

Detección de episodios epilépticos en pacientes de 3 a 12 años mediante el análisis de EEGs con algoritmos de clasificación de Machine Learning

Pablo F. Muñoz Burbano, Juan D. Cabrera Moncada.

Resumen—La epilepsia es una condición neurológica que puede afectar gravemente la calidad de vida de los pacientes debido a la imprevisibilidad de las crisis. Este trabajo propone un sistema clasificador de episodios epilépticos en pacientes pediátricos (3 a 12 años), utilizando señales EEG y técnicas de Machine Learning (ML). Se realizó una cuidadosa selección de características basadas en amplitud, frecuencia y morfología de la señal, complementada con técnicas estadísticas, selección secuencial de características (SFBS) y reducción de dimensionalidad mediante PCA. Se evaluaron distintos modelos (SVM, Random Forest, Regresión Logística, RNA) tanto con como sin el uso de datos sintéticos (SMOTE). Los resultados muestran que SVM sin PCA ni SMOTE ofrece la mayor exactitud, mientras que Random Forest sin estas técnicas logra el mejor recall para eventos ictales. Estos hallazgos sugieren un potencial uso clínico para mejorar la predicción y respuesta ante episodios epilépticos en niños.

Palabras clave—Epilepsia, EEG, SVM, Random Forest, detección de crisis, pacientes pediátricos.

I. INTRODUCCIÓN

La epilepsia es un trastorno neurológico crónico caracterizado por descargas eléctricas anormales en el cerebro, las cuales pueden causar crisis epilépticas súbitas. Esta condición afecta la autonomía de los pacientes, especialmente en edades tempranas, y representa una causa relevante de atención médica. Actualmente, los avances en el análisis de señales EEG han permitido detectar ciertos patrones previos a los episodios, pero aún persiste el reto de lograr una predicción fiable. El uso de técnicas de Inteligencia Artificial, especialmente algoritmos de Machine Learning, ha abierto nuevas posibilidades para desarrollar sistemas automáticos de detección. No obstante, existen desafíos importantes como la escasez de datos pediátricos y la heterogeneidad entre sujetos. Este trabajo aborda dichos desafíos mediante el desarrollo de un sistema clasificador de episodios ictales e interictales, con enfoque en la población infantil de 3 a 12 años.

II. DESCRIPCIÓN DEL PROBLEMA

La epilepsia es una condición neurológica crónica que se caracteriza por la presencia de crisis recurrentes debidas a descargas eléctricas anormales en el cerebro. Estas crisis pueden ser impredecibles y abruptas, lo cual representa un riesgo significativo para la integridad física de quienes la padecen, especialmente en el caso de los niños, quienes aún se encuentran en una etapa de desarrollo neurológico y psicomotor. Las crisis epilépticas constituyen

entre el 1 % y el 2 % de las urgencias médicas, y se posicionan como la segunda causa más común de atención neurológica en emergencias, así como de consulta ambulatoria neurológica, y la tercera causa de hospitalización neurológica, no solo por la crisis en sí, sino también por los accidentes secundarios que pueden desencadenar.

El principal reto clínico y tecnológico en el manejo de la epilepsia reside en su imprevisibilidad. Esta limita la autonomía de los pacientes e impone una carga significativa tanto para los cuidadores como para los sistemas de salud. Por esta razón, la detección temprana de patrones en las señales EEG que preceden a un episodio ictal se ha convertido en un objetivo prioritario para la comunidad médica y científica, con el fin de desarrollar sistemas de alerta temprana que permitan una intervención oportuna.

En este contexto, el análisis computacional de señales EEG ha sido ampliamente explorado como herramienta para la identificación de crisis. Sin embargo, persisten desafíos técnicos relevantes. Uno de los principales es la heterogeneidad de los datos clínicos, especialmente cuando se desea construir modelos generalizables que sean útiles para diferentes grupos etarios. Como destacan AbuAlrob, Itbaisha y Mesraoua (2025), los modelos entrenados principalmente con datos de adultos pierden precisión al ser aplicados a niños. Esto evidencia la necesidad de contar con bases de datos específicas y representativas de la población infantil, así como de mecanismos de mejora como el data augmentation, que busca compensar la escasez de datos de ciertos grupos mediante la generación artificial de nuevas muestras.

En adición a esto, los enfoques basados en aprendizaje no supervisado, si bien permiten modelar sin necesidad de etiquetas previas, han demostrado limitaciones importantes en este campo. Según Tautan et al. (2025), este tipo de modelos tienden a realizar una extracción de características deficiente, incluso cuando se dispone de datasets amplios de EEG. Esto se debe a que la identificación de eventos críticos como las crisis epilépticas requiere de una comprensión profunda de la morfología y dinámica de las señales, lo cual es difícil de capturar sin guía supervisada.

Frente a esta limitación, los enfoques supervisados y semi-supervisados han demostrado mejores resultados.

Estudios como el de Karasmanoglou et al. (2025) han aplicado técnicas de detección de anomalías y filtrado en el dominio de la frecuencia para la identificación de epilepsia focal en pacientes pediátricos. En dicho estudio se integraron datos EEG y ECG, concluyendo que si bien el ECG no incrementa sustancialmente el rendimiento del modelo, sí aporta a una mejor precisión en la generación de alertas tempranas. No obstante, se demostró que el desempeño del modelo basado exclusivamente en EEG sigue siendo aceptable, lo que justifica su uso en sistemas de detección sin necesidad de señal cardíaca.

En este proyecto se busca abordar estas problemáticas mediante la implementación de un sistema de clasificación automático para distinguir entre periodos interictales e ictales en pacientes pediátricos (3 a 12 años), usando señales EEG provenientes de 15 sujetos de una base de datos con registros multicanal. El enfoque propuesto combina técnicas de preprocesamiento, selección de características y entrenamiento de modelos de Machine Learning supervisado, con el fin de evaluar su efectividad para la identificación de eventos clínicamente relevantes, priorizando el análisis en población infantil subrepresentada en estudios anteriores.

III. METODOLOGÍA

La metodología propuesta para este estudio se estructuró en cuatro etapas principales: (1) selección de la población y preparación de la base de datos; (2) preprocesamiento de señales EEG; (3) extracción y selección de características relevantes; y (4) entrenamiento y evaluación de modelos de clasificación. A continuación, se describen en detalle cada una de estas fases.

a. Selección de población y preparación de la base de datos.

La base de datos utilizada contiene registros de señales EEG de 24 sujetos, de los cuales se seleccionaron únicamente aquellos que se encontraban en el rango de edad de 3 a 12 años, dando un total de 15 pacientes pediátricos. Esta decisión responde al enfoque del estudio en la población infantil, dado que los patrones cerebrales y la actividad eléctrica del cerebro varían sustancialmente según la etapa del desarrollo neurológico. Para garantizar homogeneidad en los datos analizados, se consideraron únicamente los registros que contenían 23 canales EEG específicos, definidos con base en derivaciones bipolares que buscan eliminar artefactos y destacar la actividad eléctrica local, como la epileptiforme.

Además, se seleccionaron únicamente los segmentos EEG que contenían episodios ictales (crisis epilépticas activas) e interictales (periodos sin crisis). Esta decisión metodológica se tomó para asegurar una representación balanceada de ambas clases durante el entrenamiento y la validación de los modelos. En una etapa posterior, se contempla la posibilidad de incluir segmentos

puramente interictales si es necesario aumentar la cantidad de datos para entrenar el modelo de manera más robusta.

b. Preprocesamiento de señales EEG.

El preprocesamiento de las señales EEG se llevó a cabo con el objetivo de estandarizar la base de datos y asegurar la calidad de las muestras utilizadas en el modelado. Se identificaron registros con tasas de muestreo distintas a 256 Hz; aquellos con frecuencia superior fueron remuestreados a 256 Hz para garantizar uniformidad temporal en las ventanas de análisis. Los registros con frecuencia inferior a 256 Hz fueron descartados debido a la posible pérdida de información relevante para la detección de patrones ictales.

Asimismo, se eliminaron canales duplicados o inconsistentes dentro de los segmentos EEG, con el fin de facilitar la lectura de los archivos en formato .edf y evitar sesgos en la representación espacial de la señal cerebral. Posteriormente, cada segmento fue dividido en ventanas de tiempo de 1 segundo con 50% de solapamiento, asegurando que cada ventana correspondiera claramente a una etiqueta ictal o interictal y evitando pérdidas de información o generación de artefactos debido a discontinuidades entre las ventanas de tiempo.

c. Extracción y selección de características relevantes.

En esta etapa, se aplicó una técnica de descomposición de señal mediante Transformada Discreta de Wavelet Multinivel (TDWM) de tipo Daubechies, con 3 niveles de descomposición. Esta técnica permitió redistribuir el contenido de las ventanas EEG en cuatro subbandas: D1, D2, D3 (componentes de baja frecuencia), y A3 (componente de alta frecuencia). Esta transformación no representa directamente las características usadas en la clasificación, pero sí facilita la localización espectro-temporal de la energía contenida en diferentes rangos de frecuencia, lo cual es crucial para detectar eventos ictales que a menudo se manifiestan como patrones rítmicos anómalos.

Posteriormente, se extrajeron características representativas de las señales contenidas en cada uno de los 21 canales de la ventana de tiempo, para cada una de las 4 subbandas generadas, específicamente, entropía de Shanon y energía logarítmica.

Por un lado, la entropía de Shannon es una medida que describe el grado de desorden o aleatoriedad en la distribución de energía de los coeficientes. A mayor entropía, mayor complejidad o irregularidad en la señal. Los pasos para hallar esta métrica son los siguientes:

- Primero, se calcula la potencia de cada subbanda:

$$power_i = c_i^2$$

- Luego, se normaliza la potencia dividiendo por la suma total:

$$norm_{power_i} = power_i / \sum power_i$$

- Por último, la entropía se calcula como:

$$Shannon\ Entropy = - \sum_i (norm_{power_i} * \log_2(norm_{power_i}))$$

Ahora, la energía logarítmica es una métrica que se encarga de describir la concentración o intensidad de energía de los coeficientes. Esto ayuda a detectar transiciones o eventos con energía elevada, como posibles crisis epilépticas. Lo cual fue calculado siguiendo esta fórmula:

$$Log\ Energy = \sum_i \log(c_i^2 + 10^{-10})$$

Para reducir el conjunto de características y conservar únicamente aquellas con mayor poder discriminativo, se aplicaron pruebas estadísticas:

- Primero, se utilizó el test de Kolmogorov–Smirnov para verificar la normalidad de las distribuciones.
- Luego, se aplicó el test de Mann–Whitney U para evaluar la significancia de las diferencias entre clases en cada característica extraída.

Con base en los resultados de estos tests, se utilizó un proceso de búsqueda secuencial hacia atrás (SFBS, Sequential Floating Backward Selection) para identificar el subconjunto de características que maximiza el desempeño de los modelos, medido mediante la métrica F1-Score macro. Esta métrica fue elegida debido a su capacidad para balancear precisión y exhaustividad en escenarios con clases desbalanceadas.

Finalmente, se aplicó una reducción de dimensionalidad mediante Análisis de Componentes Principales (PCA), conservando aquellas componentes principales que explicaran más del 90 % de la varianza total de los datos. Esta reducción permitió simplificar la representación de cada ventana EEG y reducir la carga computacional en la fase de clasificación, aunque se mantuvo como una etapa opcional para contrastar el desempeño de los modelos con y sin esta técnica.

A continuación, se muestra un diagrama que resume la implementación de esta fase de procesamiento de características, partiendo desde el punto en el que se tiene una ventana de 1 segundo $x(n)$ hasta la fase de uso de PCA como herramienta de extracción de nuevas características.

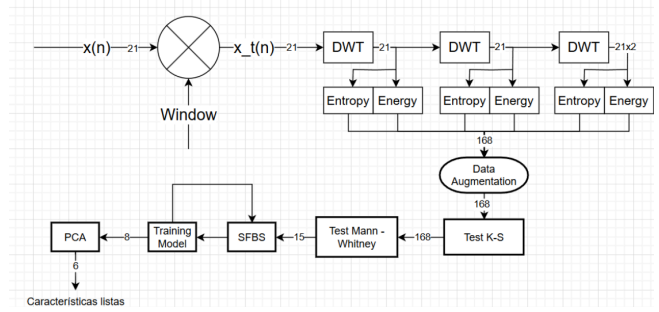


Figura 1. Diagrama de procesamiento de características.

d. Entrenamiento y evaluación de modelos de clasificación.

Para la fase de clasificación, se implementaron cuatro modelos supervisados de uso común en tareas de clasificación biomédica:

- Regresión Logística
- Random Forest
- Máquinas de Vectores de Soporte (SVM)
- Redes Neuronales Artificiales (RNA) con entrenamiento tipo backpropagation

La base de datos fue dividida en 70 % para entrenamiento y 30 % para prueba, mediante una partición estratificada, lo que garantizó que las proporciones de las clases ictal e interictal se mantuvieran constantes en ambos subconjuntos. Además, se aplicó una estrategia de exclusión por paciente, es decir, los datos de un mismo paciente no fueron usados en ambos conjuntos, con el fin de simular un escenario clínico realista en el que se desea predecir episodios en pacientes no vistos previamente por el modelo.

Cada modelo fue entrenado y ajustado en dos etapas:

- Primero, se optimizó la exactitud (accuracy) general.
- Luego, se priorizó el recall de la clase ictal, dado que detectar con alta sensibilidad los episodios críticos es el objetivo principal del sistema.

Además, se realizaron experimentos bajo distintas combinaciones de preprocesamiento:

- Con y sin generación de datos sintéticos mediante SMOTE (Synthetic Minority Over-sampling Technique).
- Con y sin reducción de dimensionalidad mediante PCA.

Esto permitió comparar objetivamente el efecto de cada técnica sobre el desempeño final del clasificador. A continuación un diagrama que resume la implementación de esta metodología.

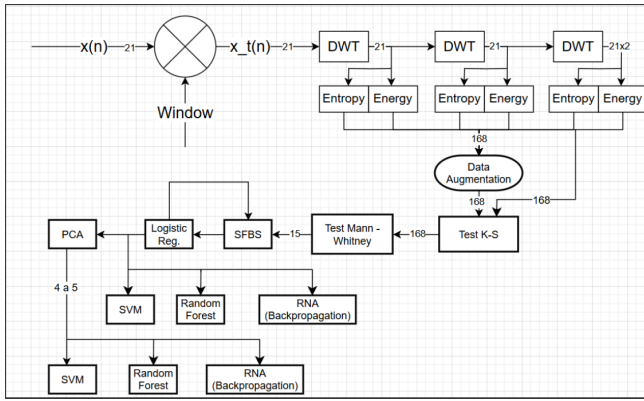


Figura 2. Diagrama final de metodología de análisis de segmentos ictales e interictales.

IV. RESULTADOS

Con el fin de evaluar el rendimiento de los modelos de clasificación implementados, se realizó una comparación sistemática bajo diferentes configuraciones de preprocesamiento, incluyendo el uso o no de PCA para la reducción de dimensionalidad y la generación de datos sintéticos mediante SMOTE. La métrica principal de evaluación fue la exactitud general (accuracy) y el recall de la clase ictal, por su relevancia clínica en la detección de eventos epilépticos. Aquí una tabla de todos los resultados obtenidos para los modelos de ML puestos en prueba.

Modelo	SMOTE	PCA	Conjunto	Accuracy	Recall (Ictal)
Regresión Logística	No	—	Entrenamiento	86%	60%
	No	—	Prueba	79%	50%
	Sí	—	Entrenamiento	83%	76%
	Sí	—	Prueba	81%	66%
Random Forest	No	No	Entrenamiento	88%	81%
	No	No	Prueba	82%	69%
	No	Sí	Entrenamiento	87%	80%
	No	Sí	Prueba	79%	67%
	Sí	No	Entrenamiento	85%	81%
	Sí	No	Prueba	81%	69%
	Sí	Sí	Entrenamiento	84%	79%
	Sí	Sí	Prueba	81%	67%
SVM	No	No	Entrenamiento	84%	68%
	No	No	Prueba	83%	66%
	No	Sí	Entrenamiento	87%	78%
	No	Sí	Prueba	79%	67%
	Sí	No	Entrenamiento	80%	72%
	Sí	No	Prueba	80%	61%
	Sí	Sí	Entrenamiento	83%	77%
	Sí	Sí	Prueba	81%	66%
RNA (Backpropagation)	No	No	Entrenamiento	82%	67%
	No	No	Prueba	81%	54%
	No	Sí	Entrenamiento	88%	68%
	No	Sí	Prueba	82%	56%
	Sí	No	Entrenamiento	83%	70%
	Sí	No	Prueba	82%	65%
	Sí	Sí	Entrenamiento	84%	74%
	Sí	Sí	Prueba	82%	63%

Tabla 1. Tabla general de resultados de pruebas con los modelos de ML.

Los resultados muestran que el modelo SVM sin aplicar SMOTE ni PCA obtuvo la mayor exactitud (83 %),

lo que lo posiciona como el más confiable para clasificar correctamente los segmentos EEG entre periodos ictales e interictales. Esto sugiere que el modelo es capaz de generalizar adecuadamente incluso en condiciones sin sobreajuste ni transformación del espacio de características.

Clase	Precisión	Recall	F1-Score	Support
Ictal	83%	66%	74%	3822
Interictal	84%	93%	88%	7200
Macro Avg.	84%	80%	81%	11022

Tabla 2. Tabla de resultados del modelo SVM sin SMOTE y sin PCA.

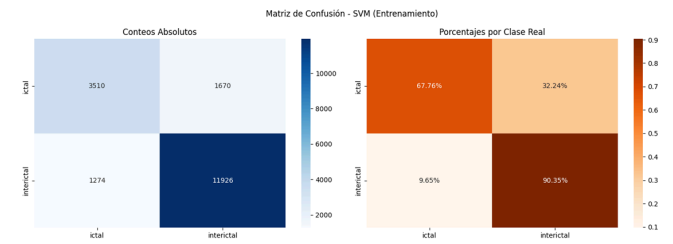


Figura 3. Matriz de confusión de entrenamiento del modelo SVM sin SMOTE y sin PCA.

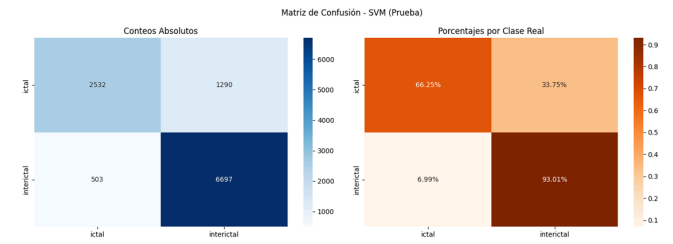


Figura 4. Matriz de confusión de prueba del modelo SVM sin SMOTE y sin PCA.

Por otra parte, el modelo Random Forest bajo las mismas condiciones (sin SMOTE ni PCA) logró el mejor recall para la clase ictal (69 %), lo cual resulta particularmente valioso en contextos clínicos donde la minimización de falsos negativos es crítica. Este resultado evidencia su capacidad para identificar con alta sensibilidad los episodios críticos, incluso si ello implica un ligero sacrificio en la precisión global.

Clase	Precisión	Recall	F1-Score	Support
Ictal	78%	69%	73%	3822
Interictal	84%	90%	87%	7200

Macro Avg.	81%	79%	80%	11022
------------	-----	-----	-----	-------

Tabla 3. Tabla de resultados del modelo Random Forest sin SMOTE y sin PCA.

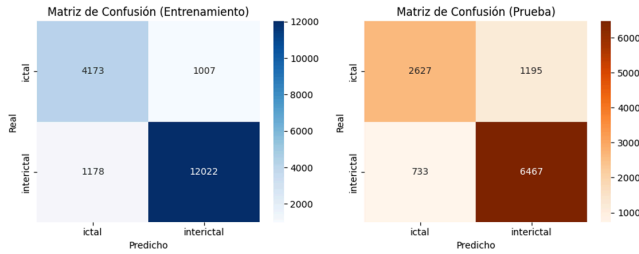


Figura 5. Matrices de confusión del modelo Random Forest sin SMOTE y sin PCA.

En general, se observó que el uso de PCA no mejoró el rendimiento de los modelos; de hecho, en la mayoría de los casos, su aplicación disminuyó tanto la exactitud como el recall, posiblemente debido a la pérdida de información discriminativa al reducir la dimensionalidad. De manera similar, la inclusión de datos sintéticos mediante SMOTE no ofreció beneficios significativos, lo que sugiere que el conjunto original de datos ya contenía suficiente representatividad de ambas clases para entrenar modelos robustos.

Finalmente, se muestra la tabla comparativa de resultados de las mejores variantes de cada modelo de ML, donde se resumen las métricas obtenidas para cada modelo y configuración evaluada:

Modelo	SMOTE	PCA	Accuracy (Prueba)	Recall Ictal (Prueba)
Regresión Logística	Sí	—	81 %	66 %
Random Forest	No	No	82 %	69 %
SVM	No	No	83 %	66 %
RNA (Backpropagation)	Sí	No	82 %	65 %

Tabla 4. Tabla de resumen de resultados de mejores modelos de ML.

V. CONCLUSIONES

Este trabajo demostró la viabilidad de aplicar modelos de aprendizaje supervisado para la detección de episodios epilépticos en pacientes pediátricos, mediante el análisis de señales EEG segmentadas y preprocesadas adecuadamente. En particular, el modelo SVM sin SMOTE ni PCA alcanzó la mayor exactitud (83 %), posicionándose como la opción más eficaz para lograr una clasificación precisa en promedio entre los periodos ictales e interictales.

Por otro lado, el modelo Random Forest sin SMOTE ni PCA obtuvo el mayor recall ictal (69 %), destacándose por su capacidad para identificar correctamente los episodios críticos. Este comportamiento lo convierte en la alternativa más robusta cuando el objetivo

clínico es minimizar los falsos negativos y garantizar una detección sensible de eventos potencialmente peligrosos.

Además, los resultados indican que ni la técnica de SMOTE ni la aplicación de PCA mejoraron el rendimiento de los modelos. Esto sugiere que un preprocesamiento adecuado junto con una buena selección de características puede ser suficiente para entrenar modelos con resultados competitivos, sin necesidad de técnicas adicionales de aumento o reducción de datos.

Como trabajo futuro, se recomienda enriquecer la caracterización de los segmentos ictales e interictales mediante el uso de nuevas variables que reflejen con mayor precisión la amplitud, la frecuencia fundamental y la morfología de las señales EEG. Aunque el presente estudio se basó principalmente en medidas como la entropía y la energía logarítmica para la extracción de características, sería valioso explorar otros descriptores como el ancho de banda espectral, la relación entre bandas de frecuencia, el análisis morfológico de complejos de onda, y métricas no lineales. Estas podrían aportar información complementaria que mejore la diferenciación entre clases y potencie el desempeño de los modelos de clasificación.

REFERENCIAS

- [1] A. H. Shoeb, *Patient-Specific Seizure Onset Detection*. Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, M.Eng. thesis, Aug. 2003.
- [2] A. Karasmanoglou, G. Giannakakis, P. Vorgia, M. Antonakakis, y M. Zervakis, "Semi-Supervised anomaly detection for the prediction and detection of pediatric focal epileptic seizures on fused EEG and ECG data," *Biomedical Signal Processing and Control*, vol. 101, p. 107083, 2025. doi: 10.1016/j.bspc.2024.107083.
- [3] A.-M. Tautan, A.-G. Andrei, C. L. Smeralda, G. Vatti, S. Rossi, y B. Ionescu, "Unsupervised learning from EEG data for epilepsy: A systematic literature review," *Artificial Intelligence in Medicine*, vol. 162, p. 103095, 2025. doi: 10.1016/j.artmed.2025.103095.
- [4] D.W. Klass and D.D. Daly. *Current Practice of Clinical Electroencephalography*. Raven Press, 1979
- [5] Gotman, J. *Automatic recognition of epileptic seizures in the eeg*. *Electroencephalography and Clinical Neurophysiology*, 54:530540, 1982
- [6] Gutttag, J. (2010). CHB-MIT Scalp EEG Database (version 1.0.0). PhysioNet. <https://doi.org/10.13026/C2K01R>.
- [7] M. A. AbuAlrob, A. Itbaisha, y B. Mesraoua, "Unlocking new frontiers in epilepsy through AI: From seizure prediction to personalized medicine," *Epilepsy & Behavior*, vol. 166, p. 110327, May 2025. doi: 10.1016/j.yebeh.2025.110327.