos en distintos estados, y obtenidos con diferentes métodos, siendo algunos de ellos registros a nivel de cuero cabelludo, y otros de tipo intracraneal. El uso de dataset de referencia es una herramienta conveniente a la hora de comparar los resultados entre diferentes modelos, y se presenta como una alternativa ante la dificultad de obtener datos propios a nivel clínico. Por otra parte, la inclusión de registros intracraneales demostró una mejora en algunos modelos[17], lo que sugiere la conveniencia de usarlos para complementar los registros superficiales[36].

Otro elemento que distingue los set de datos es la edad de los pacientes estudiados. La mayoría emplea datos recopilados en adultos, sin embargo, existen algunos clasificadores neonatales. Las diferencias morfológicas entre dichos grupos, hacen que no sea posible generalizar los resultados obtenidos en alguno de ellos al restante, pero pueden orientar respecto a qué factores, métodos y características de la señal usar.

Sobre los modelos de clasificación, difieren en las técnicas utilizadas y los niveles de complejidad de procesamiento y computación. Los más simples usan técnicas de SVM, mientras que otros más complejos usan técnicas de análisis de señales modernas o modelos de múltiples capas y parámetros. Otro elemento en el que varían es la selección de variables. Existe amplia variedad de métodos y técnicas de extracción de variables, pudiendo ser estas extraídas directamente[46], o derivadas de la señal registrada[16][53]. También se revisa modelos que no involucran la selección de variables, siendo su análisis de tipo caja negra, o end-to-end.

En general no se aprecia una relación entre complejidad del procesamiento de los datos y el rendimiento de los modelos.

Método	Tipo de clasificación	Tipo y set de dato	Método de clasificación	Accuracy	Recall	Precision
Temko et al.	Binario	Clínico: UCI Neonatal	SVM	-	-	89%
Zhang et al.	Ternario	No clínico: Set de Bonn	RF	97.35%	-	-
Golmohammadi, et al.	De 6 tipos	Mixto: Corpus de TUH	SdA	-	90%	-
Song et al.	Binario	No clínico: Set de Bonn	ELM	99%	-	-
Ullah et al.	Ternario	No clínico: Set de Bonn	CNN	99.1%	-	-
Gómez et al.	Binario	No clínico: CHB/EPILEPSIAE	FCN	99.3%	-	-
Yuan et al.	Binario	No clínico: Set de Bonn	ELM	96.5%	-	-

Tabla 2.3: Comparativa de los principales resultados de los trabajos estudiados.

En cuanto al desempeño de los modelos estudiados, según se aprecia en la tabla 2.3, todos ellos presentan rendimientos en torno al 90% o más, a la hora de identificar crisis epilépticas. Se aprecia gran variabilidad en las métricas de evaluación para los diferentes estudios y set de datos, lo que muestra la complejidad a la hora de definir modelos que puedan generalizar de buena manera. En particular, una de las principales complejidades apreciadas, es la de obtener buena generalización de los modelos para los pacientes que no han sido visto antes.

Dentro de la presente revisión del estado del arte no se encuentra ningún modelo de clasificación que incluyera otros factores de historial clínico o demográficos de cada paciente en el estudio. Sin embargo, sí se estudia un modelo con enfoque predictivo, que utiliza la existencia de crisis previas como variable relevante, lo que sugiere su utilidad como variable a incluir en modelos de clasificación.

Como ya se ha mencionado previamente, es posible obtener una mejora al complementar los registros de EEG con EEG intracranial. El presente trabajo pretende plantear una herramienta de uso sencillo, por lo que los registros a utilizar, y por tanto, la extensión de los resultados, se limita a EEG a nivel de cuero cabelludo, método poco invasivo y que no requiere de una mayor intervención del paciente.

En los objetivos del presente estudio, se plantea detectar actividad epileptiforme en el servicio de pacientes críticos del INCA. Buena parte de la literatura revisada consiste en registros hechos en condiciones ideales, ajenas al ruido o artefactos que pueden existir en las UCI, o en el SPC del INCA en particular. Existe, por otra parte, diversos mecanismos para remover dichos elementos[21] que podrían ser utilizados. Los últimos estudios revisados sugieren que no es necesario preprocesar[17] los registros para reducir artefactos o ruidos y obtener buenos resultados. En este trabajo se procesa los datos crudos, sin remoción de artefactos, considerando que se apunta a la aplicación del modelo resultante en UCI clínica real.

Respecto de la utilización o no de modelos de tipo caja negra, y previa consulta al Dr. Brunetti, no se considera como una ventaja o un inconveniente su uso. Los mejores resultados de la revisión corresponden a modelos de caja negra, lo que los posiciona como la primera opción para llevar a cabo el análisis de los datos obtenidos. Sin embargo, considerando la variedad de resultados que fue posible observar, no se limita el estudio solo a modelos de este tipo.

Capítulo 3

Análisis de datos

Si bien en primera instancia se planteó entrenar los modelos directamente con registros del INCA, debido a la disponibilidad de datos y extensión de los registros se decide replantear esta metodología. De este modo, se entrena los modelos con una bases de datos disponibles públicamente y se utilizará los registros del INCA para contrastar y evaluar resultados.

3.1. Set de datos

La base de datos a utilizar para entrenar los modelos es la de CHB-MIT Scalp EEG[41]. Esta contiene registros de EEG a nivel de cuero cabelludo de 23 casos, correspondientes a 22 sujetos (5 hombres, de 3 a 22 años; y 17 mujeres, de 1,5 a 19 años), admitidos en el Boston Children's Hospital. El montaje usado fue el 10-20 estándar, y las señales se muestrearon a 256 Hz con 16 bits de profundidad. Los registros fueron etiquetados por expertos, indicando el comienzo y el fin de las crisis. La señal de los registros es de tipo multicanal de 23 canales, y en algunos casos incluyen, además de los canales propios del EEG, el registro de electrocardiograma de los pacientes, monitores del movimiento de los ojos y relativos a la respiración, entre otros. Los canales relevantes son los 23 señalados.

Por otra parte, para abastecer a los modelos de datos más diversos, se recolecta registros del repositorio electroencefalográfico público del TUH[33].

Lo anterior significa contar con aproximadamente 3 horas de registros, de las cuales dos tercios corresponden a registro normal (sin crisis), y el tercio restante a registros con crisis epilépticas. Estos datos corresponden a registros de de un total de 35 pacientes.

3.2. Exploración de registros

Los registros en estudio presentan diferentes frecuencias de muestreo, fluctuando entre 250Hz hasta 1000Hz, por lo que se procede a ejecutar rutinas de remuestreo de las señales. Se ejecutan rutinas de *upsampling*¹ y de *downsampling*² a través de los métodos de Fourier y de interpolación lineal.

El upsampling y el downsampling son operaciones que cambian la frecuencia de muestreo de una señal, y cada uno de ellos se compone de dos pasos, en los que se cambia la frecuencia de muestreo y se filtra algunas frecuencias de la misma señal. En el caso del downsampling, al pasar de una frecuencia F a una frecuencia F_d , con $F_d < F$, se está reduciendo el ancho de banda disponible, por lo que en caso de mantener cualquier frecuencia $f > \frac{F_d}{2}$, se obtendría *aliasing*³. Por lo anterior, es necesario aplicar un filtro pasa-bajo a la señal original F que excluya las frecuencias que puedan causar aliasing. En el caso del upsampling, se pasa de una frecuencia F a una señal de frecuencia F_u , con $F_u > F$. Para esto, se inserta silencio entre las muestras de la señal original, y luego se aplica una interpolación para completar los valores de la señal durante los silencios.

La interpolación lineal se utiliza para estimar los puntos faltantes de una muestra, suponiendo que la señal se comporta como una línea recta entre dos puntos ya conocidos. La interpolación de Fourier transforma la señal al dominio de la frecuencia. Al hacer down-sample, se eliminan elementos con frecuencias intermedias y se transforma nuevamente al dominio del tiempo. Mientras que al hacer up-sample, se añade silencios al final y se vuelve a transformar al dominio del tiempo.

A continuación se procede a la separación de los canales de EEG y a la segmentación del registro en ventanas temporales de largo arbitrario, para la extracción de características. Para hacer un análisis cuantitativo, se utiliza como referencia variables relevantes útiles en clasificaciones de señales en neonatos[46], analizando así la señal en el dominio del tiempo para cada una de esas ventanas, la transformación al dominio de la frecuencia para extraer sus características y las variables asociadas a la teoría de la información. En primera instancia se genera las rutinas de importación de datos que permitan la carga masiva para el entrenamiento de los modelos, validar la utilidad de los registros y determinar la suficiencia de la cantidad de datos requeridos para continuar con el estudio.

Una vez extraídas las variables, es posible explorar visualmente si existe o no alguna característica que distinga a los dos tipos de registros que se busca clasificar. Cómo se aprecia en la Figura 3.1, existen variables para las cuales hay una evidente separación entre los registros de crisis y los normales.

¹ Remuestreo en el que la señal generada tiene mayor frecuencia que la señal original.

² Remuestreo en el que la señal generada tiene menor frecuencia que la señal original.

³ El aliasing es una distorsión que hace que diferentes señales se vuelvan indistinguibles entre sí cuando se muestrean.

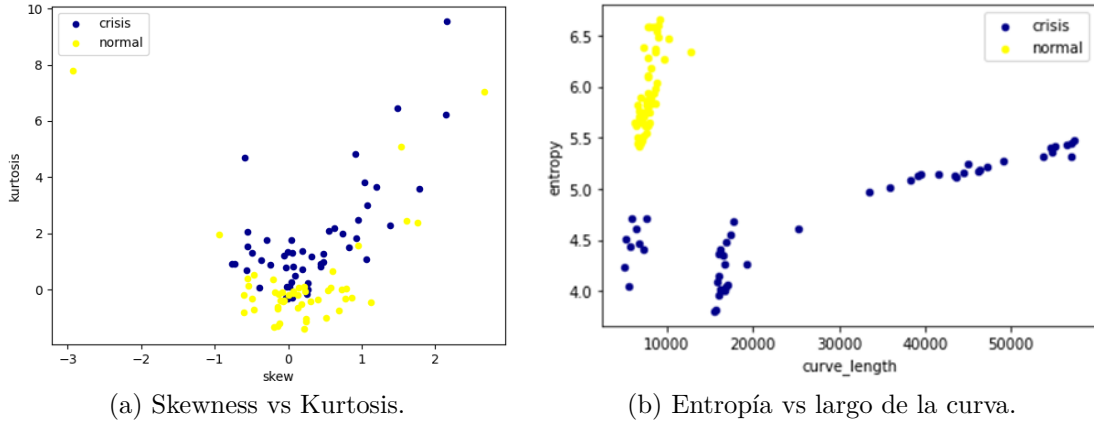


Figura 3.1: Comparativa entre variables derivadas de cada ventana temporal. Elaboración propia.

En particular en (a) se observa que la relación entre la asimetría de la señal (skewness) y distribución entre la media y las colas (kurtosis), en dónde ambos tipos de registros parecen ser igualmente asimétricos, sin embargo los registros de crisis presentan una kurtosis más elevada, lo que puede ser signo de anomalías en la señal. Por otra parte, en (b) es posible observar que los segmentos normales poseen una entropía, o grado de desorden, más elevada que los segmentos correspondientes a crisis. Esto es esperable, ya que las descargas epileptiformes que caracterizan a una crisis tienden a ser repetitivas, periódicas y estructuradas en patrones⁴.

3.3. Modelos de clasificación

Como se ha señalado en secciones anteriores, el procedimiento previo a la implementación de modelos de clasificación corresponde, como primer paso, a la separación de los registros entre set de entrenamiento del modelo y set de test. Lo anterior, procurando que registros de un mismo paciente no aparezcan más de una vez en el mismo set, para evitar *data leakage*⁵. El siguiente paso del proceso corresponde a la separación en canales, junto a la segmentación en ventanas temporales.

3.3.1. Support Vector Machine (SVM)

Como el modelo de SVM requiere de un set de variables, se procede a procesar cada una de las ventanas temporales descritas anteriormente para obtener las variables identificadas como útiles en la clasificación de señales en neonatos[46].

En el dominio del tiempo se obtienen la cantidad de mínimos y máximos de la señal en el segmento, el número de veces que la señal cruza el cero, el valor eficaz (RSM), skewness y kurtosis, además de los parámetros de Hjorth (actividad, movilidad, y complejidad), la varianza, el largo de la curva y la energía no lineal. Además, se obtiene la varianza para la primera y

⁴ Ver Anexos: Criterios para crisis no convulsivas.

⁵ En español *fuga de datos*, corresponde a la sobreestimación de la capacidad de un modelo al poner a su disposición información que no debería estar a su alcance. En este caso, se quiere evitar que parte de los datos con los que está testeando estén incluidos en el entrenamiento.

segunda derivada de la señal, así como el número de veces que dichas derivadas cruzan el cero.

En cuanto a la frecuencia, usando la transformada rápida de Fourier se obtiene el power spectral density (PSD). Luego, a partir de ello el poder total de la ventana, y la frecuencia de pico de potencia. Finalmente, en cuanto a teoría de la información, se registra la entropía de Shannon.

Lo anterior, en suma, corresponde a 19 variables numéricas para cada ventana temporal, con lo que se ejecuta clasificación binaria sobre las clases, etiquetadas como crisis (con el número 1) y normal (con el número 0). Se utiliza la clasificación de Support Vector del paquete de Scikit-learn, con un kernel de función de base radial (RBF) y un parámetro de regularización de $C = 0.8$. Dicho valor se fija por obtener los mejores resultados sobre la data disponible, luego de haber probado con valores de C entre $0.5 - 1$. Se explora además el resultado del modelo remuestrando la señal, tanto con upsampling como con downsampling.

3.3.2. Red Neuronal Convolutiva de una dimensión (CNN-1D)

Se plantea usar una CNN para ejecutar la clasificación. En este caso, las ventanas temporales descritas anteriormente serán procesadas directamente por la red, sin necesidad de extraer variables de ellas. Aparte de la señal en crudo, se tiene la etiqueta de clase, respecto a si el segmento de señal corresponde o no a una descarga epileptiforme. Para esto, se utiliza un modelo CNN de Keras, que requiere de un ajuste en la dimensionalidad de la data para hacerla coincidir con la que es capaz de recibir el modelo.

El modelo planteado tiene una arquitectura simple, consistente en una capa con 128 filtros, un kernel de tamaño 5 y una función de activación ReLU, a la cual se aplica un dropout⁶ de 0.25 como método de regularización para reducir el sobreajuste y evitar errores de generalización. El resultado se aplana y pasa a una capa densa de salida, con función de activación sigmoide.

Finalmente, se usa la función de pérdida logarítmica (*binary_crossentropy*) durante el entrenamiento junto al algoritmo de optimización de Adam para el descenso del gradiente. En una tarea de clasificación el límite de decisión en una tarea de clasificación es grande (respecto a, por ejemplo, una regresión). Dicha función de pérdida castiga las clasificaciones erróneas de un mejor modo que métodos funciones como el error cuadrático medio para este tipo de problemas. Se utiliza como métricas para la optimización *accuracy* y *recall*.

3.3.3. Red Neuronal Recurrente (RNN)

Esta clase de redes se usan para analizar datos de series temporales, tal como lo son las señales de EEG. Al igual que en el modelo anterior, la clasificación se ejecuta sobre la señal cruda. En este caso, para el modelo se usa una arquitectura específica, que es la de long short-term memory (LSTM). Para que el clasificador pueda capturar la evolución de los patrones de crisis (o de los registros interictales), la red es bidireccional. Se aplica un dropout de 0.25 y la capa de salida utiliza la función de activación sigmoide. La función de pérdida

⁶ Eliminación de nodos aleatorios durante el entrenamiento.

usada durante el entrenamiento es binary crossentropy junto al algoritmo de optimización de Adam. Se contrasta resultados usando como métricas para la optimización *accuracy* y *recall*.

3.3.4. Red Neuronal Convolutiva de dos dimensiones (CNN-2D)

Para poder aplicar el modelo, se representó cada una de las ventanas temporales anteriormente descritas de una forma visual, similar a lo planteado por Gómez et al.[17] creando una imagen 2D, en donde las filas corresponden a los distintos canales del registro EEG y las columnas a la dimensión temporal. Se codificó cada píxel en escala de grises según la amplitud del voltaje del registro.

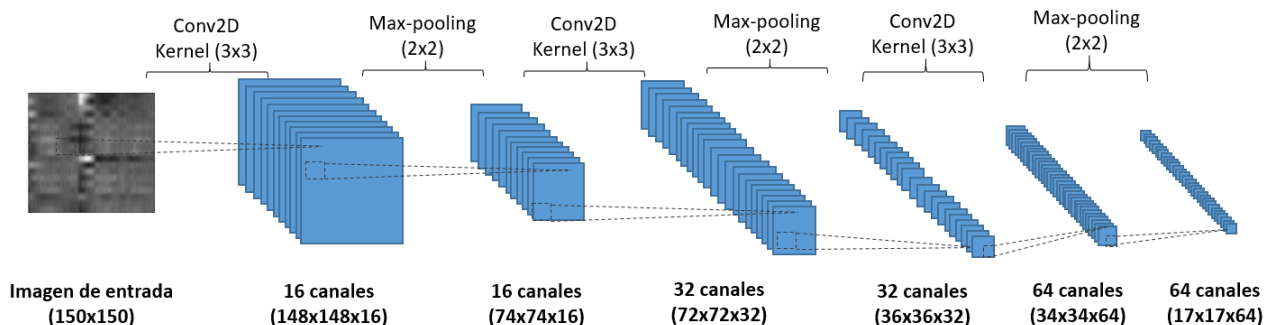


Figura 3.2: Esquema de las capas convolucionales del modelo CNN. Elaboración propia.

En la Figura 3.2 se observa el flujo de información desde la imagen de entrada, con la consiguiente generación de imágenes filtradas y la reducción dimensional a través del max-pooling. La estructura de la red es como se muestra a continuación, en la Tabla 3.1. Para la entrada a la red, se ajustan las imágenes de modo de tener entradas de 150×150 píxeles. Se utilizó max-pooling de 2×2 para reducir la dimensionalidad de las imágenes resultantes de las convoluciones.

Capa (tipo)	Dimensión de salida	# Parámetros
input_1 (InputLayer)	[(None, 150, 150, 3)]	0
conv2d (Conv2D)	(None, 148, 148, 16)	448
max_pooling2d (MaxPooling2D)	(None, 74, 74, 16)	0
conv2d_1 (Conv2D)	(None, 72, 72, 32)	4640
max_pooling2d_1 (MaxPooling2D)	(None, 36, 36, 32)	0
conv2d_2 (Conv2D)	(None, 34, 34, 64)	18496
max_pooling2d_2 (MaxPooling2D)	(None, 17, 17, 64)	0
flatten (Flatten)	(None, 18496)	0
dense (Dense)	(None, 512)	9470464
dense_1 (Dense)	(None, 1)	513

Tabla 3.1: Estructura de la red neuronal utilizada.

El tamaño del kernel usado es de 3×3 , y se utiliza como función de pérdida (*binary crossentropy*) durante el entrenamiento.

3.4. Resultados y discusión

Se ejecuta la clasificación usando los distintos modelos, y se registran métricas relevantes para medir su desempeño. Las métricas usadas son el área bajo la curva (AUC), accuracy, recall, precision y specificity. Como se ha mencionado en capítulos previos, por las características del problema estudiado, tiene mayor relevancia tener un mejor *recall*, de modo de identificar la mayor cantidad de crisis posibles.

Se evalúa los distintos modelos utilizando diferentes métodos de remuestreo y funciones de optimización. Para el downsample se usa como frecuencia objetivo la menor frecuencia entre los registros, en este caso 250Hz. Para el upsample se usa la mayor frecuencia de los registros utilizado, en este caso 1000Hz. Cada uno de esos remuestreos se ejecuta a través de los dos métodos mencionados anteriormente: método de Fourier e interpolación lineal.

Otro de los elementos que se prueba es variar la función de optimización, probando para cada modelo una optimización basada en accuracy y además una en recall.

Fn opti.	Muestra	Modelo	AUC	accuracy	recall	precision	specificity
accuracy	Downsample	CNN 1 D	0,578	0,562	0,588	0,183	0,558
accuracy	Downsample	Dense NN	0,761	0,674	0,711	0,265	0,668
accuracy	Downsample	LSTM	0,85	0,765	0,791	0,358	0,761
accuracy	Upsample	CNN 1 D	0,448	0,732	0,344	0,222	0,797
accuracy	Upsample	Dense NN	0,672	0,676	0,572	0,239	0,694
accuracy	Upsample	LSTM	0,833	0,814	0,71	0,415	0,831
Recall	Downsample	CNN 1 D	0,542	0,559	0,535	0,171	0,563
Recall	Downsample	Dense NN	0,699	0,787	0,503	0,34	0,835
Recall	Downsample	LSTM	0,828	0,781	0,767	0,373	0,783
Recall	Upsample	CNN 1 D	0,481	0,739	0,387	0,244	0,798
Recall	Upsample	Dense NN	0,588	0,718	0,385	0,223	0,774
Recall	Upsample	LSTM	0,875	0,696	0,872	0,316	0,653
accuracy	Downsample	CNN 2 D	0,673	0,846	0,613	0,326	0,681
	Upsample	SVM	0,58	0,58	0,41	0,23	0,62
	Downsample	SVM	0,62	0,669	0,401	0,281	0,738

Tabla 3.2: Comparativa de los resultados de los modelos aplicados. Remuestreo hecho a través de interpolación lineal.

Como se observa en la Tabla 3.2 los resultados son variados y, para un mismo modelo, no necesariamente homogéneos según las distintas métricas. Como se señala previamente, la métrica de interés para el problema estudiado es el recall.

Pese a la cantidad de parámetros obtenidos, el modelo de SVM muestra ser el de desempeño más pobre entre los estudiados, sobretodo considerando la métrica de interés. De los resultados se observa que el modelo de LSTM es el que obtiene los mejores resultados en esa métrica (entre 0,7 y 0,8), independiente de la función de optimización o el método de remuestreo (ver Tabla 3.3). Además, dichos modelos muestran resultados de accuracy en torno a 0,7. Por su parte, el modelo de redes convolucionales en 2 dimensiones es bastante efectivo

en su accuracy, pero no consigue destacarse en la métrica de interés para el problema. Los resultados de los demás modelos varían según la métrica o método de remuestreo usado, por lo que no es posible determinar su rendimiento global. Sí es preciso concluir que presentan resultados de recall menos efectivos que el modelo de LSTM.

Fn opti.	Muestra	Modelo	AUC	accuracy	recall	precision	specificity
accuracy	Downsample	CNN 1 D	0,553	0,605	0,517	0,186	0,620
accuracy	Downsample	Dense NN	0,592	0,783	0,323	0,280	0,860
accuracy	Downsample	LSTM	0,845	0,774	0,815	0,371	0,767
Recall	Downsample	CNN 1 D	0,544	0,664	0,422	0,194	0,705
Recall	Downsample	Dense NN	0,581	0,705	0,356	0,202	0,764
Recall	Downsample	LSTM	0,822	0,781	0,705	0,365	0,793

Tabla 3.3: Comparativa de los resultados de los modelos aplicados. Remuestreo hecho a través de Fourier.

Por otra parte, es destacable que el resultado de los modelos remuestreados con interpolación lineal presentan mejores resultados en general que aquellos con remuestreo de Fourier. En particular, las métricas de recall y precision se ven mejoradas al usar interpolación lineal, mientras que accuracy y specificity son mejores al usar el método de Fourier.

Se propone otra aproximación al problema, para estudiar qué sucede al entregar a los modelos una señal procesada en vez de la señal en crudo. En particular, se estudia los resultados al aplicar la transformada rápida de Fourier a la señal previo al entrenamiento de los modelos.

Fn opti.	Método	Sample	Modelo	AUC	accuracy	recall	precision	specificity
accuracy	Lineal	Downsample	Dense NN	0,555	0,628	0,321	0,144	0,68
accuracy	Lineal	Downsample	CNN 1 D	0,622	0,721	0,376	0,223	0,779
accuracy	Lineal	Downsample	LSTM	0,488	0,287	0,714	0,133	0,215
Recall	Lineal	Downsample	Dense NN	0,558	0,604	0,367	0,148	0,644
Recall	Lineal	Downsample	CNN 1 D	0,622	0,807	0,206	0,275	0,909
Recall	Lineal	Downsample	LSTM	0,49	0,236	0,749	0,134	0,144
accuracy	Lineal	Upsample	Dense NN	0,528	0,607	0,348	0,144	0,651
accuracy	Lineal	Upsample	CNN 1 D	0,608	0,78	0,302	0,267	0,861
accuracy	Lineal	Upsample	LSTM	0,482	0,353	0,624	0,132	0,144

Tabla 3.4: Comparativa de los resultados de los modelos aplicados. La señal está procesada usando la transformada rápida de Fourier.

Usando esta aproximación, y como se presenta en la Tabla 3.4, ninguno de los modelos logra mejorar alguna métrica con respecto a las presentadas en la Tabla 3.2, con excepción del modelo de redes convolucionales de una dimensión. Sin embargo, esta mejora en performance se da en todas las métricas excepto la de interés para el problema, que es la de recall.

De este modo, es posible afirmar que el mejor modelo es el de redes recurrentes LSTM, con un recall de hasta un 87%, accuracy de 70% y AUC de 0,875.

Capítulo 4

Prototipo expositivo

Entre los objetivos del presente trabajo se encuentra la elaboración de un prototipo expositivo que permita, a través de una interfaz visual, conocer los resultados de la clasificación realizada por el modelo.

La motivación para dicho desarrollo no es solo la de presentar las capacidades del modelo trabajado, sino que sentar las bases para una aplicación clínica en tiempo real de la solución, etapa que escapa del alcance de esta memoria.

4.1. Esquema propuesto

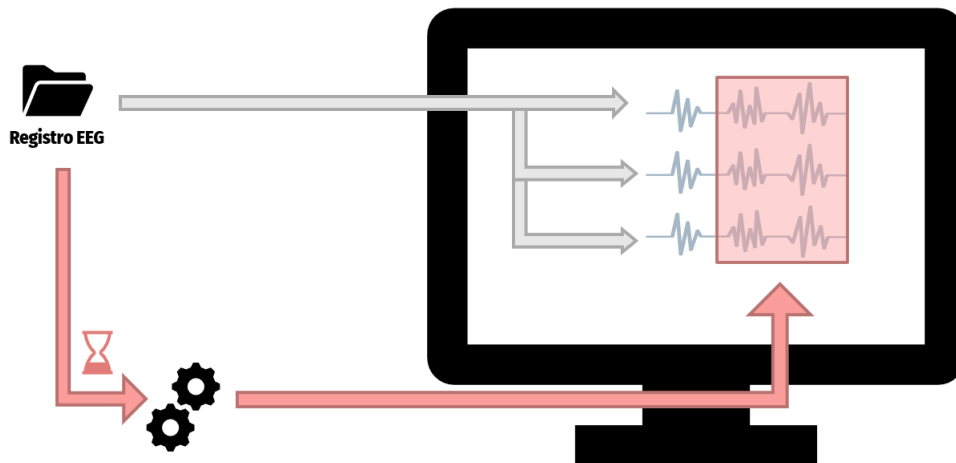


Figura 4.1: Esquema propuesto para el prototipo a realizar.

Se plantea desarrollar una aplicación web, en la que se pueda cargar un registro EEG en el formato EDF¹ que alimente una interfaz visual, en donde se represente la señal de la electroencefalografía en sus respectivos canales, a modo de serie de tiempo. Para ello, se propone

¹ El formato de datos europeo (*European Data Format*) es un formato de archivo estándar diseñado para el intercambio y almacenamiento de series de tiempo médicas.

que el archivo cargado sea transmitido a la interfaz visual en un proceso de data stream², que simule la llegada de datos en tiempo real. En paralelo, y como se diagrama en la Figura 4.1, se propone que el mismo stream de datos se paquelize para ejecutar la clasificación. Esto último, para recolectar la cantidad de muestras necesarias para que el modelo clasificador pueda operar. En principio el tamaño de cada paquete, o batch, se estima en la cantidad de muestras necesarias para completar una ventana temporal³. Una vez ejecutada la clasificación, es decir, que el modelo establezca si la ventana analizada corresponde o no a una crisis, ésta decisión se transporta a la interfaz visual. Allí se superpone visualmente para resaltar el tramo del EEG en el que se identifican crisis de tipo epileptiforme. Finalmente se propone la posibilidad de exportar el resultado, tanto en su modo visual, como en la modalidad de marcas temporales.

Las características anteriormente descritas tienen el potencial de alertar en tiempo real en el caso de una aplicación clínica, para la revisión inmediata del especialista, o de al menos acotar el área de estudio del registro en caso de su revisión posterior. Por otra parte, sujeto a las eventuales condiciones técnicas, permite un análisis remoto del registro, lo que permitiría un monitoreo continuo en UCI a nivel país.

4.2. Mockup propuesto

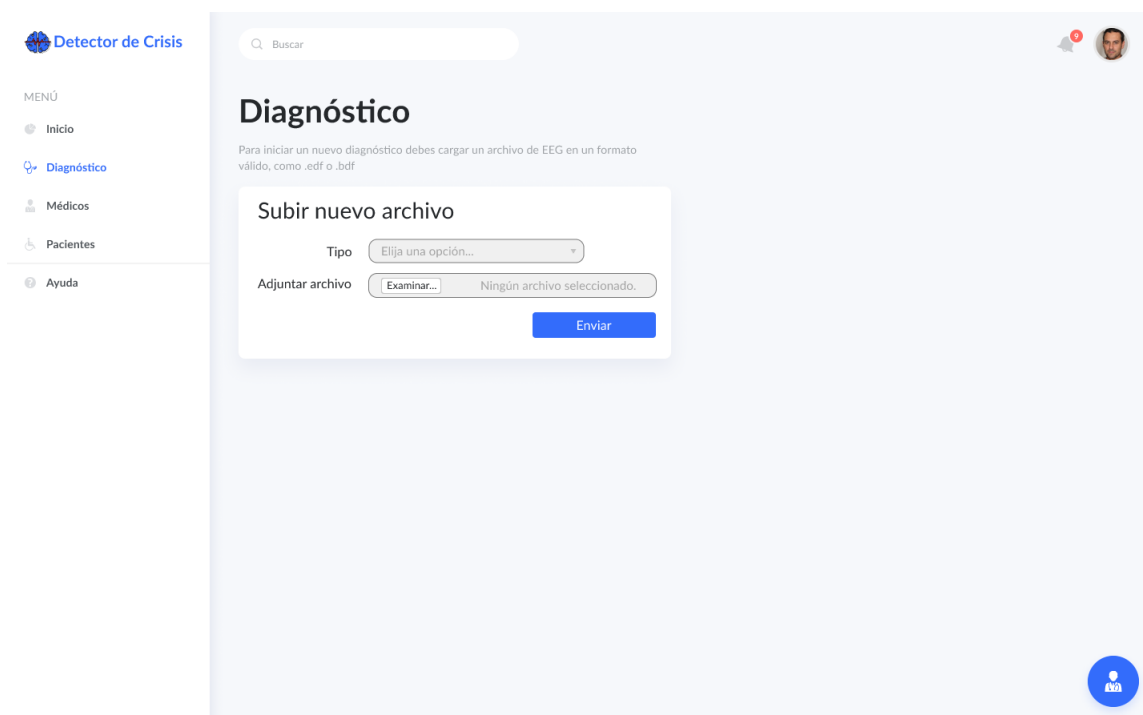


Figura 4.2: Mockup propuesto - Vista de diagnóstico antes de cargar registro EEG.

² En español Flujo de Datos, correspondiente a la llegada continua y coherente de datos que están siendo transmitidos.

³ Ventanas temporales que segmentan el registro, como se menciona en los capítulos previos.

Se desarrolla un mockup⁴ que incorpore las funcionalidades anteriormente descritas. De este modo, tal y como se muestra en la Figura 4.2, una vista o módulo de diagnóstico debe permitir agregar los archivos a estudiar desde el computador del médico. Como no existe solo un tipo de archivos o formatos que codifiquen registros electroencefalográficos es necesario identificar qué tipo de archivo es el que se está cargando a la plataforma. En una fase inicial, se debería poder cargar un archivo a la vez.

Una vez cargado el archivo, y como se muestra en la Figura 4.3, la metadata del paciente asociada al registro EEG se desplegará, de modo que el médico que sea capaz de identificarlo. A la vez, las señales de los distintos canales se graficarán, extendiéndose temporalmente en eje horizontal. El proceso de clasificación, desarrollado en paralelo, entregará los períodos marcados como crisis epilépticas segmentados, señalando sus momentos de inicio y fin, con la posibilidad de seleccionarlos para resaltarlos en la gráfica de las señales en bruto. Finalmente, se podrá exportar el registro de crisis, tanto gráficamente como se observa en la plataforma, como en el formato de anotaciones, señalando las marcas de tiempo relevantes relacionadas a las crisis detectadas. Lo anterior permitirá el trabajo coordinado entre especialista que no necesariamente se encuentren en el mismo lugar, además de la posibilidad de realizar una revisión off-line, es decir, sin estar constantemente revisando la plataforma.



Figura 4.3: Mockup propuesto - Vista de diagnóstico luego de cargar registro EEG.

Además del módulo de diagnóstico, la plataforma debe permitir al usuario (médico) revisar el registro de los archivos cargados previamente, con información relevante que resuma la información obtenida, tales como el nombre del paciente, la fecha y hora en las que se efectuó

⁴ Un mockup o maqueta web se trata de una imagen que muestra el resultado visual esperado de una página web o aplicación.

el registro, el médico a cargo y la cantidad de crisis detectadas por el algoritmo. Lo anterior se presenta en la Figura 4.4, en donde también es posible apreciar el menú de navegación lateral propuesto. Se pretende una interfaz de fácil navegación para el usuario médico, adaptable a dispositivos móviles y de escritorio.

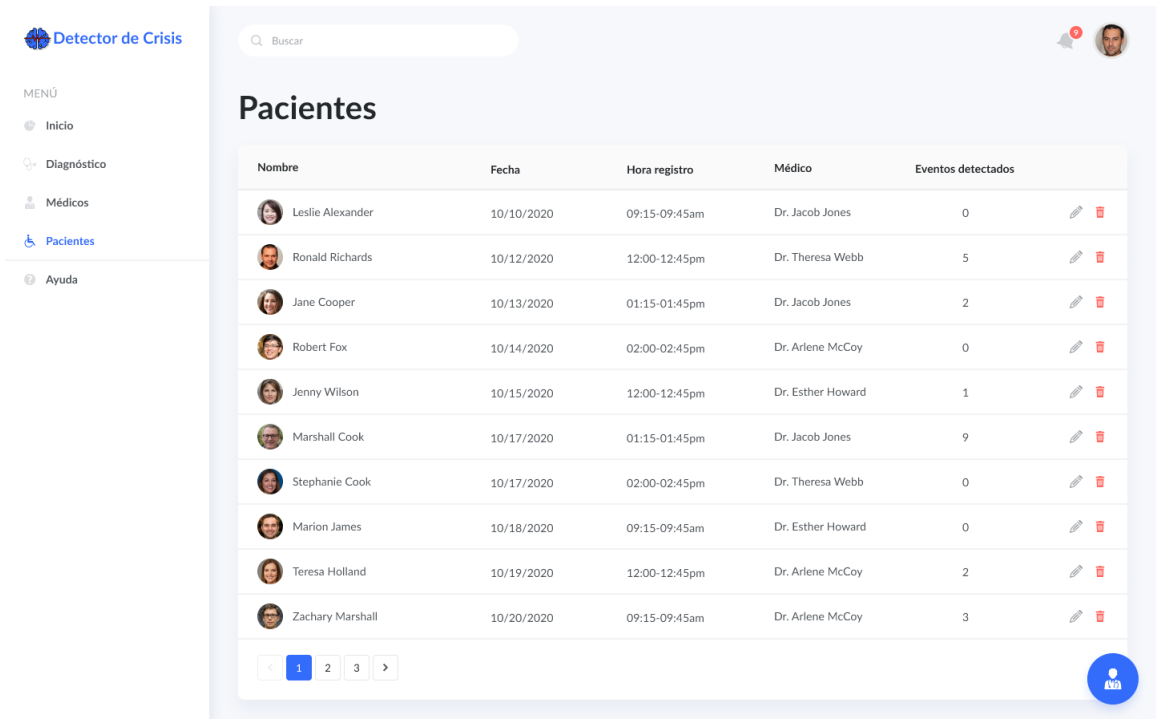


Figura 4.4: Esquema propuesto para el prototipo a realizar.

Las vistas del mockup presentado fueron desarrolladas en Figma⁵.

4.3. Prototipo desarrollado

Según lo anterior, se desarrolla un prototipo con las funcionalidades básicas: recibir un archivo de electroencefalografía, graficarlo y ejecutar la clasificación. Para ello, se crea una aplicación montada en un servidor local basada en Django⁶, con una interfaz de usuario en React⁷.

Este prototipo se limita a trabajar con archivos de tipo .edf. Una vez cargado el archivo, se procede a graficar la señal en bruto. Otra de las limitaciones del prototipo desarrollado es que, pese a que el funcionamiento corresponde al planteado en las partes anteriores, éste se desarrolla solo para el primer canal del registro cargado. De este modo, y como se muestra en la Figura 4.5, la aplicación es capaz de representar el registro EEG junto a regiones del registro que señalan las zonas en las que se identifica una crisis según el modelo.

⁵ Figma es una herramienta de diseño de UI (interfaz de usuario) y UX (experiencia de usuario).

⁶ Django es un framework web escrito en Python.

⁷ React es una biblioteca Javascript de código abierto diseñada para crear interfaces de usuario en desarrollo de aplicaciones.

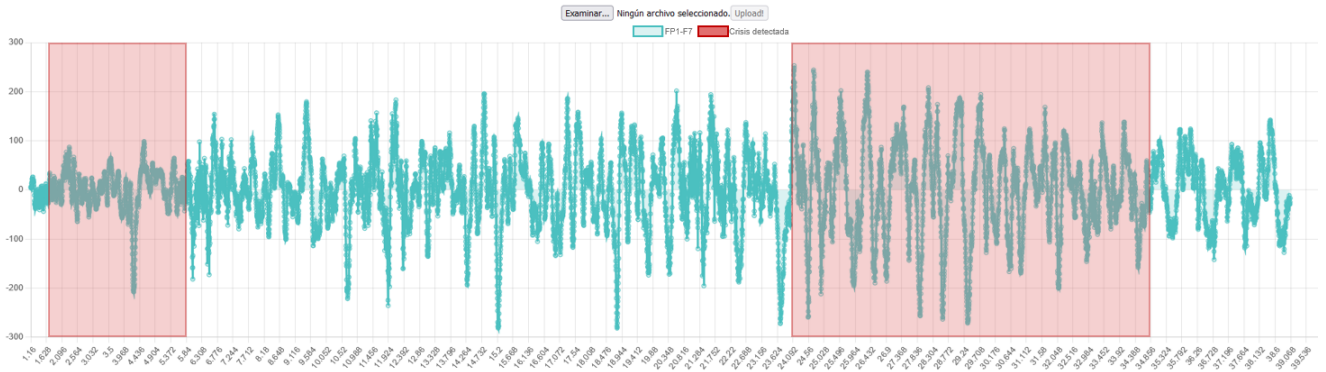


Figura 4.5: Ejemplo del funcionamiento de la plataforma.

Para ejecutar la tarea de clasificación, la aplicación procesa el archivo cargado de un modo análogo al desarrollado para el entrenamiento de los modelos: segmenta la señal en canales para luego generar ventanas temporales. Las ventanas temporales se evalúan a través de un modelo precargado en la plataforma, guardado de la parte anterior para persistencia del modelo.

Un mayor desarrollo del prototipo, o el desarrollo completo de la plataforma de acuerdo a las especificaciones descritas en las partes previas, se proponen como trabajo posterior al presente.

4.4. Impacto económico y social

El presente trabajo tiene como objetivo ser un insumo para mejorar la capacidad diagnóstica de crisis epilépticas en pacientes críticos, para así propiciar tratamientos tempranos que deriven en una mejora de los pronósticos asociados. A continuación se busca cuantificar el beneficio que sería posible generar de implementar el uso de un modelo de clasificación automatizada de crisis epilépticas, tanto a nivel individual como a nivel institucional del instituto INCA. Sin perjuicio de lo anterior, los resultados son generalizables a otras instituciones de salud que cuenten con el mismo problema.

Como se señala en partes anteriores, la prevalencia e incidencia del estado epiléptico no convulsivo es subestimada por su complejidad diagnóstica[30]. Por ello se explora en primera instancia desde el diagnóstico de epilepsia. La epilepsia es una enfermedad del sistema nervioso central que engloba un conjunto de síndromes, definidos clínicamente por episodios recurrentes de crisis epilépticas (convulsivas o no convulsivas)[40], y afecta a más de 50 millones de personas en el mundo[14], con una incidencia aproximada en Chile de 114 casos cada 100.000 habitantes al año[9].

En el estudio de esta condición, el último tiempo se ha puesto el foco en factores anteriormente poco valorados o estudiados. Así, se ha observado en las personas con epilepsia la existencia de alteraciones de la memoria y la atención, cambios de comportamiento y de reacciones frente a la cronicidad de la enfermedad, y trastornos del estado de ánimo, como depresión[15] o ansiedad. Lo anterior es una muestra de las formas en las que esta condición

puede afectar la calidad de vida de las personas. Dichas consecuencias también son sufridas por su entorno. Quienes han tomado el rol de cuidadores de personas con epilepsia informan ansiedad, depresión e insomnio[23].

En el lado económico, en el caso de Estados Unidos, se tiene que los costos médicos directos anuales promedio para los cuidadores de niños con frecuencias de convulsiones baja y alta son 3,2 millones y 8,1 millones de pesos chilenos, respectivamente. Los costos para los cuidadores de adultos con discapacidad intelectual se estiman entre 3,9 y 6,8 millones. Los costos indirectos asociados al cuidado de un niño son de 16,3 y 31,8 millones, para convulsiones poco frecuentes y frecuentes, respectivamente; mientras los de los cuidadores de un adulto fueron entre 11 y 22,5 millones. Por otra parte, los costos anuales de atención médica fueron de aproximadamente 2,2 millones y los costos de pérdida de productividad fueron de 4 millones de pesos por persona. Al extrapolar a la población estadounidense de cuidadores de personas con epilepsia, éstas se vieron a una carga económica 4 veces más grande que el resto de la población [23].

	Por persona en EEUU	Por persona en Chile	Total Chile	Estado epiléptico
Pérdida de productividad	\$ 4.000.000	\$ 832.885	\$ 18.684.383.159	\$ 9.276.632.340
Costos de atención médica	\$ 2.200.000	\$ 458.087	\$ 10.276.410.738	\$ 5.102.147.787

Tabla 4.1: Comparación de costos estimados por país.

Es posible estimar los costos respectivos en Chile mediante un par de supuestos. En la tabla 4.1 se presentan los costos por pérdida de productividad y de atención médica. Al ajustar según el PIB per cápita (ver Tabla 4.2), se obtiene el costo estimado por persona en Chile. Considerando la incidencia de 114 personas cada 100.000 habitantes, además de la población estimada por el INE de Chile de 19.678.363 personas, se obtiene el gasto estimado del total de personas afectadas por epilepsia en Chile al año. Al ajustar por la incidencia estimada del Estado Epiléptico de 56,6 casos por cada 100.000 adultos[39], se obtiene que solo en pérdida de productividad y costos de atención médica directa, esta condición implica un perjuicio económico de 14.400 millones de pesos al año. Estos costos son asumidos por las personas o por el estado, considerando las prestaciones que deben efectuarse tanto en salud pública como privada, además de los efectos de la pérdida de la productividad asociada.

País	PIB per cápita
EEUU	\$ 63.543
Chile	\$ 13.231

Tabla 4.2: Comparación de PIB per cápita al año 2020, según datos de las cuentas nacionales del Banco Mundial y archivos de datos de las cuentas nacionales de la OCDE.

Considerando que un tratamiento precoz del EE muestra hasta un 71,4% de efectividad[47], es posible estimar el efecto de la detección temprana de crisis epilépticas. Considerando que el mejor modelo entrenado (LSTM) es capaz de detectar el 87,2% de las crisis presentadas, si fuese posible replicar el resultado en clínica, detectando los casos de estado epiléptico a

tiempo para efectuar un tratamiento efectivo, y que dicho tratamiento deriva en una recuperación rápida y efectiva, sería posible salvar o disminuir las secuelas de alrededor de 1144 personas al año, solo en Chile.

Medicamento	% de éxito
Fenobarbital	71,4
Lorazepam	63
Diazepam y fenitoína	53,3
Fenitoína	41,5

Tabla 4.3: Comparación de efectividad en tratamiento de Estado Epiléptico.
Fuente: Treiman & Walker [47]

La estimación anterior considera la población estimada para Chile por el INE para el año 2021 de 19.678.363 personas, la incidencia estimada por Romano [39] de 56,6 casos por cada 100.000 adultos de EE, la proporción de casos de EE que son no convulsivos [39], la mortalidad estimada por Young [52], la sensibilidad reportada previamente por el modelo LSTM y la probabilidad de éxito del mejor tratamiento reportado en la Tabla 4.3.

Siguiendo el mismo razonamiento anterior, ahora aplicado a los costos de pérdida de productividad y de atención médica, es posible estimar una potencial reducción anual de estos costos de hasta 950 millones de pesos y 524 millones de pesos, respectivamente. Lo anterior totaliza 1.470 millones de pesos anualmente de ahorro agregado, como techo de rendimiento.

Es necesario indicar que, tal como se señala, las estimaciones consideran el uso del mejor medicamento, de la mejor capacidad de detección y de buenas respuestas a los tratamientos. Cada una de las condiciones anteriores podría no darse, y de hecho es esperable que no se den simultáneamente. Sin embargo, es relevante para los efectos del presente trabajo conocer cuales serían los efectos de su implementación, aún en condiciones ideales.

Suponiendo un caso más realista, si el sistema fuese capaz de detectar un 70 % de las crisis que se le presentan, con la posibilidad de otorgar tratamiento en el 30 % de los casos, y con una efectividad de 50 % en el tratamiento, entonces sería posible salvar o disminuir secuelas en hasta 190 pacientes al año, ahorrar costos de atención médica en cerca de 90 millones de pesos, y evitar pérdidas en productividad de hasta 170 millones de pesos al año, solo en Chile. A lo anterior se debe agregar el impacto personal y social, la oportunidad de mejorar el bienestar y tener una mejor calidad de vida, tanto para los pacientes como su entorno.

Para las estimaciones anteriores no se consideró un costo de desarrollo ni de implementación de la solución, por estar en etapas muy tempranas para definir un producto viable.

Capítulo 5

Conclusiones

A continuación se expone los principales resultados del presente trabajo de investigación. Se hace referencia a las restricciones durante el desarrollo del trabajo y se hace sugerencias para trabajos futuros relacionados.

5.1. Conclusiones generales

En los objetivos del trabajo se plantea desarrollar un modelo capaz de clasificar actividad epileptiforme en registros de pacientes del Instituto de Neurocirugía Asenjo. Asociado a este objetivo se plantea la hipótesis de investigación, de que es efectivamente posible identificar dicha actividad cerebral para los registros mencionados.

Al respecto, es necesario comentar la dificultad en tener acceso y disponibilidad de datos del Instituto para su análisis. Lo anterior, fue previsto como uno de los eventuales riesgos de trabajar con data médica, por su carácter confidencial. Por lo anterior, se debió recurrir a bases de datos con registros EEG de repositorio. Sin duda, dicha decisión afecta el cumplimiento de los objetivos, planteados en el escenario de la utilización de registros de pacientes neurocríticos, y en particular del INCA. Pese a ello, se consideró que aún así era posible cumplir con los objetivos propuestos, con diferente fuente de datos.

Así, respecto a los objetivos específicos:

- Respecto al estado del arte, fue posible identificar un área de estudio bastante avanzada, con una variedad de métodos cada vez más amplia, y con la introducción de técnicas que permiten tener resultados cercanos al gold standard. La exploración de los trabajos actuales fue la base para la elaboración de un modelo de clasificación capaz de clasificar crisis epilépticas.
- Como ya se señala previamente, se debió recurrir a diferentes fuentes de datos, distintas de la originalmente planeada para poder entrenar los modelos propuestos. La cantidad de data disponible públicamente es bastante amplia, lo que no significa que sea suficiente para realizar todo tipo de tareas. No fue posible sistematizar data correspondiente a pacientes neurocríticos en la condición que se planteó al inicio de este trabajo, a partir de data pública. Lo anterior sugiere que la recolección, sistematización y puesta a disposición de data clínica por parte de las instituciones médicas es necesaria, no solo para

tener procesos rutinarios más efectivos, sino que, en la eventualidad de surgir oportunidades de mejora, para ser capaces de tener los insumos necesarios para crear soluciones a la medida de sus propias necesidades.

- El modelo que obtuvo los mejores resultados con los registros disponibles fue el de redes neuronales recurrentes LSTM, presentando un recall de hasta un 87 %, accuracy de 70 % y AUC de 0,875. Dicho resultado se presenta al nivel superior de las alternativas comerciales estudiadas. A falta de una comparación con un especialista in situ, se considera dichas alternativas como referencia de efectividad, por lo que se evalúa positivamente el resultado obtenido.
- Al tener un modelo con un desempeño adecuado se esquematizó un prototipo capaz de hacer visibles las potencialidades de uso de un sistema de clasificación automática. La maqueta o mockup sobrepasó la capacidad de desarrollo del autor en los plazos establecidos, sin embargo se presentó un prototipo capaz de realizar las tareas básicas planteadas inicialmente. Se considera un objetivo logrado, por cuanto establece las bases del desarrollo a utilizar de cara a una futura utilización, y la dirección en la cual realizar dicho desarrollo.
- Respecto al impacto económico estimado de una aplicación clínica del modelo propuesto, y como se señaló en la parte anterior, las estimaciones monetarias son optimistas al considerar que todos los pacientes pueden ser tratados, omitiendo la existencia de pacientes que podrían no estar bajo supervisión médica al momento de presentar crisis epilépticas, la influencia de otras patologías en el pronóstico final, o la disponibilidad de medicinas o especialistas para efectuar el tratamiento planteado. Sin perjuicio de lo anterior, se establece que el desarrollo de una aplicación con las características descritas sería beneficiosa tanto social como económicamente para los pacientes, para los equipos médicos, las instituciones de salud y gobierno.

Por lo anterior, se considera plenamente cumplidos 4 de los 5 objetivos del presente trabajo de título, con excepción del referente a “Construir un set de datos con muestras de pacientes reales para el entrenamiento de modelos.”. Al haber tenido que usarse fuentes alternativas de datos, como lo son repositorios de registros EEG públicos, no se está cumpliendo el espíritu de dicho objetivo, por cuanto se pretendía que dichos registros fueran los del INCA. Sin embargo, al haberse establecido un set de datos a partir de registros de repositorios, con el cual fue efectivamente posible entrenar y testear modelos, se considera dicho objetivo parcialmente realizado.

Respecto a la hipótesis de investigación planteada inicialmente, los resultados del trabajo indican que es posible detectar las crisis epilépticas de forma automática, lo que implica eventuales mejoras en los tiempos de aplicación de tratamiento a los pacientes, menores tiempos de revisión de registros por parte de los especialistas y un consecuente progreso en la capacidad de la gestión diagnóstica de crisis epilépticas. Es por ello, que se considera la hipótesis validada. Sin perjuicio de lo anterior, se considera necesario como paso siguiente el validar la hipótesis en una experiencia clínica que involucre un trabajo directo con los registros del INCA y datos como los tiempos de respuestas efectivos ante crisis, tasas de recuperación de pacientes, tiempos de revisión de los registros y tasas de ocupación de los especialistas médicos involucrados.

5.2. Trabajos futuros

A partir de la experiencia del trabajo realizado se sugieren para trabajos futuros los siguientes puntos:

- Establecer de antemano la fuente de datos, asegurando que los instrumentos de medición se correspondan entre muestras y que exista cierta estandarización en el tipo de dato. De este modo será posible evitar lidiar con, por ejemplo, distintos montajes de los electrodos que puedan afectar el resultado de un experimento. Es importante a la vez considerar las diferencias fisiológicas entre pacientes de, por ejemplo, diferentes edades.
- Un espacio de mejora del INCA es tener a disposición datos de pacientes reales, con las respectivas etiquetas que hagan posible su procesamiento en aprendizaje de máquinas supervisado.
- Por los puntos anteriores se sugiere replicar los análisis ya realizados en el presente trabajo, una vez que los datos estén disponibles, de modo de tener resultados que sean relevantes para la institución.
- A partir de la cantidad de datos disponibles, se sugiere el uso de herramientas de big data para la estructuración de un set de datos con las características que se quieren estudiar. En el caso del presente trabajo dicha característica era la de ser registros de pacientes neurocríticos.
- Respecto a las cargas de trabajo, se sugiere dividir el trabajo de exploración de modelos que puedan ser efectivos en la clasificación de crisis epilépticas y el trabajo de desarrollo de la plataforma capaz de presentar dichos resultados a un usuario médico.
- En cuanto al prototipo propuesto, en trabajos futuros lo esperable es que la carga de archivos a la plataforma, el proceso de análisis y clasificación y reportería se haga de forma automática, sin la necesidad de introducir de forma manual el archivo a estudiar. Dicho trabajo implica manejo de bases de datos y arquitectura de redes, además de los módulos de interfaz de usuario, por lo que se sugiere abordarlo separadamente desde ambas perspectivas.

Bibliografía

- [1] Saad Albawi, Tareq Abed Mohammed, and Saad Al-Zawi. Understanding of a convolutional neural network. In *2017 International Conference on Engineering and Technology (ICET)*, pages 1–6. Ieee, 2017.
- [2] Alvear, S., Canteros, J., Jara, J., & Rodríguez, P. Costos reales de tratamientos intensivos por paciente y día cama. *Revista médica de Chile*, 141(2):202–208, 2013.
- [3] Web Intelligence Centre. Quienes somos, (s. f.); archivado en *Wayback Machine*. URL <https://web.archive.org/web/20151009225408/http://wic.uchile.cl/quienes-somos/>.
- [4] Web Intelligence Centre. Home, (s. f.); revisado el 25 de noviembre de 2020, en. URL <https://wic.uchile.cl/>.
- [5] Cynthia C Chernecky and Barbara J Berger. *Laboratory Tests and Diagnostic Procedures-E-Book*. Elsevier Health Sciences, 2012.
- [6] Constanza, A., & Alberto, A. Estado Epiléptico no Convulsivo. *Revista Memoriza.com*, 10:7–12, 2013.
- [7] De Lorenzo R.J., Hauser W.A., Towne A.R., et al. A prospective, populationbased epidemiologic study of status epilepticus in Richmond, Virginia. *Neurology*, 46:1029–1035, 1996.
- [8] Instituto de Neurocirugía Dr. Alfonso Asenjo. Quienes somos, (s. f.); consultado el 30 de noviembre de 2020, en. URL <https://www.institutodeneurocirugia.cl/quienes-somos/>.
- [9] Ministerio de Salud del Gobierno de Chile. Guía clínica auge, epilepsia adultos, 2014. URL https://www.minsal.cl/sites/default/files/files/GUIA%20CLINICA_EPILEPSIA%20ADULTOS_web.pdf.
- [10] Wouter Deburchgraeve, PJ Cherian, Maarten De Vos, RM Swarte, JH Blok, Gerhard Henk Visser, Paul Govaert, and Sabine Van Huffel. Automated neonatal seizure detection mimicking a human observer reading eeg. *Clinical Neurophysiology*, 119(11): 2447–2454, 2008.
- [11] AD Dongare, RR Kharde, Amit D Kachare, et al. Introduction to artificial neural network. *International Journal of Engineering and Innovative Technology (IJEIT)*, 2 (1):189–194, 2012.
- [12] Ebersole, J. S., & Pedley, T. A. (Eds.). *Current practice of clinical electroencephalography*. Lippincott Williams & Wilkins, 2003.
- [13] Solari B. F. Estado epiléptico, consideraciones sobre manejo y tratamiento. *Revista Médica Clínica Las Condes*, 24(6):938–945, 2013. doi: [https://doi.org/10.1016/S0716-8640\(13\)70247-5](https://doi.org/10.1016/S0716-8640(13)70247-5).

- [14] Joaquim Ferreira and Tiago Mestre. Eslicarbazepine acetate: a new option for the treatment of focal epilepsy. *Expert opinion on investigational drugs*, 18(2):221–229, 2009.
- [15] Frank G Gilliam. Diagnosis and treatment of mood disorders in persons with epilepsy. *Current opinion in neurology*, 18(2):129–133, 2005.
- [16] Meysam Golmohammadi, Amir Hossein Harati Nejad Torbati, Silvia Lopez de Diego, Iyad Obeid, and Joseph Picone. Automatic analysis of eegs using big data and hybrid deep learning architectures. *Frontiers in Human Neuroscience*, 13:76, 2019. ISSN 1662-5161. doi: 10.3389/fnhum.2019.00076. URL <https://www.frontiersin.org/article/10.3389/fnhum.2019.00076>.
- [17] Catalina Gómez, Pablo Arbeláez, Miguel Navarrete, Catalina Alvarado-Rojas, Michel Le Van Quyen, and Mario Valderrama. Automatic seizure detection based on imaged-eeg signals through fully convolutional networks. *Scientific reports*, 10(1):1–13, 2020.
- [18] Jean Gotman. Automatic detection of seizures and spikes. *Journal of Clinical Neurophysiology*, 16(2):130–140, 1999.
- [19] Alioth Guerrero-Aranda. Electroencefalografía continua en pacientes críticamente enfermos ii: desafíos y perspectivas. *Acta Colombiana de Cuidado Intensivo*, 20(1):39 – 44, 2020. ISSN 0122-7262. doi: <https://doi.org/10.1016/j.acci.2019.11.008>. URL <http://www.sciencedirect.com/science/article/pii/S0122726219300722>.
- [20] Gastaut H. Classification of status epilepticus. En: Delgado-Escueta AV, Wasterlain CG, Treiman DT, Porter RJ, eds. *Status Epilepticus. Advances in Neurology*, 34:15–35, 1983.
- [21] MM Hartmann, K Schindler, TA Gebbink, G Gritsch, and T Kluge. Pureeeg: Automatic eeg artifact removal for epilepsy monitoring. *Neurophysiologie Clinique/Clinical Neurophysiology*, 44(5):479–490, 2014.
- [22] R Hopfengärtner, F Kerling, V Bauer, and H Stefan. An efficient, robust and fast method for the offline detection of epileptic seizures in long-term scalp eeg recordings. *Clinical Neurophysiology*, 118(11):2332–2343, 2007.
- [23] Shaun A Hussain, Jesse D Ortendahl, Tanya GK Bentley, Amanda L Harmon, Shaloo Gupta, Charles E Begley, Ibrahim Khilfeh, and Russell L Knoth. The economic burden of caregiving in epilepsy: An estimate based on a survey of us caregivers. *Epilepsia*, 61(2):319–329, 2020.
- [24] Instituto Nacional de Estadísticas. Inicio, (s. f.); consultado el 10 de noviembre de 2020, en. URL <https://www.ine.cl/>.
- [25] Serkan Kiranyaz, Onur Avcı, Osama Abdeljaber, Turker Ince, Moncef Gabbouj, and Daniel J Inman. 1d convolutional neural networks and applications: A survey. *Mechanical systems and signal processing*, 151:107398, 2021.
- [26] Johannes Koren, Sebastian Hafner, Moritz Feigl, and Christoph Baumgartner. Systematic analysis and comparison of commercial seizure-detection software. *Epilepsia*, 62(2): 426–438, 2021.
- [27] Laroche S., Haider H. *Handbook of ICU EEG monitoring*. Springer Publishing Company, 2018.
- [28] Lawson P. B., Quintana F. J. C. Status epiléptico no convulsivo: una entidad heterogénea

- de difícil diagnóstico. Reporte de dos casos con presentación inusual. *Rev. méd. Chile*, 131:1045–1050, 2003.
- [29] Oded Maimon and Lior Rokach. *Data mining and knowledge discovery handbook*. Springer, 2005.
- [30] Andrea Yosajany Morales-del Ángel, Daniel Ramón Hernández-Salcedo, Raúl Valencia-López, and Joel Orozco-Paredes. El estado epiléptico no convulsivo en el paciente geriátrico. *Medicina interna de México*, 34(5):704–714, 2018.
- [31] Neurospec. Besa epilepsy 2.0, (s. f.); revisado el 25 de agosto de 2021, en. URL <https://shop.neurospec.com/besa-epilepsy/>.
- [32] Ernst Niedermeyer and FH Lopes da Silva. *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.
- [33] Iyad Obeid and Joseph Picone. The temple university hospital eeg data corpus. *Frontiers in neuroscience*, 10:196, 2016.
- [34] Epilepsy Foundation of America’s Working Group on Status Epilepticus. Treatment of convulsive status epilepticus: recommendations of the Epilepsy Foundation of America’s Working Group on Status Epilepticus. *JAMA*, 270:854–859, 1993.
- [35] Ivan Osorio, Mark G Frei, and Steven B Wilkinson. Real-time automated detection and quantitative analysis of seizures and short-term prediction of clinical onset. *Epilepsia*, 39(6):615–627, 1998.
- [36] Georgia Ramantani, Louis Maillard, and Laurent Koessler. Correlation of invasive eeg and scalp eeg. *Seizure*, 41:196–200, 2016. ISSN 1059-1311. doi: <https://doi.org/10.1016/j.seizure.2016.05.018>. URL <https://www.sciencedirect.com/science/article/pii/S1059131116300656>.
- [37] Godoy F. Jaime & Uribe S. Reinaldo. Estado epiléptico del adulto. *Revista Médica Clínica Las Condes*, 24(6):947 – 952, 2013. ISSN 0716-8640. doi: [https://doi.org/10.1016/S0716-8640\(13\)70248-7](https://doi.org/10.1016/S0716-8640(13)70248-7).
- [38] Gonzalo M Rojas, Carolina Alvarez, Carlos E Montoya, María de la Iglesia-Vayá, Jaime E Cisternas, and Marcelo Gálvez. Study of resting-state functional connectivity networks using eeg electrodes position as seed. *Frontiers in neuroscience*, 12:235, 2018.
- [39] Romano, L. M. Incidencia del status epiléptico en el adulto: estudio epidemiológico prospectivo basado en población cautiva en Argentina. *Neurología Argentina*, 2(3):172–179, 2010.
- [40] Antoni Rossiñol, Irene Molina, Tomeu Rossiñol, and Alexandre Garcia-Mas. Calidad de vida y percepción de salud general de personas con epilepsia en función de las crisis, la afectación neurocognitiva, la electroencefalografía, su respuesta al tratamiento y los efectos secundarios. *Rev Neurol*, 57(1):17–24, 2013.
- [41] Ali Hossam Shoeb. *Application of machine learning to epileptic seizure onset detection and treatment*. PhD thesis, Massachusetts Institute of Technology, 2009.
- [42] Yuedong Song, Jon Crowcroft, and Jiaxiang Zhang. Automatic epileptic seizure detection in eegs based on optimized sample entropy and extreme learning machine. *Journal of Neuroscience Methods*, 210(2):132–146, 2012. ISSN 0165-0270. doi: <https://doi.org/10.1016/j.jneumeth.2012.05.018>.

1016/j.jneumeth.2012.07.003. URL <https://www.sciencedirect.com/science/article/pii/S0165027012002671>.

- [43] Cornelis J Stam. Nonlinear dynamical analysis of eeg and meg: review of an emerging field. *Clinical neurophysiology*, 116(10):2266–2301, 2005.
- [44] Aaron F Struck, Berk Ustun, Andres Rodriguez Ruiz, Jong Woo Lee, Suzette M LaRoche, Lawrence J Hirsch, Emily J Gilmore, Jan Vlachy, Hiba Arif Haider, Cynthia Rudin, et al. Association of an electroencephalography-based risk score with seizure probability in hospitalized patients. *JAMA neurology*, 74(12):1419–1424, 2017.
- [45] William O Tatum IV. *Handbook of EEG interpretation*. Demos Medical Publishing, 2014.
- [46] A. Temko, E. Thomas, W. Marnane, G. Lightbody, and G. Boylan. Eeg-based neonatal seizure detection with support vector machines. *Clinical Neurophysiology*, 122(3):464–473, 2011. ISSN 1388-2457. doi: <https://doi.org/10.1016/j.clinph.2010.06.034>. URL <https://www.sciencedirect.com/science/article/pii/S1388245710006036>.
- [47] David M Treiman and MC Walker. Treatment of seizure emergencies: convulsive and non-convulsive status epilepticus. *Epilepsy research*, 68:77–82, 2006.
- [48] Trinkka E., Kälviäinen R. 25 years of advances in the definition, classification and treatment of status epilepticus. *Seizure*, 44:65–73, 2017.
- [49] Ihsan Ullah, Muhammad Hussain, Emad ul Haq Qazi, and Hatim Aboalsamh. An automated system for epilepsy detection using eeg brain signals based on deep learning approach. *Expert Systems with Applications*, 107:61–71, 2018. ISSN 0957-4174. doi: <https://doi.org/10.1016/j.eswa.2018.04.021>. URL <https://www.sciencedirect.com/science/article/pii/S0957417418302513>.
- [50] Berk Ustun and Cynthia Rudin. Methods and models for interpretable linear classification. *arXiv preprint arXiv:1405.4047*, 2014.
- [51] Sholom M Weiss and Nitin Indurkha. *Predictive data mining: a practical guide*. Morgan Kaufmann, 1998.
- [52] Young G.B., Jordan K.G., Doig G.S. An assessment of nonconvulsive seizures in the intensive care unit using continuous EEG monitoring: an investigation of variables associated with mortality. *Neurology*, Jul;47(1):83–9, 1996.
- [53] Tao Zhang, Wanzhong Chen, and Mingyang Li. Ar based quadratic feature extraction in the vmd domain for the automated seizure detection of eeg using random forest classifier. *Biomedical Signal Processing and Control*, 31:550–559, 2017. ISSN 1746-8094. doi: <https://doi.org/10.1016/j.bspc.2016.10.001>. URL <https://www.sciencedirect.com/science/article/pii/S1746809416301525>.

Anexo A

Capítulo 1

A.1. Criterios para crisis no convulsivas

Criterios para las crisis no convulsivas

Criterios Primarios

1. Descargas epileptiformes focales o generalizadas repetitivas a un ritmo de $\geq 3\text{Hz}$
2. Descargas epileptiformes focales o generalizadas repetitivas a un ritmo de $<3\text{Hz}$ y el criterio secundario
3. Ondas secuenciales rítmicas, periódicas o cuasiperiódicas a 1 Hz y evolución inequívoca de la frecuencia (aumentando o disminuyendo gradualmente por al menos 1 Hz), morfología o extensión espacial. Excluye evolución solo en amplitud, o solo cambios en la forma .

Criterio Secundario

1. Mejora significativa en el estado clínico o aparición de patrones normales de EEG previamente ausentes en respuesta a la administración aguda de un fármaco antiepiléptico (FAE), como benzodiazepina.

Definición de evolución

1. Frecuencia: al menos dos cambios consecutivos en la misma dirección en al menos 0,5/s, por ejemplo, de 2 a 2,5 a 3/s, o de 3 a 2 a 1,5/s
2. Morfología: al menos dos cambios consecutivos a una morfología nueva
3. Ubicación: propagación secuencial hacia dentro o hacia fuera de al menos dos posiciones de electrodos estándar de 10-20 diferentes.

Para calificar como presente, una sola frecuencia o ubicación debe persistir durante al menos tres ciclos. Los criterios de evolución deben alcanzarse sin que el patrón permanezca inalterado en frecuencia, morfología o ubicación durante ≥ 5 min.

Fuente: Ebersole & Pedley [12]