

**Implementación de EEG y modelos de aprendizaje automático para la detección de estados
de confusión en la educación a distancia**

Natalia Perea Calderón

Asesor

Rafael Gaitán Ospina

Universidad Nacional Abierta y a Distancia – UNAD

Escuela de Ciencias Básicas, Tecnología e Ingeniería

Especialización Ciencia de Datos y Analítica

2024

Dedicatoria

A todos los estudiantes y educadores que enfrentan los retos del aprendizaje en entornos virtuales, especialmente en épocas de transformación y adaptación. De manera especial, a la docente y madre Natali Calderón, cuya dedicación y pasión por la enseñanza han sido fuente de inspiración y motivación para este trabajo. Que este estudio sirva como un aporte para elevar la calidad de la educación, promoviendo un aprendizaje más inclusivo, personalizado y transformador.

Agradecimientos

Agradezco a la Universidad Nacional Abierta y a Distancia, por ser el espacio donde se sentaron las bases de este trabajo, proporcionando las herramientas, el conocimiento y el entorno académico necesarios para su desarrollo.

A los profesores que han sido parte fundamental de mi formación profesional, cuya paciencia, dedicación y compromiso con la enseñanza se ven reflejados en este trabajo. Con profundo reconocimiento, agradezco a mi asesor de proyecto, Rafael Gaitán, por su invaluable guía y orientación durante el desarrollo de esta investigación. Su experiencia y apoyo fueron esenciales para alcanzar los objetivos planteados.

Finalmente, a mi familia y amigos, cuya confianza, amor y apoyo incondicional me motivaron a esforzarme y dar lo mejor de mí en cada etapa de este proyecto.

Este trabajo es también fruto de su inspiración y aliento constante.

Resumen

El estudio analiza la detección de confusión en estudiantes durante el aprendizaje a distancia mediante señales EEG y modelos de aprendizaje automático. Utilizando datos de un dispositivo de 14 canales en 8 estudiantes, se aplicaron técnicas de balanceo (Undersampling y Oversampling) y se entrenaron modelos de clasificación (árboles de decisión, KNN y regresión logística). Los resultados revelan que el balanceo de datos mediante Oversampling ofrece un equilibrio óptimo entre sensibilidad, especificidad y precisión, destacándose el modelo KNN en la banda beta, asociada con la atención activa y la concentración. Además, se identificó que los electrodos en regiones temporales, parietales y occipitales (T7, T8, P7, P8, O1 y O2) tienen una mayor asociación con los estados de confusión, mientras que los electrodos frontales presentan menor relevancia.

Palabras claves: Aprendizaje virtual, balanceo de datos, EEG, estados de confusión, modelos de aprendizajes.

Abstract

The study examines the detection of confusion in students during remote learning using EEG signals and machine learning models. Data from a 14-channel device in 8 students were processed using balancing techniques (Undersampling and Oversampling), and classification models (decision trees, KNN, and logistic regression) were trained. The results reveal that data balancing through Oversampling provides an optimal balance between sensitivity, specificity, and precision, with the KNN model standing out in the beta band, associated with active attention and concentration. Additionally, electrodes in temporal, parietal, and occipital regions (T7, T8, P7, P8, O1, and O2) showed a stronger association with confusion states, while frontal electrodes were less relevant.

Keywords: Confusion states, data balancing, EEG, machine learning models, virtual learning.

Tabla de Contenido

Glosario.....	14
Introducción	16
Planteamiento del Problema	18
Justificación	20
Objetivos.....	22
Objetivo General.....	22
Objetivos Específicos.....	22
Marco Conceptual.....	23
Marco Teórico.....	28
Metodología	30
Análisis Estadístico para la Identificación de Patrones en Datos EEG	33
Correlaciones.....	34
Dominio Tiempo	34
Banda de Frecuencia Alfa	35
Banda de Frecuencia Beta	36
Banda de Frecuencia Gama.....	37
Prueba de Contingencia	39
Banda de Frecuencia Alfa	39
Banda de Frecuencia Beta	40
Banda de Frecuencia Gama.....	41
Modelos de Aprendizaje Automático	44
Árboles de Decisión.....	44

Datos sin Balanceo.....	45
Clasificación para Todos los Electrodo.....	45
Clasificación para Electrodo Específicos.....	47
Datos con Undersampling.....	49
Clasificación para Todos los Electrodo.....	49
Clasificación para Electrodo Específicos.....	52
Datos con Oversampling.....	53
Clasificación para Todos los Electrodo.....	53
Clasificación para Electrodo Específicos.....	56
KNN.....	57
Datos sin Balanceo.....	58
Clasificación para Todos los Electrodo.....	58
Clasificación para Electrodo Específicos.....	60
Datos con Undersampling.....	62
Clasificación para Todos los Electrodo.....	62
Clasificación para Electrodo Específicos.....	64
Datos con Oversampling.....	66
Clasificación para Todos los Electrodo.....	66
Clasificación para Electrodo Específicos.....	68
Regresión Logística.....	71
Datos sin Balanceo.....	72
Clasificación para Todos los Electrodo.....	73
Clasificación para Electrodo Específicos.....	75

Datos con Undersampling	76
Clasificación para Todos los Electrodo.....	76
Clasificación para Electrodo Específicos.....	78
Datos con Oversampling	81
Clasificación para Todos los Electrodo.....	81
Clasificación para Electrodo Específicos.....	83
Comparación de Modelos	85
Conclusiones	93
Referencias.....	94

Lista de Tablas

Tabla 1 <i>Coefficientes de Correlación y Valores p en el Dominio Tiempo.....</i>	34
Tabla 2 <i>Coefficientes de Correlación y Valores p en Banda Alfa.</i>	35
Tabla 3 <i>Coefficientes de Correlación y Valores p en Banda Beta.....</i>	36
Tabla 4 <i>Coefficientes de Correlación y Valores p en Banda Gama.</i>	37
Tabla 5 <i>Tabla de Contingencia Banda Alfa.....</i>	39
Tabla 6 <i>Tabla de Contingencia Banda Beta</i>	40
Tabla 7 <i>Tabla de Contingencia Banda Gama.....</i>	41
Tabla 8 <i>Comparación de Modelos sin Balanceo para Todos los Canales</i>	86
Tabla 9 <i>Comparación de Modelos sin Balanceo para Canales Específicos.</i>	87
Tabla 10 <i>Comparación de Modelos con Undersampling para Todos los Canales.</i>	88
Tabla 11 <i>Comparación de Modelos con Undersampling para Canales Específicos.</i>	89
Tabla 12 <i>Comparación de Modelos con Oversampling para Todos los Canales.</i>	90
Tabla 13 <i>Comparación de Modelos con Oversampling para Canales Específicos.....</i>	91

Lista de Figuras

Figura 1 <i>Sistema 10-20</i>	25
Figura 2 <i>Posicionamiento de Electrodo en el Sistema 10-20</i>	30
Figura 3 <i>Metodología</i>	32
Figura 4 <i>Curvas de Aprendizaje de Árboles de Decisión para Datos sin Balanceo para Todos los Canales</i>	46
Figura 5 <i>Matrices de Confusión de Árboles de Decisión para Datos sin Balanceo para Todos los Canales</i>	47
Figura 6 <i>Curva de Aprendizaje de Árboles de Decisión para Datos sin Balanceo para Canales Específicos</i>	48
Figura 7 <i>Matrices de Confusión de Árboles de Decisión para Datos sin Balanceo para Canales Específicos</i>	49
Figura 8 <i>Curvas de Aprendizaje de Árboles de Decisión para Datos con Undersampling para Todos los Canales</i>	50
Figura 9 <i>Matrices de Confusión de Árboles de Decisión para Datos con Undersampling para Todos los Canales</i>	51
Figura 10 <i>Curvas de Aprendizaje de Árboles de Decisión para Datos con Undersampling para Canales Específicos</i>	52
Figura 11 <i>Matrices de Confusión de Árboles de Decisión para Datos con Undersampling para Canales Específicos</i>	53
Figura 12 <i>Curvas de Aprendizaje de Árboles de Decisión para Datos con Oversampling para Todos los Canales</i>	54
Figura 13 <i>Matrices de Confusión de Árboles de Decisión para Datos con Oversampling para</i>	

<i>Todos los Canales.</i>	55
Figura 14 <i>Curvas de Aprendizaje de Árboles de Decisión para Datos con Oversampling para Canales Específicos.</i>	56
Figura 15 <i>Matrices de Confusión de Árboles de Decisión para Datos con Oversampling para Canales Específicos.</i>	57
Figura 16 <i>Curvas de Aprendizaje de KNN para Datos sin Balanceo para Todos los Canales.</i>	59
Figura 17 <i>Matrices de Confusión de KNN para Datos sin Balanceo para Todos los Canales.</i>	60
Figura 18 <i>Curvas de Aprendizaje de KNN para Datos sin Balanceo para Canales Específicos.</i>	61
Figura 19 <i>Matrices de Confusión de KNN para Datos sin Balanceo para Canales Específicos.</i>	62
Figura 20 <i>Curvas de Aprendizaje de KNN para Datos con Undersampling para Todos los Canales.</i>	63
Figura 21 <i>Matrices de Confusión de KNN para Datos con Undersampling para Canales Específicos.</i>	64
Figura 22 <i>Curvas de Aprendizaje de KNN para Datos con Undersampling para Canales Específicos.</i>	65
Figura 23 <i>Matrices de Confusión de KNN para Datos con Undersampling para Canales Específicos.</i>	66
Figura 24 <i>Curvas de Aprendizaje de KNN para Datos con Sobremuestro para Todos los Canales.</i>	67
Figura 25 <i>Matrices de Confusión de KNN para Datos con Sobremuestro para Todos los Canales.</i>	68
Figura 26 <i>Curvas de Aprendizaje de KNN para Datos con Sobremuestro para Canales Específicos.</i>	69

Figura 27 <i>Matrices de Confusión de KNN para Datos con Sobremuestro para Canales Específicos.</i>	71
Figura 28 <i>Curvas de Aprendizaje de Regresión Logística para Datos sin Balancear para Todos los Canales.</i>	73
Figura 29 <i>Matrices de Confusión de Regresión Logística para Datos sin Balancear para Todos los Canales.</i>	74
Figura 30 <i>Curvas de Aprendizaje de Regresión Logística para Datos sin Balancear para Canales Específicos.</i>	75
Figura 31 <i>Matrices de Confusión de Regresión Logística para Datos sin Balancear para Canales Específicos.</i>	76
Figura 32 <i>Curvas de Aprendizaje de Regresión Logística para Datos con Undersampling para Todos los Canales.</i>	77
Figura 33 <i>Matrices de Confusión de Regresión Logística para Datos con Undersampling para Todos los Canales.</i>	78
Figura 34 <i>Curvas de Aprendizaje de Regresión Logística para Datos con Undersampling para Canales Específicos.</i>	79
Figura 35 <i>Matrices de Confusión de Regresión Logística para Datos con Undersampling para Canales Específicos.</i>	80
Figura 36 <i>Curvas de Aprendizaje de Regresión Logística para Datos con Oversampling para Todos los Canales.</i>	82
Figura 37 <i>Matrices de Confusión de Regresión Logística para Datos con Oversampling para Todos los Canales.</i>	83
Figura 38 <i>Curvas de Aprendizaje de Regresión Logística para Datos con Undersampling para</i>	

Canales Específicos. 84

Figura 39 *Matrices de Confusión de Regresión Logística para Datos con Oversampling para*

Todos Canales Específicos. 85

Glosario

Aprendizaje automático: Rama de la inteligencia artificial que utiliza algoritmos para identificar patrones a partir de datos y hacer predicciones o clasificaciones basadas en ellos.

Árboles de decisión: Modelo de aprendizaje automático que representa decisiones y sus posibles consecuencias en forma de un árbol jerárquico.

Dominio frecuencia: Análisis de señales que se centra en las frecuencias que componen la señal, útil en estudios EEG para identificar patrones en diferentes bandas de frecuencia.

Dominio tiempo: Representación de señales en función del tiempo, que muestra cómo varían las amplitudes en intervalos temporales específicos.

EEG: Siglas de electroencefalografía, técnica no invasiva para medir la actividad eléctrica cerebral mediante electrodos colocados en el cuero cabelludo.

FFT: Siglas de Transformada Rápida de Fourier, algoritmo eficiente para calcular la Transformada Discreta de Fourier (DFT) y su inversa (IDFT).

Emotive Epoc X: Dispositivo EEG de 14 canales utilizado para registrar señales cerebrales, diseñado para investigaciones y aplicaciones prácticas en neurociencia.

Especificidad: Métrica que mide la proporción de verdaderos negativos correctamente identificados por un modelo, evaluando su capacidad para evitar falsos positivos.

Inteligencias múltiples: Teoría de Howard Gardner que postula que las personas poseen diferentes tipos de inteligencias.

KNN: Algoritmo de aprendizaje automático no paramétrico basado en el principio de que una instancia se clasifica según la categoría de sus vecinos más cercanos.

Oversampling: Técnica de balanceo de datos que incrementa artificialmente las muestras de la clase minoritaria para equilibrar conjuntos de datos desbalanceados.

Precisión: Métrica que mide la proporción de predicciones correctas realizadas por un modelo en relación con todas sus predicciones.

Regresión logística: Modelo estadístico utilizado para clasificar datos binarios, basado en una función logística que estima la probabilidad de que una instancia pertenezca a una categoría.

Sensibilidad: Métrica que mide la proporción de verdaderos positivos correctamente identificados por un modelo, reflejando su capacidad para detectar la clase positiva.

Sistema 10-20: Sistema internacional para la colocación de electrodos en el cuero cabelludo durante mediciones EEG, basado en distancias porcentuales entre electrodos.

Undersampling: Técnica de balanceo de datos que reduce el número de muestras de la clase mayoritaria para equilibrar conjuntos de datos desbalanceados.

Transformada de Fourier: Técnica matemática para descomponer señales en componentes de diferentes resoluciones, utilizada en análisis EEG para identificar características específicas en distintas bandas de frecuencia.

Introducción

El aumento exponencial en el uso de plataformas digitales como medio de aprendizaje ha transformado significativamente la educación, facilitando el acceso al conocimiento en diversos contextos y escalas. Sin embargo, esta transición ha suscitado preocupaciones en la comunidad académica sobre el impacto en la calidad del aprendizaje y en los procesos cognitivos de los estudiantes. En entornos virtuales, la falta de comunicación no verbal, una herramienta esencial en la educación presencial limita la capacidad de los docentes para identificar estados emocionales y cognitivos, como la confusión, lo que dificulta intervenir eficazmente en el proceso de aprendizaje. Este desafío se agrava en la educación a distancia, donde los estudiantes pueden enfrentarse a mayores distracciones y barreras para la comprensión. La detección temprana de estados de confusión es fundamental, ya que permite implementar estrategias personalizadas que contribuyen a mejorar el rendimiento académico, fomentar la motivación y fortalecer los procesos cognitivos necesarios para una comprensión profunda del contenido. En este contexto, el uso de tecnologías innovadoras, como la electroencefalografía (EEG), y el análisis de datos mediante modelos de aprendizaje automático, surge como una solución prometedora. Estas herramientas permiten evaluar de manera no invasiva y en tiempo real el estado cognitivo de los estudiantes, identificando patrones de confusión a partir de la actividad cerebral.

Este documento presenta un enfoque interdisciplinario que combina señales EEG con modelos de aprendizaje automático, evaluando su efectividad en la detección de estados de confusión en estudiantes que participan en entornos de aprendizaje virtual. A través del análisis de señales EEG captadas mediante dispositivos especializados, este estudio identifica los momentos en que los estudiantes experimentan confusión, permitiendo intervenciones inmediatas y personalizadas. Además, se evalúan diferentes estrategias de preprocesamiento y balanceo de

datos, como el Undersampling y el Oversampling, para mejorar la precisión de los modelos de clasificación y abordar los retos del desbalance de clases en los datos.

Planteamiento del Problema

El uso de plataformas digitales para el aprendizaje ha experimentado un crecimiento acelerado en los últimos años, impulsado, en gran medida, por la pandemia del COVID-19. Esta tendencia, según las proyecciones, continuará expandiéndose de manera exponencial (Sethi, Dabas, Dua, Dalawat, & Sethia, 2018). Sin embargo, este auge no está exento de desafíos. Diversos académicos han expresado preocupaciones acerca de la calidad y efectividad del aprendizaje en línea, señalando que el simple uso de herramientas digitales no garantiza una experiencia educativa de alta calidad (Arana-Llanes, González-Serna, Pineda-Tapia, & Rendón-Miranda, 2016). Aunque los avances tecnológicos han facilitado la transición al aprendizaje virtual, también han introducido nuevas complejidades, como la necesidad de abordar estilos de aprendizaje diversos y considerar condiciones de los estudiantes, que incluyen trastornos de atención, déficits de memoria, problemas de percepción, barreras lingüísticas y dificultades en el desarrollo (Lancheros-Cuesta, Arias, Forero, & Duran, 2018). Estudios como el de Tambe y Khachane (2016) han señalado que la falta de atención y concentración, comúnmente derivada de estados de confusión, es una de las principales causas de deserción en los entornos educativos. Este problema es particularmente prevalente en la educación a distancia, donde la interacción remota con contenidos educativos dificulta la identificación de estados cognitivos críticos, como la confusión, por parte de los educadores. La incapacidad para detectar estos estados a tiempo afecta significativamente el rendimiento académico de los estudiantes (Chen, Wang, & Yu, 2017).

Actualmente, las evaluaciones estandarizadas son herramientas frecuentes para medir el progreso de los estudiantes, pero su aplicación retrospectiva dificulta realizar ajustes oportunos en el proceso educativo. Dado que el aprendizaje es un proceso dinámico y continuo, es

imprescindible que los programas educativos se adapten de manera frecuente con base en retroalimentaciones inmediatas sobre el desempeño de los estudiantes (Hu & Kuo, 2017).

A pesar del creciente interés en el estudio de los comportamientos de los estudiantes en modalidades virtuales, existe una notable carencia de herramientas efectivas y confiables que permitan identificar y abordar la confusión de los estudiantes en tiempo real. La ausencia de estas herramientas limita la capacidad de los instructores para desarrollar contenidos personalizados y diseñar metodologías que se ajusten a los diferentes estilos de aprendizaje y necesidades individuales (Arana-Llanes, González-Serna, Pineda-Tapia, & Rendón-Miranda, 2016).

Justificación

La detección temprana y precisa de los estados de confusión en los estudiantes es fundamental para intervenir de manera efectiva y proporcionar el apoyo necesario que les permita superar los obstáculos en el proceso de aprendizaje. Cuando los estudiantes experimentan confusión, su capacidad para comprender y retener la información se ve comprometida, lo que afecta no solo su desempeño académico, sino también su motivación para continuar aprendiendo. Este problema es especialmente crítico en los entornos de aprendizaje a distancia o virtuales, donde la falta de interacción física limita la comunicación no verbal entre docentes y estudiantes, dificultando la detección de señales de confusión (Arana-Llanes, González-Serna, Pineda-Tapia, & Rendón-Miranda, 2016).

En este contexto, la implementación de tecnologías avanzadas como la electroencefalografía (EEG) y modelos de aprendizaje automático ofrece una solución innovadora. Los sensores EEG permiten monitorear de manera objetiva y no intrusiva la actividad cerebral de los estudiantes mientras participan en actividades de aprendizaje en línea. Estas señales contienen información valiosa sobre su estado emocional y cognitivo, que puede ser analizada mediante algoritmos de aprendizaje automático para identificar patrones asociados a la confusión.

Este análisis no solo permite generar alertas en tiempo real, sino también desarrollar estrategias pedagógicas personalizadas que respondan a las necesidades específicas de cada estudiante (Trigka, Dritsas, & Mylonas, 2023). Facilitando la implementación de metodologías y actividades educativas más efectivas y adaptativas. Al identificar los momentos en que los estudiantes experimentan confusión, los docentes pueden realizar ajustes oportunos en el contenido y en las estrategias de enseñanza, mejorando significativamente la calidad de la

educación a distancia. Esto contribuye a una experiencia de aprendizaje personalizada, donde los estudiantes se sienten más motivados y apoyados, lo que se traduce en un mejor rendimiento académico y una mayor satisfacción con el proceso educativo (Hu & Kuo, 2017).

La incorporación de tecnologías como el EEG en combinación con aprendizaje automático no solo aborda un problema crítico en la educación a distancia, sino que también representa un avance significativo hacia una educación más inclusiva, eficiente y adaptativa, optimizando la experiencia tanto para los estudiantes como para los educadores.

Objetivos

Objetivo General

Detectar estados de confusión en estudiantes que se forman en la modalidad de educación a distancia o virtual a través de señales de electroencefalografía (EEG) y modelos de aprendizaje automático.

Objetivos Específicos

Identificar patrones asociados a estados de confusión en señales de EEG de estudiantes.

Evaluar la efectividad y precisión de diferentes modelos de aprendizaje automático para detectar estados cognitivos.

Marco Conceptual

La educación es un proceso cultural profundamente vinculado a la esencia de la humanidad y refleja la dinámica de las sociedades modernas. Este proceso implica la adquisición de conocimientos del entorno y desempeña un papel esencial en la satisfacción de la necesidad humana de pertenencia, al tiempo que preserva la individualidad de cada persona (León, 2007). Además, la educación constituye una herramienta indispensable para la universalización del conocimiento, permitiendo a las personas mantener su identidad única mientras desarrollan habilidades para enfrentar los desafíos del entorno.

En este contexto, el sistema educativo colombiano se fundamenta en un marco normativo que responde a las necesidades de la sociedad. Según el Artículo 2.5.3.2.2.5 del Decreto 1075 de 2015, modificado por el Decreto 1330 de 2019, la modalidad educativa se define como:

” Modo utilizado que integra un conjunto de opciones organizativas y/o curriculares que buscan dar respuesta a requerimientos específicos del nivel de formación y atender características conceptuales que faciliten el acceso a los estudiantes, en condiciones diversas de tiempo y espacio” (Mineducación, 2022).

La modalidad educativa es un eje crucial en la estructura del sistema, ya que permite a las instituciones adaptar sus ofertas académicas a las necesidades específicas de los estudiantes, independientemente de su ubicación o disponibilidad horaria. De acuerdo con el Ministerio de Educación (2022) existen cuatro modalidades principales en el sistema educativo: presencial, remota, virtual y dual. La modalidad remota facilita el acceso a recursos educativos mediante métodos no presenciales, sin depender completamente de las tecnologías de la información y las comunicaciones (TIC). Por otro lado, la modalidad virtual aprovecha las TIC para ofrecer experiencias de aprendizaje más interactivas, integrando recursos como plataformas digitales y

entornos telemáticos. Este enfoque ha sido ampliamente adoptado por instituciones de educación superior, ya que promueve estrategias de aprendizaje autónomo y fomenta el desarrollo de competencias específicas (Betancourt, Bruges, & Sfer, 2017).

Sin embargo, el éxito de estas modalidades no depende exclusivamente de la tecnología. Al-Khatir Al-Arimi (2014) destacó la importancia de desarrollar competencias específicas para garantizar un desempeño académico exitoso en entornos de aprendizaje a distancia. La elección de modelos educativos adecuados, que se alineen con las necesidades individuales de los estudiantes, es crucial para optimizar los resultados de aprendizaje y minimizar factores como la falta de atención o concentración, los cuales suelen generar confusión y bajo rendimiento (Arana-Llanes, González-Serna, Pineda-Tapia, & Rendón-Miranda, 2016).

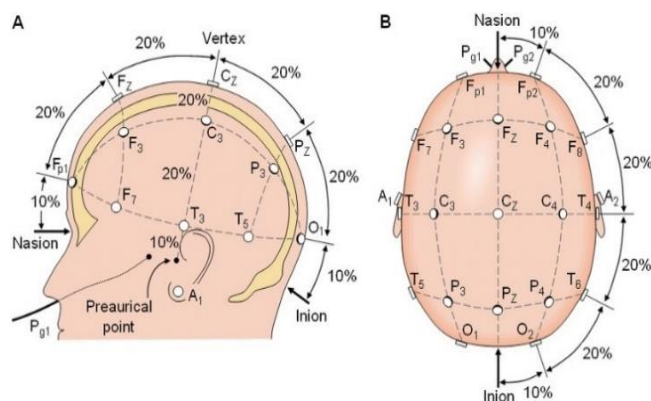
El estado emocional de confusión se reconoce como un estado cognitivo complejo que surge con frecuencia durante las tareas de aprendizaje y resolución de problemas (Liang, Liu, Qiu, & Zhang, 2018). En la educación virtual o a distancia, donde las oportunidades de interacción entre profesores y estudiantes son limitadas, los estudiantes frecuentemente pueden experimentar confusión debido a la falta de comprensión o a las distracciones. Adicionalmente, la evaluación de los estados de confusión de los estudiantes se complica por la ausencia de señales no verbales, que normalmente utilizan los profesores para evaluar el nivel de comprensión de los estudiantes durante la educación presencial (Luo, Zhou, Yang, & He, 2023). En este contexto, es crucial desarrollar métodos que permitan detectar estados de confusión en estudiantes que utilizan modalidades virtuales, para así adaptar los procesos de enseñanza y mejorar el rendimiento académico. Una de las técnicas emergentes para lograr esto es el análisis de señales de electroencefalografía (EEG), que permite monitorear la actividad cerebral y correlacionarla con estados cognitivos específicos.

El cerebro, órgano central de las funciones cognitivas humanas, está compuesto por varias regiones que desempeñan roles distintos. El lóbulo frontal es responsable de la planificación, organización, resolución de problemas y regulación emocional. Incluye el área prefrontal, asociada con el control ejecutivo, y el área frontal, que participa en la atención y memoria de trabajo. El temporal se encarga de procesar estímulos auditivos, reconocimiento facial y memoria episódica. Es fundamental para el procesamiento del lenguaje. El lóbulo parietal está involucrado en el procesamiento sensorial y la percepción espacial. Mientras que el lóbulo occipital es responsable del procesamiento visual, incluido el reconocimiento de formas, colores y movimientos (Fuentes-Martinez, 2024) (Vachha, Massoud, & Huang, 2022).

El EEG es una técnica no invasiva utilizada para medir la actividad eléctrica del cerebro mediante electrodos colocados en el cuero cabelludo. Introducida por Hans Berger en 1929, esta técnica permite analizar las señales cerebrales en términos de frecuencia y amplitud (Casson, Abdulaal, Dulabh, Kohli, & Trimble, 2018). El sistema internacional 10/20 (figura 1) es el estándar más utilizado para la colocación de electrodos, lo que facilita la comparación de datos entre diferentes estudios y poblaciones (UNAM, 2019).

Figura 1

Sistema 10-20



Nota. Imagen tomada de Novo-Olivas y cols.,

Las señales EEG se clasifican en bandas de frecuencia asociadas con estados específicos de actividad cerebral. La banda Delta (<4 Hz) está vinculada con el sueño profundo y la actividad de la materia gris. Theta (4-8 Hz) se relaciona con estados de relajación profunda y meditación. Alfa (8-13 Hz) predomina durante la relajación con los ojos cerrados. Beta (13-30 Hz) está presente en estados de atención, resolución de problemas y toma de decisiones y Gama (>30 Hz) se asocia con la hiper alerta y los procesos cognitivos complejos (Abo-Zahhad, Ahmed, & Abbas, 2015).

El procesamiento de señales EEG comienza con la adquisición de datos, seguida de la identificación de períodos relevantes, la eliminación de artefactos y el filtrado de ruido. Posteriormente, se extraen características clave, que son normalizadas y utilizadas en algoritmos de clasificación para identificar patrones en los datos (Casson, Abdulaal, Dulabh, Kohli, & Trimble, 2018). Los clasificadores son un tipo de modelo de aprendizaje automático diseñado para desarrollar predicciones o clasificaciones a partir de conjuntos de datos extensos. Clasificados en dos tipos principales de aprendizaje: aprendizaje supervisado y no supervisado, los clasificadores permiten la interpretación de datos complejos y proporcionan información sobre los patrones de datos (Casson, Abdulaal, Dulabh, Kohli, & Trimble, 2018). En el aprendizaje supervisado, el modelo se entrena con datos etiquetados y proporciona retroalimentación directa. Por el contrario, en el aprendizaje no supervisado, el modelo se entrena con datos sin etiquetar y no proporciona retroalimentación. El aprendizaje supervisado se utiliza para construir modelos de regresión que puedan hacer predicciones sobre datos futuros basándose en datos conocidos. Específicamente, se aprovecha para predecir datos cuantitativos cuando las entradas son del mismo tipo y se pueden ordenar de acuerdo con lo que se espera de la salida. De manera similar, los modelos de clasificación se utilizan cuando el resultado es categórico (Kyriakides & Margaritis, 2019). Los modelos de aprendizaje supervisado se utilizan ampliamente en el campo

de las neurociencias. Varios investigadores como Ni, Yuskel, Mandel & Xie (2017), Sun & otros (2019) y Dakure, Benlamine & Frasson (2021) han aplicado algoritmos refinados de aprendizaje automático. Entre los clasificadores comúnmente utilizados se encuentra K-Nearest Neighbors (KNN), un método no paramétrico que clasifica nuevos datos basándose en los vecinos más cercanos en el espacio de características (Dakoure, Benlamine, & Frasson, 2021). También está el modelo de regresión logística, que modela la probabilidad de un resultado binario en función de variables independientes y ha sido utilizado en estudios para detectar estados cognitivos como la confusión en estudiantes (García Amaya, Aceves Fernández, Tovar Arriaga, & Ramos Arreguín, 2020). Por último, los árboles de decisión son modelos predictivos basados en reglas derivadas de los datos, útiles para interpretar patrones complejos (Dakoure, Benlamine, & Frasson, 2021).

Marco Teórico

Diversos investigadores han centrado su atención en estudiar los mecanismos del cerebro durante el aprendizaje y en desarrollar técnicas que optimicen la educación, haciéndola más personalizada y significativa para los estudiantes. Este interés ha cobrado especial relevancia tras el incremento en el uso de plataformas digitales para el aprendizaje durante la pandemia, lo que ha intensificado la necesidad de comprender cómo se da el aprendizaje en estos entornos.

Por ejemplo, Ahire, Awale & Wagh (2023) implementaron modelos de aprendizaje automático como KNN, Naive Bayes y Random Forest para clasificar señales de EEG mientras los participantes realizaban actividades de aprendizaje. Su estudio se enfocó en sujetos con Trastorno por Déficit de Atención e Hiperactividad (TDAH) y encontró que el clasificador Naive Bayes alcanzó una precisión del 96 %. Las características más relevantes para el diagnóstico se localizaron en las regiones frontales, centrales y parietales del cerebro, áreas críticas para la atención y el control ejecutivo.

De manera complementaria, García Amaya, Aceves Fernández, Tovar Arriaga & Ramos Arreguín (2020) centraron sus esfuerzos en identificar características del EEG útiles para diagnosticar TDAH mediante aprendizaje automático. Utilizaron regresión logística para analizar la banda beta en electrodos posicionados en las regiones centrales, parietales y frontales del cerebro. El modelo de regresión logística alcanzó una precisión del 80.30 %, mientras que el modelo de Máxima Verosimilitud logró una precisión del 84.40 %. Estos resultados confirmaron que la actividad en la banda beta es un indicador sólido para diferenciar entre individuos con TDAH y grupos de control. Además, los modelos utilizados destacaron por su simplicidad y eficacia, siendo comparables con técnicas más avanzadas como SVM y análisis de componentes principales.

Por otro lado, estudios como el de Fuentes Martínez (2024) proponen plataformas EEG para monitorear la atención grupal en entornos educativos. Este trabajo buscó correlacionar la actividad en las bandas alfa y beta con niveles de atención percibidos mediante cuestionarios autoadministrados. Se encontró una fuerte correlación entre la actividad en la banda beta y estados de atención sostenida durante tareas cognitivas demandantes. La plataforma permitió a los docentes identificar momentos de baja atención grupal, sugiriendo intervenciones pedagógicas oportunas.

Sin embargo, también evidenció diferencias individuales en la respuesta a estímulos educativos, destacando la necesidad de adaptar métodos de enseñanza a perfiles específicos.

En un contexto similar, Hu, Li, Sun & Ratcliffe (2016) presentaron un sistema para identificar niveles de atención en estudiantes durante el aprendizaje afectivo en educación a distancia. Utilizando el clasificador KNN combinado con el método de selección de atributos basada en correlación (CFS), lograron una tasa de clasificación correcta del 80.84 %. Los niveles de atención se evaluaron mediante la escala SAM (Self-Assessment Manikin), que clasifica emociones en términos de valencia (positivo-negativo) y activación (calma-excitación). Se observaron correlaciones significativas entre la actividad EEG en las bandas alfa y beta y los niveles de atención reportados, donde la actividad beta aumentó en estados de alta atención y la actividad alfa predominó en estados de baja atención.

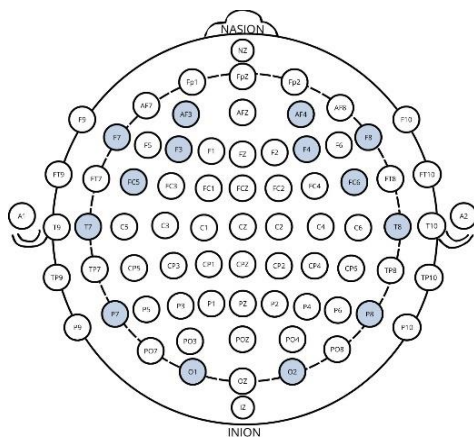
Estos estudios resaltan el potencial de las señales EEG y los modelos de aprendizaje automático para comprender mejor los procesos cognitivos durante el aprendizaje y optimizar la educación personalizada, tanto en entornos presenciales como virtuales.

Metodología

Se implementó la base de datos EEG data / Distance learning disponible en Kaggle (Omar, 2022), que recoge señales de electroencefalografía (EEG) utilizando el headset Emotiv Epoc X de 14 canales. La figura 2 muestra la disposición de los electrodos en el casco, según el sistema internacional 10-20. Durante el estudio, un grupo de 8 estudiantes con diferentes niveles educativos participó en una actividad interactiva en la que se les consultó sobre los temas que les resultaban más fáciles y difíciles de aprender. Los participantes realizaron una lectura en línea mientras sus señales EEG eran registradas. Además, se les hicieron preguntas para evaluar su nivel de comprensión, lo que permitió seleccionar videos adaptados a su nivel de entendimiento. Posteriormente, se presentaron los videos seleccionados a cada estudiante mientras se registraban las señales de EEG. A partir de las respuestas obtenidas y de una serie de preguntas relacionadas con la comprensión del contenido, se construyó una variable binaria, donde un valor de 1 indicaba que el estudiante comprendió el contenido, mientras que un valor de 0 señalaba que no lo comprendió.

Figura 2

Posicionamiento de Electroodos en el Sistema 10-20



El análisis y procesamiento de la base de datos se llevó a cabo utilizando el entorno Python, donde se verificó la integridad de los datos, asegurándose de que estuvieran completos y libres de valores atípicos. El procedimiento se describe en la figura 3.

En primer lugar, se realizó un preprocesamiento de los datos, visualizando cada canal en el dominio de la frecuencia para identificar posibles ruidos provocados por artefactos y electromagnetismo. Posteriormente, se aplicó un filtro para reducir la tendencia de la señal utilizando una ventana móvil. Con el fin de evitar sesgos en el modelo, se realizó un proceso de shuffling, lo que permitió mezclar todos los datos. Al examinar la distribución de la variable objetivo, se identificó un desbalance entre las clases, por lo que se implementaron técnicas de remuestreo para equilibrar los datos, incluyendo el Undersampling y el Oversampling. Como resultado, se generaron tres conjuntos de datos: uno sin balanceo, otro con balanceo utilizando Undersampling, y otro con balanceo mediante Oversampling. A cada uno de estos conjuntos se les realizaron pruebas estadísticas para extraer características clave que permitieran determinar la relación de cada canal de EEG con la variable objetivo.

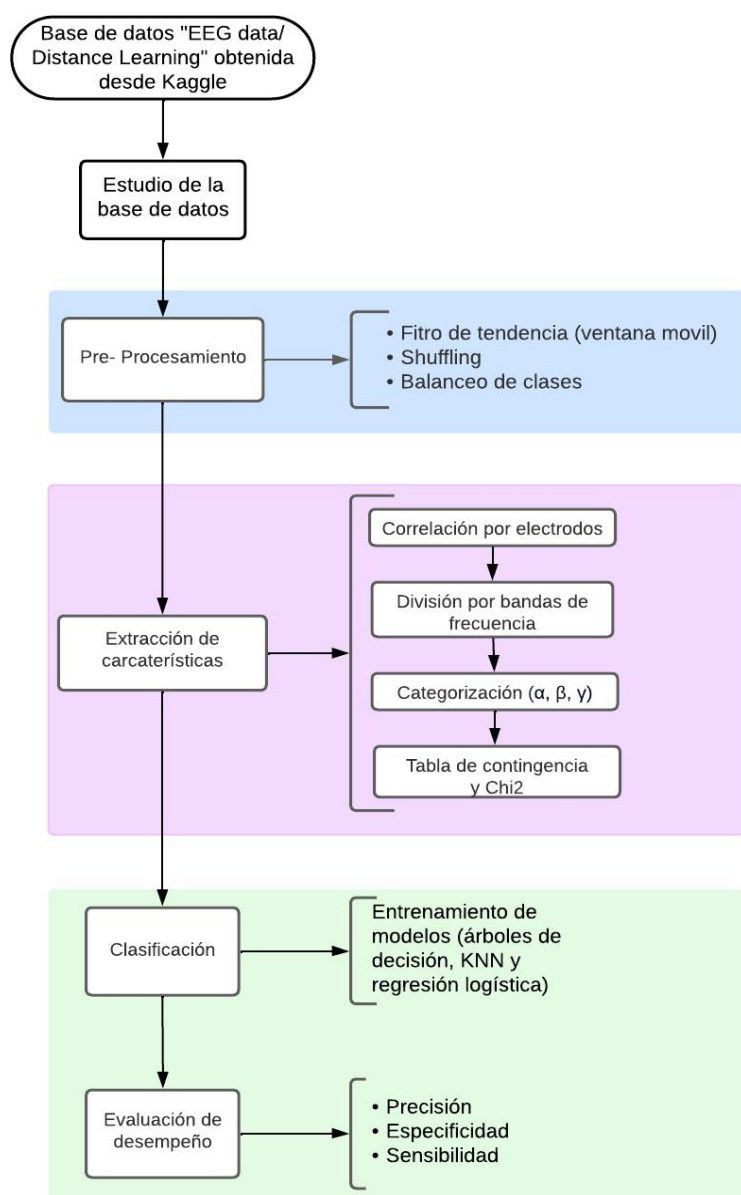
Se calculó la correlación de Pearson para todos los canales en el dominio del tiempo y en las bandas de frecuencia seleccionadas (alfa, beta y gama). Además, se calculó la estadística descriptiva para cada banda de frecuencia. Posteriormente, utilizando la mediana como umbral discriminante, se categorizó la señal en valores altos y bajos. Con esta categorización, se realizó una prueba de contingencia para estimar los valores de chi-cuadrado y su valor p.

Finalmente, se entrenaron tres modelos de aprendizaje automático árboles de decisión, KNN y regresión logística utilizando tanto los datos de todos los electrodos como los de aquellos electrodos que presentaron los valores más altos en correlación y en la prueba de contingencia. Esto resultó en 18 modelos distintos, para cada uno de ellos se estimaron las métricas de

precisión, especificidad y sensibilidad, con el fin de determinar qué procesamiento de datos, qué electrodos y qué bandas de frecuencia son más significativos y permiten obtener los mejores resultados. Los modelos fueron evaluados para identificar cuáles de estos factores (pre-procesamiento, selección de canales y bandas) optimizan el rendimiento del modelo predictivo en términos de clasificación de la comprensión del contenido por parte de los estudiantes.

Figura 3

Metodología



Análisis Estadístico para la Identificación de Patrones en Datos EEG

La identificación de patrones asociados con estados de confusión en estudiantes puede ofrecer información crucial para mejorar la personalización de las estrategias educativas, especialmente en entornos de aprendizaje virtual. Este capítulo se centra en la aplicación de métodos estadísticos para identificar y caracterizar patrones de actividad cerebral relacionados con momentos de confusión en estudiantes, utilizando datos de EEG obtenidos durante actividades de aprendizaje en línea.

Inicialmente, se estiman las correlaciones de Pearson entre las señales EEG de diferentes regiones cerebrales usando los diferentes canales de electrodos con la variable objetivo (estados de confusión), lo que permite evaluar las relaciones entre la actividad cerebral regional y los estados de comprensión o confusión de los estudiantes. A continuación, se implementa un proceso de categorización que facilita la aplicación de una prueba de contingencia, con el objetivo de examinar la relación entre los electrodos y la variable objetivo en las diferentes bandas de frecuencia. Este enfoque estadístico permite extraer variables clave que podrían diferenciar de manera efectiva entre los estados de comprensión y los momentos de confusión.

Los análisis realizados incluyen el cálculo de los coeficientes de Pearson y los valores p para cada banda de frecuencia relevante (alfa, beta y gama), así como la estimación de la chi-cuadrado en función de la categorización realizada. Estos resultados permiten identificar características distintivas en las señales EEG que pueden ser útiles para predecir y diferenciar los estados cognitivos de los estudiantes durante el aprendizaje. Al final de este capítulo, se sientan las bases para el desarrollo de modelos predictivos que, mediante el análisis de señales EEG, puedan identificar los momentos de confusión.

Correlaciones

A través de la estimación de la correlación de Pearson, se busca cuantificar el grado de asociación lineal entre los electrodos, las distintas bandas de frecuencia y el dominio temporal, en relación con los estados de confusión experimentados por los estudiantes. Para ello, se empleó la función `pearsonr` de la librería SciPy, que permite calcular el coeficiente de correlación de Pearson para cada electrodo en función de la variable objetivo, junto con el valor p correspondiente, el cual determina la significancia estadística de la correlación observada. Este procedimiento fue implementado tanto para los datos en el dominio temporal como para las bandas de frecuencia alfa, beta y gama, que han demostrado, según la literatura, estar asociadas con procesos cognitivos complejos. La evaluación de estas correlaciones proporciona una base sólida para identificar las características de las señales EEG más relevantes para la detección de estados de confusión, permitiendo avanzar hacia el desarrollo de modelos predictivos más precisos.

Dominio Tiempo

La tabla 1 presenta los resultados de la correlación entre los datos en el dominio del tiempo y la variable objetivo, proporcionando una visión general de la relación lineal a lo largo de las señales EEG.

Tabla 1

Coefficientes de Correlación y Valores p en el Dominio Tiempo

Electrodo	Coefficiente de Pearson	Valor- p
AF3	0.200039	0.000000e+00
F7	0.173616	0.000000e+00
F3	0.098953	2.738866e-149
FC5	0.122829	1.638414e-229

Electrodo	Coefficiente de Pearson	Valor-p
T7	0.232229	0.000000e+00
P7	0.251239	0.000000e+00
O1	0.257016	0.000000e+00
O2	0.205388	0.000000e+00
P8	0.249387	0.000000e+00
T8	0.299826	0.000000e+00
FC6	-0.109299	6.967783e-182
F4	0.113391	1.091280e-195
F8	0.019453	3.330544e-07
AF4	-0.083702	3.116604e-107

Nota. Coeficientes de correlación de Pearson y valores p para cada electrodo en el dominio tiempo.

Banda de Frecuencia Alfa

La tabla 2 muestra las correlaciones obtenidas entre los datos y la variable objetivo específicamente en la banda de frecuencia alfa, la cual se asocia comúnmente con estados de relajación y procesamiento cognitivo.

Tabla 2

Coeficientes de Correlación y Valores p en Banda Alfa

Electrodo	Coefficiente de Pearson	Valor-p
AF3	0.054161	7.000891e-46
F7	0.045437	8.671897e-33
F3	0.029665	7.037635e-15
FC5	0.030865	5.549568e-16

Electrodo	Coefficiente de Pearson	Valor-p
T7	0.063009	1.717605e-61
P7	0.069409	2.960041e-74
O1	0.068978	2.340861e-73
O2	0.054286	4.375389e-46
P8	0.067478	2.835241e-70
T8	0.083292	3.351398e-10
FC6	-0.02599	9.016668e-12
F4	0.034467	1.502730e-19
F8	0.005261	1.675379e-01
AF4	-0.01758	3.959604e-06

Nota. Coeficientes de correlación de Pearson y valores p para cada electrodo en la banda de frecuencia alfa.

Banda de Frecuencia Beta

La tabla 3 presenta las correlaciones correspondientes a la banda de frecuencia beta, que está vinculada a la atención activa y la concentración

Tabla 3

Coeficientes de Correlación y Valores p en Banda Beta

Electrodo	Coefficiente de Pearson	Valor-p
AF3	0.104705	5.295861e-167
F7	0.092764	2.307788e-131
F3	0.047994	2.172958e-36
FC5	0.063170	8.489714e-62
T7	0.123248	4.439677e-231

P7	0.134326	1.776707e-274
O1	0.138155	2.249357e-290
O2	0.107786	6.555476e-177
P8	0.127773	2.433510e-248
T8	0.158685	0.000000e+00
FC6	-0.065955	3.255904e-67
F4	0.054293	4.262371e-46
F8	0.001861	6.253673e-01
AF4	-0.050014	2.260146e-39

Nota. Coeficientes de correlación de Pearson y valores p para cada electrodo en la banda de frecuencia beta.

Banda de Frecuencia Gama

La tabla 4 muestra los resultados de la correlación en la banda de frecuencia gama, que se asocia con procesos cognitivos de alto nivel, como la resolución de problemas y la percepción consciente.

Tabla 4

Coeficientes de Correlación y Valores p en Banda Gama

Electrodo	Coeficiente de Pearson	Valor-p
AF3	0.104147	3.074091e-165
F7	0.090608	2.144907e-125
F3	0.050284	8.809409e-40
FC5	0.060669	3.914221e-57
T7	0.121329	5.972558e-224
P7	0.127673	5.970368e-248

O1	0.129802	2.730799e-256
O2	0.104349	7.063377e-166
P8	0.129434	7.738090e-255
T8	0.155203	0.000000e+00
FC6	-0.055232	1.207176e-47
F4	0.060124	3.837199e-56
F8	0.014006	2.383652e-04
AF4	-0.041241	2.660718e-27

Nota. Coeficientes de correlación de Pearson y valores p para cada electrodo en la banda de frecuencia gama.

Los resultados obtenidos en las tablas 1, 2, 3 y 4 para los valores p muestran que todas las correlaciones son estadísticamente significativas al estar por debajo del umbral común del valor p 0.05. Esto respalda la existencia de una asociación entre las señales registradas por los electrodos y la variable objetivo.

Los electrodos T7, P7, O1, O2, P8 y T8 destacan de manera consistente en todas las tablas por presentar coeficientes de correlación más altos en comparación con los demás. Esto sugiere que las regiones temporal, parietal y occipital tienen una relación relativamente más fuerte con el estado de confusión. Sin embargo, los coeficientes son bajos en términos absolutos, lo que implica que la magnitud de la correlación es débil. Por otro lado, electrodos como F8, FC6, F3 y AF4 presentan coeficientes de correlación cercanos a cero o incluso negativos, lo que indica que su relación con la variable objetivo es limitada o, en algunos casos, inversa. Lo cual supone que las regiones frontales están menos involucradas en procesos cognitivos relacionados con la confusión y concentración.

Prueba de Contingencia

Para evaluar la existencia de una relación significativa entre los electrodos y la variable objetivo, se implementó una prueba de contingencia. Esta prueba permite comparar las frecuencias observadas de las categorías de los electrodos con las frecuencias esperadas bajo la hipótesis de independencia. Para ello, se realizó una categorización de los datos basada en medidas estadísticas. La mediana de cada electrodo se utilizó como umbral para clasificar los valores de cada señal en dos categorías: alta o baja, para cada una de las bandas de frecuencia consideradas (alfa, beta y gama). Una vez realizada esta clasificación, se construyó una tabla de contingencia que muestra las frecuencias observadas para cada combinación de las categorías de los electrodos (alta o baja) y la variable objetivo. Posteriormente, se calculó el estadístico chi-cuadrado sumando las diferencias entre las frecuencias observadas y las frecuencias esperadas, divididas por las frecuencias esperadas. Este valor se utiliza para calcular el valor p, el cual permite evaluar la significancia estadística de la relación entre las categorías de los electrodos y la variable objetivo.

Banda de Frecuencia Alfa

La tabla de contingencia 5 muestra la estimación del valor chi-cuadrado y la significancia estadística resultante de la categorización de cada canal de los electrodos en comparación con la variable objetivo.

Tabla 5

Tabla de Contingencia Banda Alfa

Electrodo	Chi2	Valor-p
AF3	99.743239	1.734919e-23
F7	86.052857	1.751899e-20

F3	28.952555	7.417290e-08
FC5	32.672378	1.090763e-08
T7	150.184606	1.579826e-34
P7	183.066257	1.037439e-41
O1	196.563889	1.174086e-44
O2	151.688307	7.412237e-35
P8	158.329608	2.621926e-36
T8	270.423727	9.174028e-61
FC6	37.057912	1.146724e-09
F4	42.487757	7.112604e-11
F8	0.761464	3.828707e-01
AF4	6.743847	9.407157e-03

Banda de Frecuencia Beta

La tabla de contingencia 6 presenta la estimación del valor chi-cuadrado y la significancia estadística obtenida al categorizar los datos de cada canal de los electrodos en la banda de frecuencia beta, comparándolos con la variable objetivo. Esta tabla facilita la evaluación de la relación entre las categorías de los electrodos y los estados de confusión o comprensión, permitiendo verificar si existe una dependencia significativa entre las señales de EEG en la banda beta y la variable objetivo.

Tabla 6

Tabla de Contingencia Banda Beta

Electrodo	Chi2	Valor-p
AF3	317.769223	4.434228e-71
F7	353.014631	9.346928e-79

F3	97.837829	4.540893e-23
FC5	225.824471	4.852790e-51
T7	480.667384	1.529405e-106
P7	544.648081	1.837639e-120
O1	632.633443	1.336822e-139
O2	514.999532	5.181672e-114
P8	490.971730	8.757399e-109
T8	838.794773	1.984382e-184
FC6	208.361197	3.129051e-47
F4	117.418430	2.324667e-27
F8	0.436344	5.088921e-01
AF4	70.313250	5.059670e-17

Banda de Frecuencia Gama

La tabla de contingencia 7 muestra los resultados del chi-cuadrado y la significancia estadística para la categorización de los electrodos en la banda de frecuencia gama. Al igual que en las otras bandas, se evalúa la relación entre las categorías de los electrodos y la variable objetivo, con el fin de determinar si existe una asociación significativa en esta frecuencia, que podría ser indicativa de patrones relacionados con los estados de confusión en los estudiantes.

Tabla 7

Tabla de Contingencia Banda Gama

Electrodo	Chi2	Valor-p
AF3	347.985003	1.163958e-77
F7	361.542490	1.299311e-80
F3	108.711059	1.877542e-25

FC5	219.168330	1.373389e-49
T7	498.265291	2.266779e-110
P7	487.554005	4.853280e-108
O1	617.079111	3.228808e-136
O2	527.373933	1.052608e-116
P8	562.754417	2.115647e-124
T8	930.509157	2.288855e-204
FC6	165.951008	5.669801e-38
F4	135.959067	2.036864e-31
F8	2.976303	8.449206e-02
AF4	69.452828	7.826411e-17

Las tablas 5, 6 y 7 evidencia que, en las tres bandas de frecuencia, los electrodos T8, T7, P7, P8, O1 y O2 se destacan por sus altos valores de Chi-cuadrado y asociaciones estadísticamente significativas. Concordando con los resultados obtenidos en la prueba de correlación de Pearson, se puede establecer que estas regiones podrían estar más involucradas en los procesos corticales asociados con el estado de confusión.

Las bandas beta y gama muestran asociaciones más fuertes en comparación con alfa, lo que podría implicar que las señales de EEG de mayor frecuencia son más relevantes para comprender el estado de confusión. Por otro lado, los electrodos frontales como F8 y AF4 presentan asociaciones más débiles, sugiriendo que estas regiones podrían no ser tan relevantes en este contexto.

Estos resultados destacan la importancia de las regiones cerebrales representadas por los electrodos occipitales, parietales y temporales en los procesos mentales vinculados a los estados de confusión y comprensión. Las asociaciones identificadas en estas áreas sugieren que

desempeñan un papel central en la dinámica neuronal relacionada con dichos estados cognitivos. Esto es consistente con su participación en funciones como la integración sensorial, el procesamiento visual-espacial y la memoria, que son esenciales en la comprensión y la regulación cognitiva (Fraga, 2024) (Impulsa, 2018).

Modelos de Aprendizaje Automático

Los resultados presentados en el Capítulo 1 revelaron una mayor asociación de los electrodos ubicados en las regiones temporales, parietales y occipitales con la variable objetivo. En base a estos hallazgos, se propone entrenar y evaluar tres modelos de aprendizaje automático: árboles de decisión, K-Nearest Neighbors (KNN) y regresión logística. Se abordarán dos enfoques para el entrenamiento de los modelos: uno utilizando todos los canales de los electrodos, con el objetivo de evaluar el desempeño global considerando toda la información disponible, y otro enfocado exclusivamente en los canales de las regiones temporales, parietales y occipitales, seleccionados con base en su relevancia en los análisis de correlación y pruebas de contingencia.

Este segundo enfoque busca optimizar la precisión de los modelos y explorar si el uso de canales específicos mejora su rendimiento, al mismo tiempo que se reduce la dimensionalidad de los datos, lo cual podría facilitar el proceso de interpretación y aumentar la eficiencia computacional. Además, se evaluarán los tres modelos utilizando tres tipos de conjuntos de datos: sin balanceo, con Undersampling y con Oversampling. Esto dará lugar al entrenamiento de un total de 18 modelos, lo que permitirá comparar el impacto del balanceo de datos en el rendimiento de cada modelo.

Árboles de Decisión

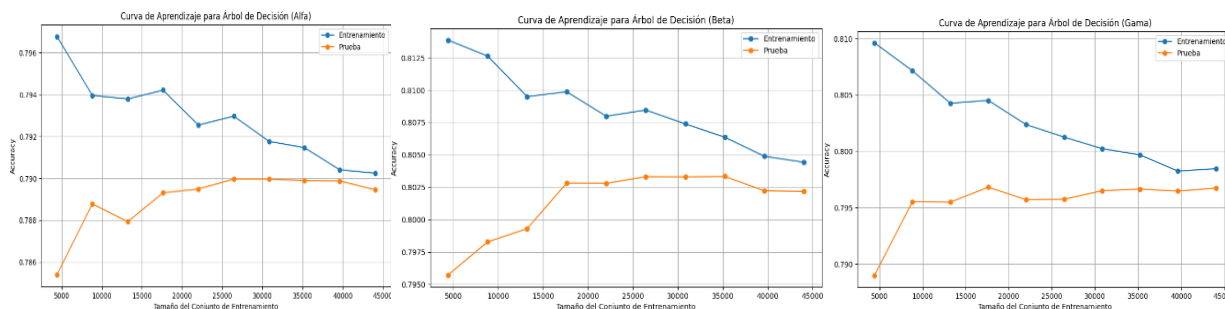
Se implementó la función Decision Tree Classifier de la librería scikit-learn para entrenar los modelos de árboles de decisión. Estos modelos fueron configurados utilizando el criterio de impureza de Gini y una profundidad máxima de 3 niveles para cada conjunto de datos. Esta configuración permitió controlar la complejidad de los árboles, evitando el sobreajuste y facilitando la interpretación de los modelos. Los árboles fueron entrenados y evaluados en los tres conjuntos de datos generados: sin balanceo, Undersampling y Oversampling, para analizar

cómo estas técnicas de manejo de clases desbalanceadas afectan el rendimiento del modelo.

Datos sin Balanceo

La sección presenta la aplicación de los modelos de árbol de decisión sobre el conjunto de datos sin balancear. Esta sección se subdivide en dos categorías: una que utiliza todos los electrodos disponibles y otra que se enfoca únicamente en aquellos electrodos seleccionados por su mayor relevancia estadística, determinada a partir de los análisis de correlación y pruebas de contingencia. De esta manera, se busca comparar el rendimiento del modelo en función de la cantidad de información utilizada, evaluando si la selección de los electrodos más relevantes mejora la precisión del modelo al reducir la dimensionalidad de los datos.

Clasificación para Todos los Electrodo. La figura 4 muestra las curvas de aprendizaje del modelo de árboles de decisión aplicadas a las bandas de frecuencia, utilizando todos los electrodos y datos sin balancear. En las tres bandas de frecuencia se observa un comportamiento similar: la curva azul, que representa la precisión del modelo en el conjunto de entrenamiento, disminuye a medida que aumenta el número de muestras, mientras que la curva naranja, que muestra la precisión del modelo en los datos de prueba, experimenta un incremento conforme se incrementan las muestras. Este patrón sugiere que el modelo alcanza un punto de equilibrio al aprender de los datos de entrenamiento y mejora gradualmente su capacidad predictiva en los datos de prueba.

Figura 4*Curvas de Aprendizaje de Árboles de Decisión para Datos sin Balanceo para Todos los Canales*

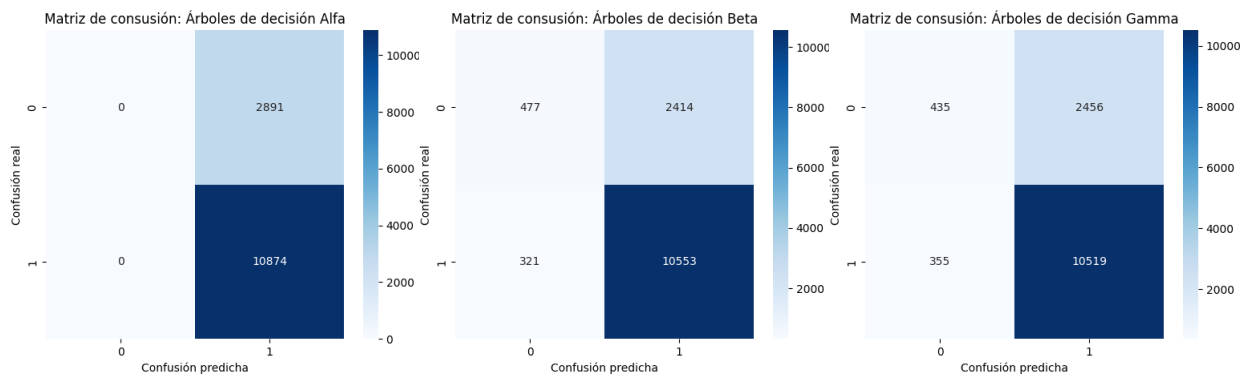
Nota. Curvas de aprendizaje del modelo árboles de decisión para todos los electrodos sin técnicas de balanceo aplicadas en las tres bandas de frecuencia (alfa, beta y gama).

A partir del modelo entrenado se calculó la matriz de confusión para cada banda de frecuencia. La figura 5 muestra que el modelo de árboles de decisión tiene una alta precisión al predecir los casos en los que el estudiante comprendió, pero presenta dificultades al clasificar correctamente los casos de no comprensión. En la figura correspondiente a la banda de frecuencia alfa, se observa que el modelo clasifica todas las muestras como positivas, asumiendo que el estudiante comprendió, lo que indica una falta de discriminación en los casos negativos. En contraste, en las bandas de beta y gama el modelo muestra una mejora en la capacidad de identificar correctamente los casos negativos, lo que sugiere un avance en la precisión para predecir cuando el estudiante no comprendió. Sin embargo, en ambos casos sigue existiendo una proporción considerable de falsos positivos, lo que significa que el modelo continúa clasificando erróneamente los casos de no comprensión como si el estudiante hubiera entendido. Además, se observa un número significativo de verdaderos positivos, lo que indica que el modelo es capaz de reconocer los casos de comprensión correctamente. Este comportamiento sugiere que el modelo tiene dificultades para distinguir con precisión los casos de no comprensión, especialmente en la

banda de frecuencia alfa, donde el modelo muestra una menor capacidad para diferenciar entre los estados de comprensión y no comprensión.

Figura 5

Matrices de Confusión de Árboles de Decisión para Datos sin Balanceo para Todos los Canales



Nota. Matrices de confusión del modelo árboles de decisión para todos los electrodos sin técnicas de balanceo aplicadas en las tres bandas de frecuencia (alfa, beta y gama).

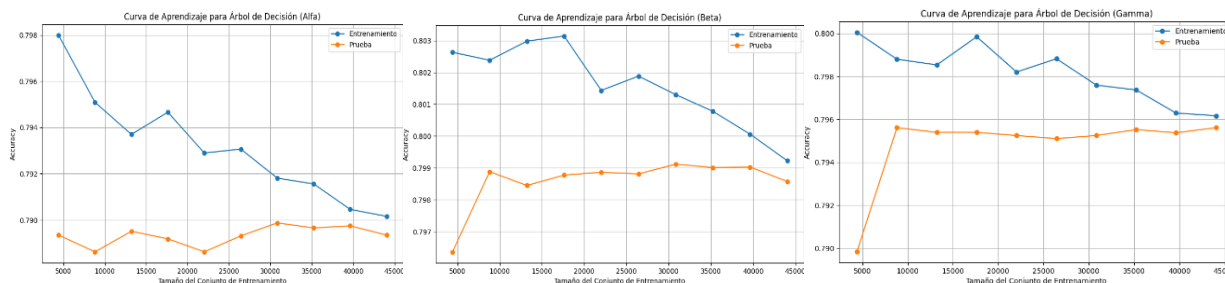
Clasificación para Electrodo Específicos. La figura 6 muestra las curvas de aprendizaje del modelo de árboles de decisión aplicado a los datos sin balancear, utilizando únicamente los electrodos T8, T7, P7, P8, O1 y O2. La línea azul representa la precisión del modelo en el conjunto de entrenamiento, donde se observa una reducción en la precisión a medida que aumenta el número de muestras. Por otro lado, la línea naranja ilustra el comportamiento del modelo en los datos de prueba, mostrando un aumento en la precisión conforme se incrementan las muestras. Esto sugiere que el modelo está aprendiendo de los datos de entrenamiento y ajustándose gradualmente a los datos de prueba. Sin embargo, la banda alfa no presenta una mejora continua en la curva de prueba, sino que oscila alrededor de un valor constante de precisión. Este comportamiento indica que el modelo tiene dificultades para generalizar correctamente, lo que podría ser indicativo de que los datos de prueba no aportan nueva información significativa para mejorar el rendimiento en

esa banda en particular.

Figura 6

Curva de Aprendizaje de Árboles de Decisión para Datos sin Balanceo para Canales

Específicos



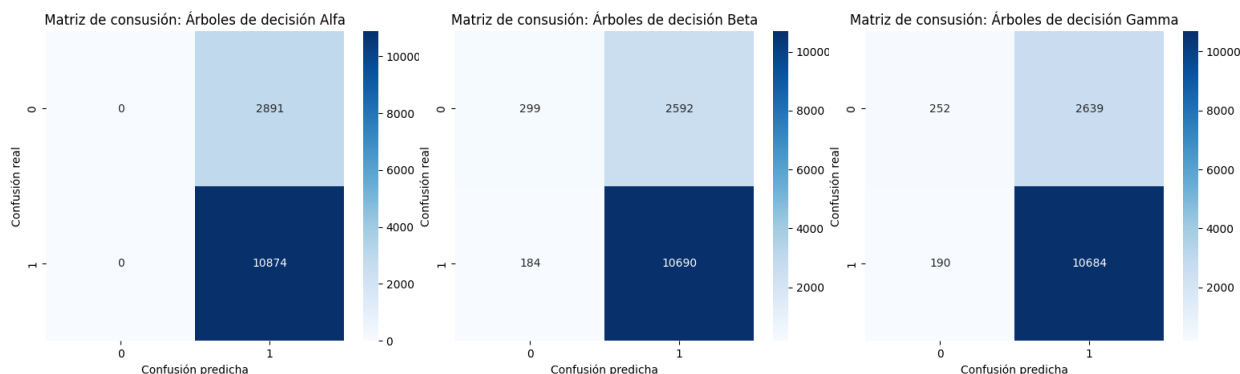
Nota. Curvas de aprendizaje del modelo árboles de decisión para electrodos específicos sin técnicas de balanceo aplicadas en las tres bandas de frecuencia (alfa, beta y gama).

La figura 7 presenta las matrices de confusión para cada banda de frecuencia. Al igual que en el análisis realizado con todos los electrodos, el modelo muestra un sesgo hacia las predicciones positivas, clasificando correctamente los casos en los que el estudiante comprendió, pero fallando al predecir los casos en los que no comprendió. Este comportamiento indica que la selección de solo los electrodos con mayor asociación estadística a la variable objetivo no parece tener un impacto significativo en los resultados obtenidos con el conjunto de datos sin balancear. Además, el modelo muestra un mejor rendimiento en las bandas de frecuencia beta y gama, donde se observa una mayor cantidad de verdaderos negativos. En contraste, en la banda alfa, el modelo tiende a predecir todas las muestras como positivas, lo que refleja una falta de discriminación en los casos en los que el estudiante no comprendió. Este análisis resalta la necesidad de ajustar el modelo para mejorar su capacidad de distinguir entre comprensión y no comprensión, especialmente en la banda de frecuencia alfa.

Figura 7

Matrices de Confusión de Árboles de Decisión para Datos sin Balanceo para Canales

Específicos



Nota. Matrices de confusión del modelo árboles de decisión para electrodos específicos sin técnicas de balanceo aplicadas en las tres bandas de frecuencia (alfa, beta y gama).

Datos con Undersampling

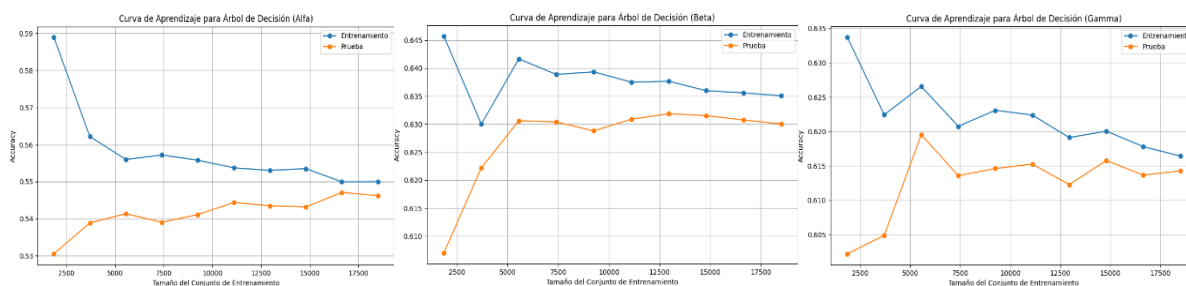
En la sección se presenta el análisis detallado de los resultados obtenidos del entrenamiento de modelos de árboles de decisión, aplicados a los datos con Undersampling. Este análisis se lleva a cabo para dos grupos de electrodos: uno compuesto por aquellos que muestran una mayor asociación estadística con la variable objetivo y otro que incluye todos los electrodos disponibles. A través de este enfoque, se busca evaluar cómo el balanceo de los datos mediante Undersampling y la selección de electrodos relevantes afectan el rendimiento del modelo, tanto en términos de precisión como en su capacidad para diferenciar entre estados de comprensión y no comprensión. La comparación entre los resultados obtenidos para estos dos conjuntos de electrodos permitirá identificar si la reducción de la dimensionalidad y el tratamiento de los desequilibrios en los datos contribuyen a una mejora en la capacidad predictiva del modelo.

Clasificación para Todos los Electrodos. Los resultados del entrenamiento se presentan

en las figuras 8 y 9. La curva de entrenamiento de la banda alfa disminuye hasta estabilizarse en 0,55, mientras que la de prueba oscila en torno a 0,525, evidenciando un infra ajuste. La figura de la banda beta muestra un mejor desempeño del modelo, donde ambas curvas convergen con el aumento de datos, aunque persisten oscilaciones en la curva de prueba por ruido o variabilidad. El modelo gama presenta un buen balance entre sesgo y varianza, con curvas que convergen de forma más estable, reflejando un rendimiento más generalizable.

Figura 8

Curvas de Aprendizaje de Árboles de Decisión para Datos con Undersampling para Todos los Canales



Nota. Curvas de aprendizaje del modelo árboles de decisión para todos los electrodos con técnica de Undersampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

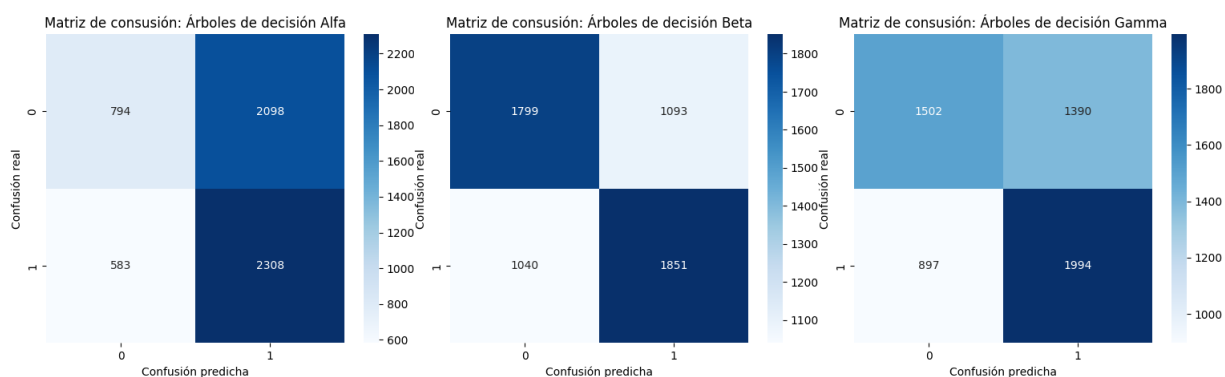
La matriz de confusión presentada en la figura 9 muestra una evolución en el desempeño del modelo al predecir verdaderos negativos. En la banda alfa, se observa un elevado número de falsos positivos en comparación con los verdaderos negativos, lo que sugiere que el modelo tiende a clasificar incorrectamente muchas instancias de la clase negativa como positivas. Aunque el modelo muestra un buen desempeño identificando correctamente los verdaderos positivos, su precisión general se ve afectada significativamente por estos errores en la clase negativa. En la banda beta, la matriz de confusión evidencia un mejor balance. El modelo logra clasificar

correctamente un número considerable de ejemplos tanto de la clase positiva como de la negativa, acompañado de una reducción notable en la cantidad de falsos positivos y falsos negativos en comparación con alfa. Esto indica un ajuste más equilibrado en la predicción de ambas clases.

Por su parte, la banda gama presenta un desempeño aún más robusto, con un número más equilibrado de verdaderos positivos y negativos en comparación con alfa, y una cantidad menor de falsos negativos que beta. Sin embargo, persiste un nivel significativo de falsos positivos, lo que sugiere que aún hay margen para mejorar la precisión en la clasificación de la clase negativa. En general, los resultados indican una mejora progresiva del modelo al predecir las muestras negativas tras la aplicación del Undersampling. Este proceso parece haber reducido la desproporción entre los errores de clasificación en las diferentes configuraciones evitando el sesgo del modelo.

Figura 9

Matrices de Confusión de Árboles de Decisión para Datos con Undersampling para Todos los Canales

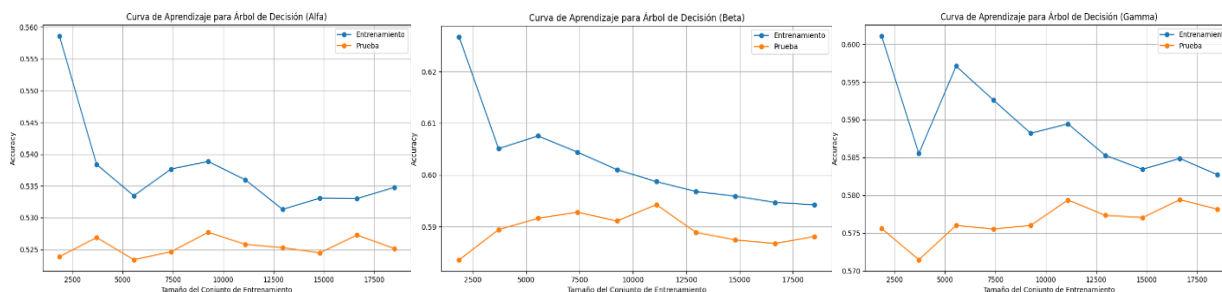


Nota. Matrices de confusión del modelo árboles de decisión para todos los electrodos con técnica de Undersampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Clasificación para Electrodo Específicos. En la figura 10 se muestran los resultados de los árboles de decisión aplicados con Undersampling. La curva de alfa ilustrada en la figura 10 revela que el modelo no aprende adecuadamente, con la curva de prueba estable en torno a 0.525 y la de entrenamiento disminuyendo con más muestras, indicando un alto sesgo y baja capacidad predictiva. En la banda beta, aunque hay un leve incremento en precisión, el modelo sigue sin capturar características clave. La banda gama muestra mejor balance entre sesgo y varianza, pero con una precisión máxima de 0.58, lo que refleja insuficiencia en el ajuste del modelo, posiblemente por la complejidad de los datos o falta de ajuste en los hiperparámetros.

Figura 10

Curvas de Aprendizaje de Árboles de Decisión para Datos con Undersampling para Canales Específicos



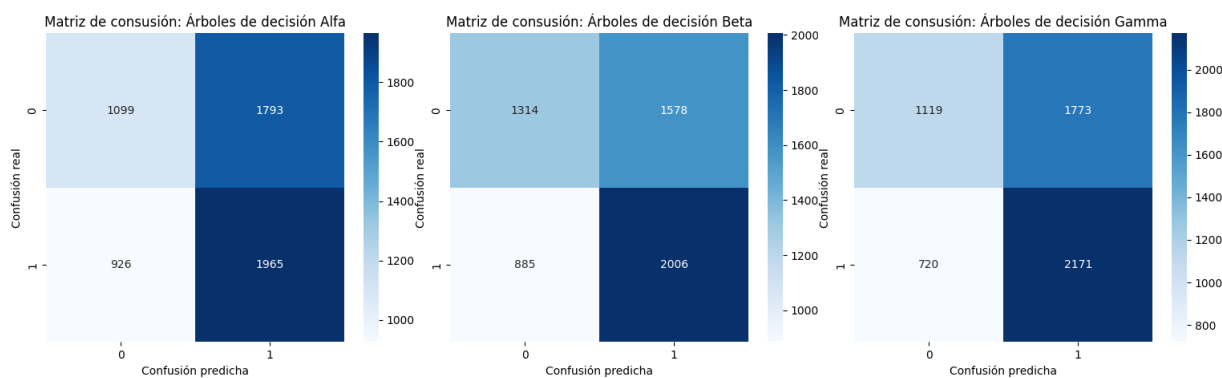
Nota. Curvas de aprendizaje del modelo árboles de decisión para electrodos específicos con técnica de Undersampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Las matrices de confusión que se muestran en la figura 11 evidencian que los árboles de decisión tienen dificultades para ajustar correctamente los datos en las bandas. En la banda alfa, el modelo presenta un sesgo significativo hacia la clase 1, con un alto número de falsos positivos, mientras que en la banda beta mejora ligeramente el balance entre clases, pero persisten errores importantes. La banda gama muestra el mejor rendimiento, con menor cantidad de falsos

negativos y un mejor equilibrio entre las predicciones correctas de ambas clases, aunque la precisión general sigue siendo baja. Estos resultados sugieren que el modelo es demasiado simple para capturar las características de los datos.

Figura 11

Matrices de Confusión de Árboles de Decisión para Datos con Undersampling para Canales Específicos



Nota. Matrices de decisión del modelo árboles de decisión para electrodos específicos con técnica de Undersampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Datos con Oversampling

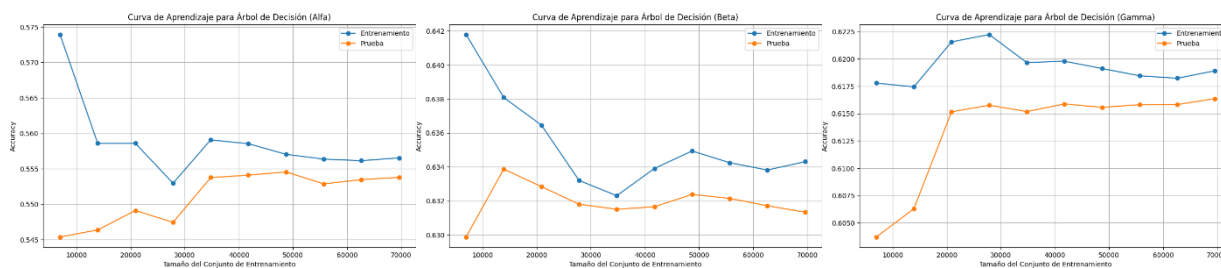
Clasificación para Todos los Electrodo. La figura 12 presenta las curvas de aprendizaje del modelo de árbol de decisión aplicado a los datos sobremuestreados en todos los canales. En alfa y beta, se identifica un claro subajuste: la curva de entrenamiento muestra una alta precisión inicial que disminuye a medida que aumenta el número de muestras, mientras que la curva de prueba permanece significativamente inferior y estable, sin mejoras con el incremento del conjunto de datos. Este comportamiento evidencia que el modelo no está capturando patrones relevantes ni logrando una representación adecuada de los datos, lo que limita su capacidad predictiva.

Por otro lado, la banda gama muestra un desempeño notablemente mejor, la curva de

entrenamiento incrementa de forma progresiva su precisión conforme crece el tamaño del conjunto de entrenamiento, hasta converger con la curva de entrenamiento. Este comportamiento sugiere un equilibrio adecuado entre sesgo y varianza, lo que indica que el modelo está logrando aprender patrones relevantes y generalizar de manera efectiva. Sin embargo, a pesar de esta mejora, el rendimiento global del modelo podría seguir viéndose afectado por la complejidad de los datos o la necesidad de ajustar mejor los hiperparámetros para maximizar la capacidad predictiva y reducir errores residuales.

Figura 12

Curvas de Aprendizaje de Árboles de Decisión para Datos con Oversampling para Todos los Canales



Nota. Curvas de aprendizaje del modelo árboles de decisión para todos los electrodos con técnica de Oversampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Las matrices de confusión de la figura 13 revelan diferencias significativas en el desempeño del modelo para las bandas alfa, beta y gama. En el caso de alfa, los errores se concentran principalmente en los falsos negativos, lo que indica que el modelo tiene dificultades para identificar correctamente ciertas clases. Este comportamiento puede atribuirse al subajuste, donde el modelo es demasiado simple para capturar las relaciones complejas entre las características de los datos, lo que afecta negativamente su capacidad predictiva.

En las bandas beta y gama, se observa una mejora en la clasificación. Hay un aumento en

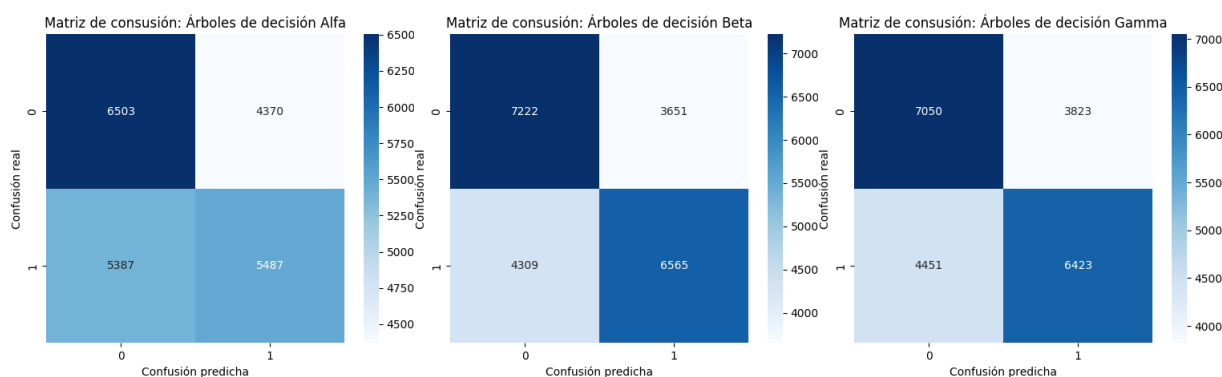
los verdaderos positivos y negativos, acompañado de una reducción en los falsos positivos y negativos. Esto refleja un desempeño más equilibrado en la clasificación de las clases, aunque el análisis de las curvas de aprendizaje sugiere que el modelo aún no alcanza un ajuste óptimo.

En beta, aunque la reducción de errores es significativa, persiste un margen de mejora en la capacidad de generalización.

Gama, por su parte, muestra el mejor desempeño entre las tres bandas. La disminución de falsos positivos y negativos es más pronunciada, lo que indica una mejora en la precisión y la sensibilidad del modelo. Sin embargo, el análisis general sugiere que el modelo aún podría beneficiarse de ajustes adicionales en los hiperparámetros o de un enfoque más avanzado en el preprocesamiento de los datos para mitigar posibles problemas como el ruido o la distribución desequilibrada de las clases.

Figura 13

Matrices de Confusión de Árboles de Decisión para Datos con Oversampling para Todos los Canales

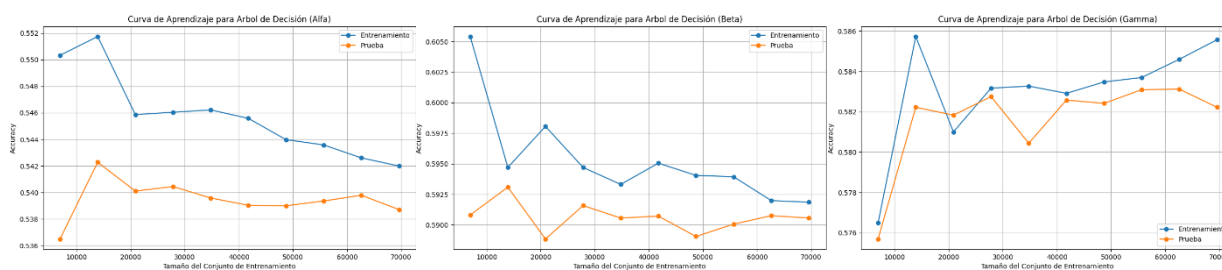


Nota. Matrices de confusión del modelo árboles de decisión para todos los electrodos con técnica de Oversampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Clasificación para Electrodos Específicos. En la imagen que presenta la figura 14 se aprecia que existe un subajuste en alfa y en beta, pues presenta una disminución progresiva de la precisión en entrenamiento y un bajo rendimiento en prueba sin mejoras significativas al incrementar los datos. gama logra un mejor balance, con curvas que convergen a medida que crece el conjunto de entrenamiento, indicando una adecuada capacidad de generalización.

Figura 14

Curvas de Aprendizaje de Árboles de Decisión para Datos con Oversampling para Canales Específicos



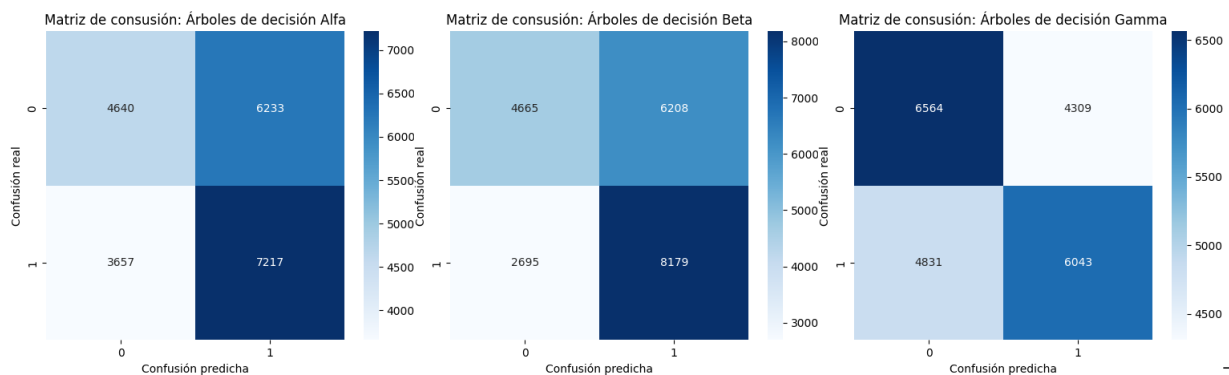
Nota. Curvas de aprendizaje del modelo árboles de decisión para electrodos específicos con técnica de Oversampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

La matriz de confusión para cada banda de frecuencia, presentada en la Figura 15, revela diferencias significativas en el desempeño del modelo para predecir la comprensión de los estudiantes. En la banda alfa, el modelo demuestra ser moderadamente efectivo para identificar estudiantes que comprendieron el contenido, con un número considerable de verdaderos positivos. No obstante, el elevado número de falsos positivos indica que el modelo tiende a clasificar erróneamente como comprendido a estudiantes que realmente estaban confundidos, lo que podría deberse a características poco discriminantes en esta banda de frecuencia. En la banda beta, se observa un desempeño superior en términos de sensibilidad, logrando un mayor número de verdaderos positivos y reduciendo los falsos negativos en comparación con la banda alfa. Esto

sugiere que las señales capturadas en esta banda están más correlacionadas con la comprensión del contenido. Sin embargo, el modelo mantiene un alto número de falsos positivos, lo que refleja dificultad para distinguir entre estudiantes que estaban confundidos y los que no. Finalmente, en la banda gama, el modelo presenta el mejor desempeño en términos de especificidad, al reducir significativamente los falsos positivos y aumentar los verdaderos negativos. Esto indica que es más preciso al identificar a los estudiantes que no comprendieron el contenido. No obstante, este logro se ve contrarrestado por un número elevado de falsos negativos, lo que implica que el modelo no logra captar correctamente a una proporción considerable de estudiantes que sí comprendieron.

Figura 15

Matrices de Confusión de Árboles de Decisión para Datos con Oversampling para Canales Específicos



Nota. Matrices de confusión del modelo árboles de decisión para electrodos específicos con técnica de Undersampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

KNN

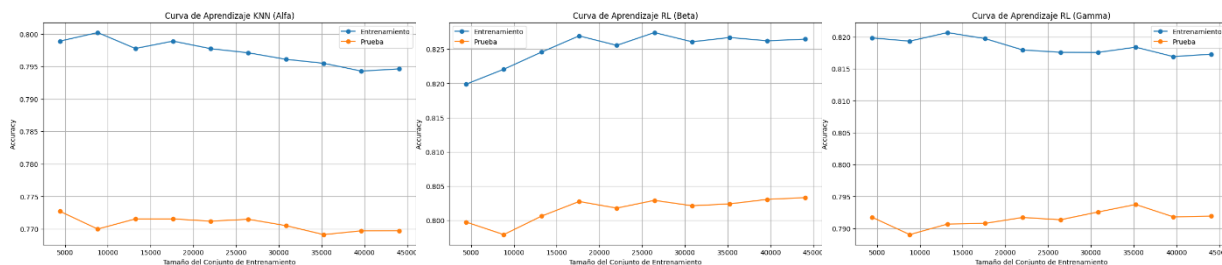
El modelo de K-Nearest Neighbors (KNN) se implementó utilizando la función KNeighbors Classifier de la biblioteca scikit-learn. Para todas las configuraciones de datos (sin

balanceo, con Undersampling y con Oversampling), se estableció el parámetro de $k=10$, lo que implica que el modelo considera los 10 vecinos más cercanos en el espacio de características para realizar la clasificación. El objetivo de esta configuración fue evaluar el impacto del balanceo de datos en el desempeño del modelo KNN en la predicción de la variable objetivo. Las pruebas se llevaron a cabo de manera independiente para cada banda de frecuencia, permitiendo así identificar cómo contribuyen las características específicas de cada banda al desempeño global del modelo. El objetivo de esta sección es determinar cómo las técnicas para abordar el desequilibrio de clases afectan la capacidad del modelo para clasificar correctamente los estados de confusión y comprensión. Además, se analiza si la reducción de dimensionalidad, al restringir el análisis a los electrodos más relevantes, mejora la precisión general y la robustez del modelo. La comparación de los resultados entre ambos conjuntos de electrodos permite explorar si la inclusión de información menos relevante afecta negativamente el rendimiento del modelo, o si, por el contrario, un enfoque más específico contribuye a una clasificación más efectiva y eficiente.

Datos sin Balanceo

Clasificación para Todos los Electrodo. La Figura 16 muestra las curvas de aprendizaje del modelo KNN en datos sin balancear. Se observa una brecha significativa entre las curvas de entrenamiento y prueba, indicando sobreajuste y dificultades para generalizar a nuevos datos. La precisión en el entrenamiento permanece alta y constante, mientras que en la prueba es más baja, reforzando el diagnóstico de sobreajuste.

Entre las bandas, beta y gama muestran un mejor equilibrio entre las curvas, sugiriendo mayor capacidad de generalización. En contraste, la banda alfa presenta el peor desempeño, con baja precisión en ambos conjuntos, lo que indica menor discriminación en las características de esta banda.

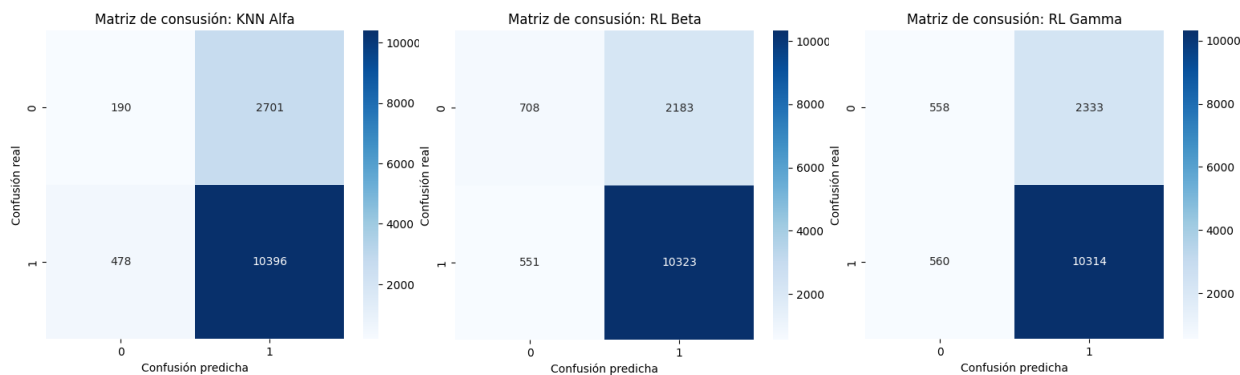
Figura 16*Curvas de Aprendizaje de KNN para Datos sin Balanceo para Todos los Canales*

Nota. Curvas de aprendizaje del modelo KNN para todos los electrodos con técnica de Undersampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

La figura 17 presenta las matrices de confusión correspondientes al modelo KNN aplicado a los datos sin balancear para todos los canales de EEG. Se observa que la banda beta destaca por lograr el mayor número de verdaderos positivos, indicando que es la más efectiva para detectar estudiantes que comprendieron el contenido, aunque presenta un alto número de falsos positivos, lo que compromete su precisión global. La banda gama presenta un comportamiento similar, con un alto número de verdaderos positivos, pero con menos verdaderos negativos, lo que la hace menos confiable para distinguir entre las dos clases. La banda alfa, aunque identifica correctamente a una gran cantidad de estudiantes que comprendieron, tiene un desempeño general más limitado debido a su elevada tasa de falsos positivos, lo que sugiere una menor capacidad discriminativa.

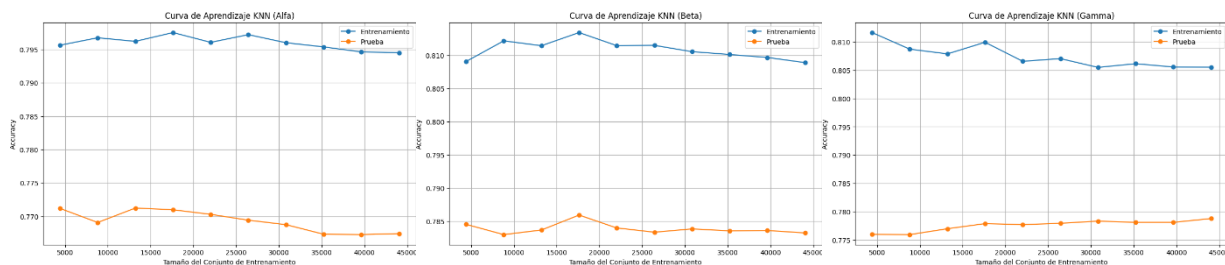
Figura 17

Matrices de Confusión de KNN para Datos sin Balanceo para Todos los Canales



Nota. Matrices de confusión del modelo KNN para todos los electrodos con técnica de Undersampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Clasificación para Electroodos Específicos. Las curvas de aprendizaje para el modelo KNN aplicadas a los electrodos específicos en los datos con sobre muestreo se presentan en la figura 18. La banda beta destaca por lograr el mejor equilibrio entre precisión en el conjunto de entrenamiento y el conjunto de prueba, lo que indica que las características de esta banda tienen una mayor capacidad de generalización en comparación con alfa y gama. Por otro lado, la banda alfa presenta un desempeño más limitado, con una brecha significativa entre la precisión de entrenamiento y prueba, lo que sugiere un mayor nivel de sobreajuste y una capacidad discriminativa moderada. La banda gama, aunque alcanza la mayor precisión en entrenamiento, exhibe una brecha considerable con la precisión de prueba, lo que refleja un sobreajuste severo y una menor estabilidad en la generalización.

Figura 18*Curvas de Aprendizaje de KNN para Datos sin Balanceo para Canales Específicos*

Nota. Curvas de aprendizaje del modelo KNN para electrodos específicos con técnica de Undersampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

En la figura 19, se observa un desempeño variado del modelo KNN en las distintas bandas de frecuencia. La banda beta destaca por ofrecer el mejor equilibrio entre sensibilidad y especificidad, con 517 verdaderos negativos y 10,287 verdaderos positivos. Sin embargo, aunque su rendimiento es superior en términos generales, aún presenta un número moderado de falsos negativos y un alto número de falsos positivos, lo que implica cierta dificultad en la correcta clasificación de la clase negativa.

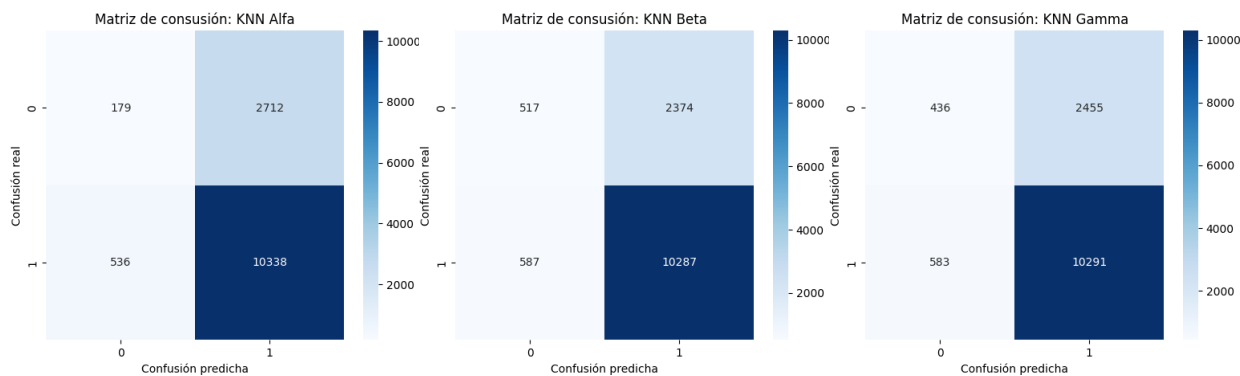
La banda gama, por su parte, muestra una reducción en los falsos positivos respecto a alfa, lo que mejora la precisión en la identificación de la clase negativa. Sin embargo, esto viene acompañado de un incremento en los falsos negativos, indicando que el modelo tiene menor sensibilidad y podría no detectar todos los casos positivos. Este comportamiento sugiere que el modelo en gama está más optimizado para evitar falsas alarmas en la clase negativa, aunque a costa de sacrificar algunos casos verdaderamente positivos.

En contraste, la banda alfa tiene el peor desempeño relativo. Presenta el mayor número de falsos positivos y la menor cantidad de verdaderos negativos, lo que limita su capacidad para identificar correctamente la clase negativa. A pesar de esto, logra un buen número de verdaderos

positivos, lo que indica que tiene cierta efectividad en la detección de la clase positiva, aunque con un costo significativo en términos de precisión global.

Figura 19

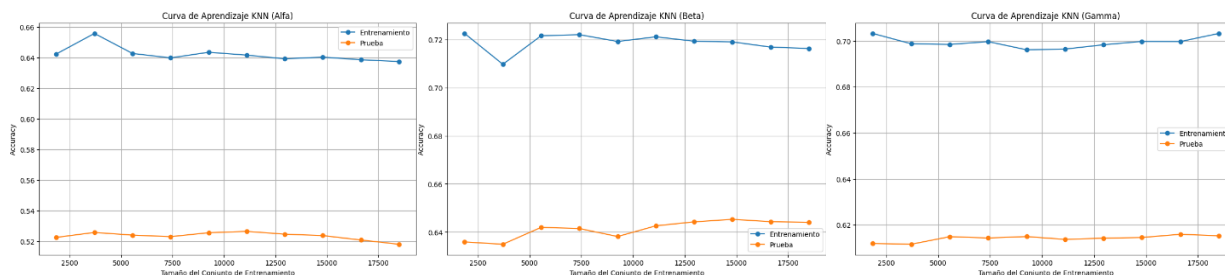
Matrices de Confusión de KNN para Datos sin Balanceo para Canales Específicos



Nota. Matrices de confusión del modelo KNN para electrodos específicos sin técnicas de balanceo aplicadas en las tres bandas de frecuencia (alfa, beta y gama).

Datos con Undersampling

Clasificación para Todos los Electrodo. Las curvas de aprendizaje del modelo KNN aplicadas a los datos con Undersampling, que se presentan en la figura 22, revelan que en la banda beta tiene el mejor equilibrio entre las curvas de entrenamiento y prueba, mostrando una menor brecha en comparación con las demás bandas. Por otro lado, la banda alfa presenta un desempeño moderado, aunque con una brecha más amplia, lo que sugiere un mayor sobreajuste. La banda gama muestra una alta precisión en el entrenamiento, pero la precisión en el conjunto de prueba es baja, indicando un sobreajuste y menor capacidad de generalización. En conjunto, los resultados destacan que el modelo KNN que busca maximizar la precisión en datos balanceados por Undersampling presenta un sobreajuste con los datos de entrenamiento y tiene un desempeño más bajo en los datos de entrenamiento.

Figura 20*Curvas de Aprendizaje de KNN para Datos con Undersampling para Todos los Canales*

Nota. Curvas de aprendizaje del modelo KNN para todos los electrodos con técnica de Undersampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

En la figura 23, se evidencia un desempeño diferenciado del modelo KNN en las distintas bandas de frecuencia al clasificar los datos. La banda beta destaca como la más equilibrada, logrando 1,681 verdaderos negativos y 1,995 verdaderos positivos, lo que indica una sólida capacidad para identificar tanto casos positivos como negativos. Además, su número de falsos negativos es relativamente bajo en comparación con las demás bandas, lo que refuerza su efectividad al minimizar los errores en la detección de estudiantes que no comprendieron el contenido. Este balance sugiere que beta captura patrones significativos en los datos, ofreciendo un modelo más confiable.

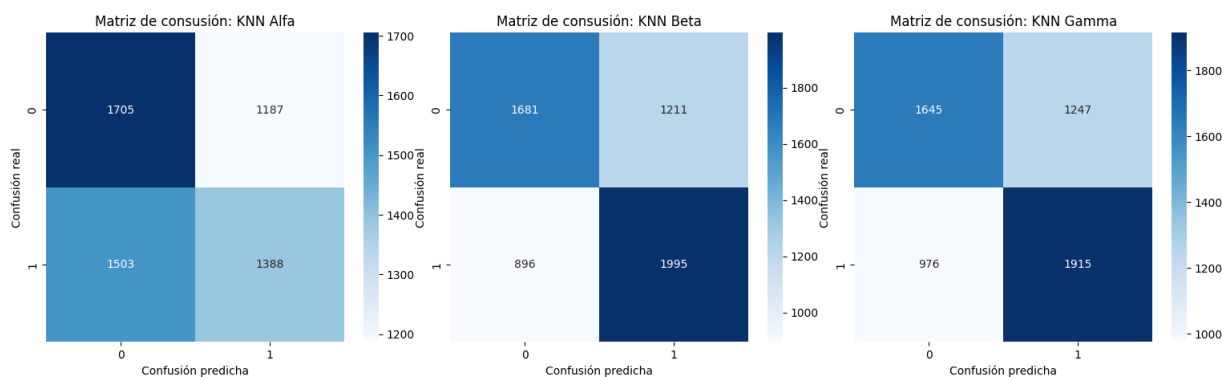
En contraste, la banda gama muestra un comportamiento mixto. Aunque reduce los falsos negativos en comparación con alfa, lo que implica una mejor detección de estudiantes que comprendieron el contenido, presenta un mayor número de falsos positivos. Esto sugiere que el modelo clasifica incorrectamente más casos como positivos, posiblemente debido a una mayor sensibilidad que compromete la especificidad. A pesar de este incremento en los falsos positivos, gama muestra una mejora relativa en la sensibilidad, lo que podría ser útil en contextos donde es más importante identificar todos los casos positivos, aunque se permita un margen de error en la

clase negativa.

Por otro lado, la banda alfa exhibe el desempeño más limitado entre las tres bandas. Registra el mayor número de falsos negativos y falsos positivos, lo que refleja dificultades tanto para detectar correctamente los casos de comprensión como para identificar de manera precisa los casos de confusión. Este desempeño sugiere que las características extraídas de alfa no son suficientemente discriminativas, lo que afecta significativamente la precisión y sensibilidad del modelo en esta banda.

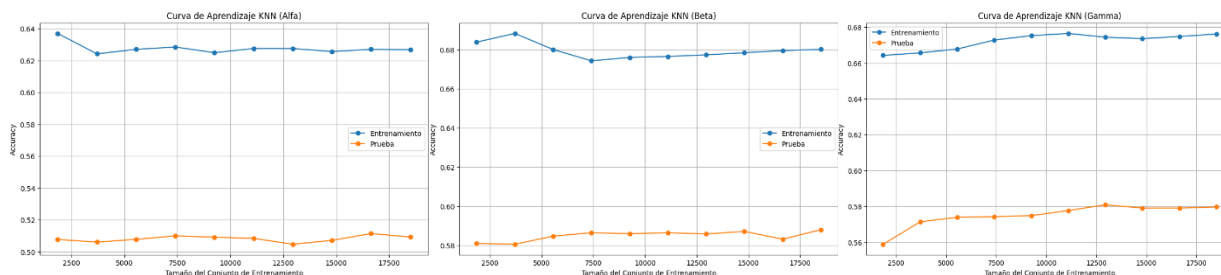
Figura 21

Matrices de Confusión de KNN para Datos con Undersampling para Canales Específicos



Nota. Matrices de confusión del modelo KNN para electrodos específicos con técnica de Undersampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Clasificación para Electrodo Específicos. Los resultados que presenta la figura 22 presentan un patrón consistente en las tres bandas, donde existe un sobre ajuste a los datos de entrenamiento en los que se presenta una mejor precisión y permanece estable con el aumento del tamaño del conjunto de entrenamiento. Sin embargo, la precisión en los conjuntos de prueba es significativamente más baja. Esto sugiere que el modelo no generaliza bien para los datos de prueba.

Figura 22*Curvas de Aprendizaje de KNN para Datos con Undersampling para Canales Específicos*

Nota. Curvas de aprendizaje del modelo KNN para electrodos específicos con técnica de Undersampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Las matrices de confusión para los canales específicos tras el balanceo por Undersampling, presentadas en la figura 23, muestran un desempeño moderado del modelo, con valores predichos correctamente en ambas clases, pero un número significativo de errores tanto en las clases 0 como 1. Esto evidencia las limitaciones del modelo para distinguir con precisión entre las dos clases.

En la banda alfa, el modelo clasifica incorrectamente 1,716 casos de la clase 1 y 1,131 de la clase 0, reflejando dificultades importantes en la detección tanto de estudiantes confundidos como de aquellos que comprendieron el contenido. Aunque el rendimiento es bajo, alfa proporciona una base de comparación útil frente a las otras bandas.

La banda beta muestra una ligera mejora en la clasificación de la clase 1, con 1,309 errores, lo que indica un aumento en la sensibilidad del modelo. Sin embargo, persisten 1,121 errores en la clase 0, lo que señala que, aunque el modelo ha mejorado en la identificación de estudiantes que comprendieron el contenido, aún enfrenta retos significativos al clasificar correctamente a los estudiantes confundidos.

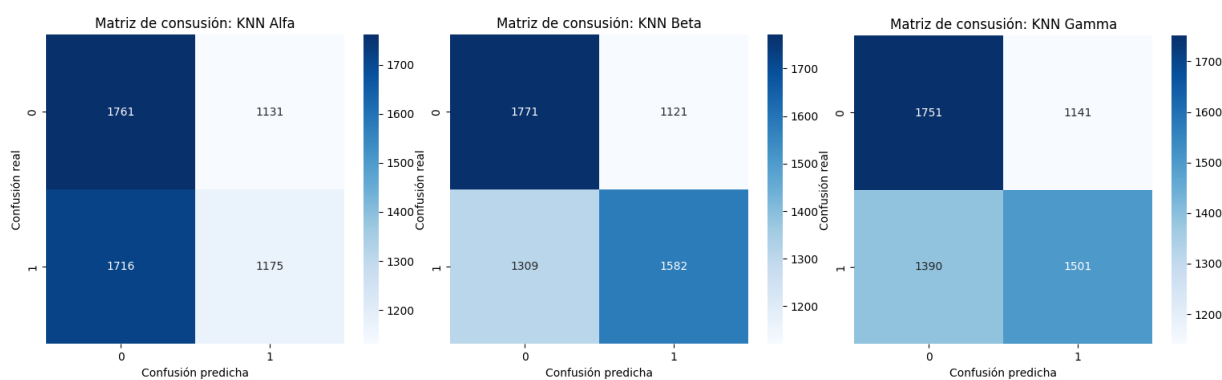
Finalmente, la banda gama presenta un comportamiento intermedio. Con 1,390 errores en

la clase 1 y 1,141 en la clase 0, el modelo muestra un equilibrio relativo, pero no alcanza una precisión óptima. Si bien la mejora en la clase 1 no es tan marcada como en beta, gama logra un desempeño más balanceado en ambas clases.

A pesar de estas limitaciones, las matrices reflejan una mejora en la predicción de la clase 0, es decir, en la capacidad del modelo para identificar estudiantes confundidos, en comparación con modelos evaluados previamente sin balanceo.

Figura 23

Matrices de Confusión de KNN para Datos con Undersampling para Canales Específicos



Nota. Matrices de confusión del modelo KNN para electrodos específicos con técnica de Undersampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

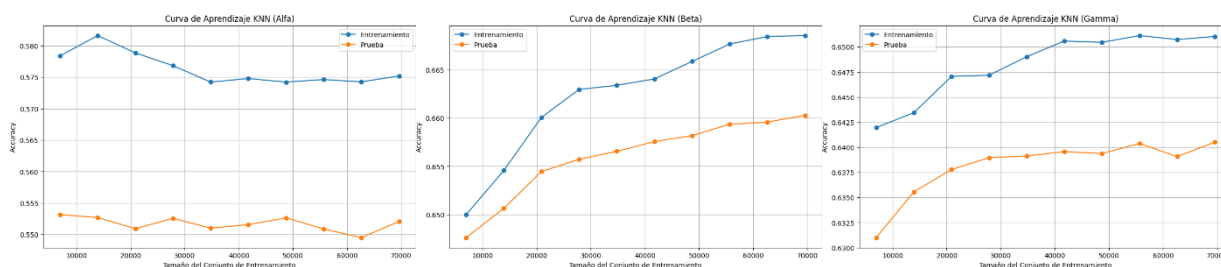
Datos con Oversampling

Clasificación para Todos los Electrodo. Las curvas de aprendizaje de la figura 24 muestra que para la banda alfa la precisión en el entrenamiento se mantiene relativamente constante, mientras que la precisión en prueba es baja, indicando un desempeño deficiente en generalización. En beta, se observa una mejora notable en la precisión tanto en el entrenamiento como en la prueba a medida que aumenta el tamaño del conjunto de datos, mostrando una tendencia hacia una mejor generalización. Para gama, se aprecia un comportamiento similar al

de beta, con incrementos progresivos en la precisión del conjunto de prueba, aunque se mantiene una brecha considerable entre ambas curvas, indicando una mayor varianza. Los resultados sugieren una mejora al usar el método de Oversampling en la capacidad de generalización en las bandas beta y gama, pero en alfa el modelo sigue presentando dificultades significativas.

Figura 24

Curvas de Aprendizaje de KNN para Datos con Sobremuestra para Todos los Canales



Nota. Curvas de aprendizaje del modelo KNN para todos los electrodos con técnica de Oversampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Los resultados de las matrices de confusión de la figura 25 reflejan una clara correspondencia con los patrones observados en las curvas de aprendizaje. Para la banda alfa, el modelo muestra un desempeño limitado, con 5,337 falsos negativos y 4,462 falsos positivos. Estos valores indican una notable dificultad para diferenciar correctamente entre las clases, lo que sugiere que las características extraídas de la banda alfa no proporcionan información suficientemente discriminativa para una clasificación precisa. Este alto número de errores en ambas clases evidencia una combinación de baja sensibilidad y especificidad, lo que afecta significativamente la utilidad del modelo en escenarios prácticos.

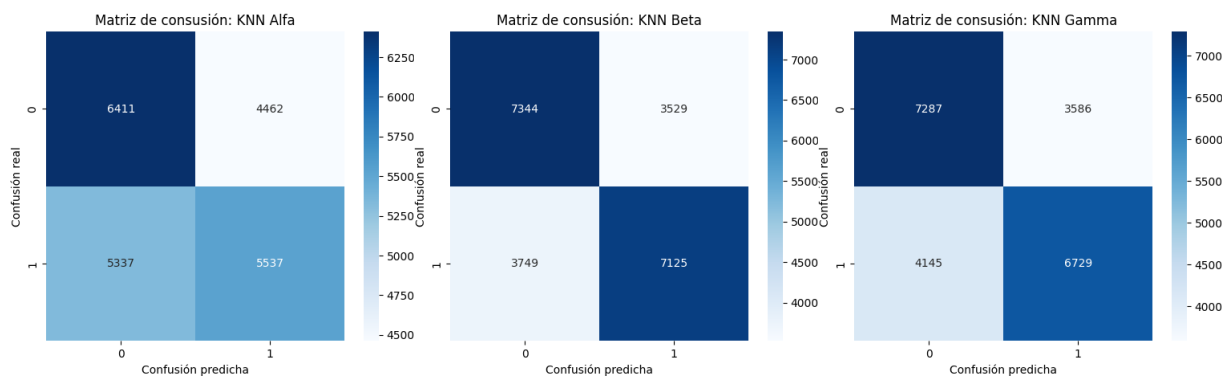
Por otro lado, las bandas beta y gama presentan un desempeño considerablemente mejor. En la banda beta, se observa una reducción notable en los errores, con una disminución

significativa tanto en los falsos negativos como en los falsos positivos. Esto refleja una mejora en el balance entre sensibilidad y especificidad, lo que permite al modelo identificar de manera más efectiva tanto los casos positivos como los negativos. Este desempeño más robusto sugiere que las características de la banda beta son más representativas y relevantes para la tarea de clasificación.

La banda gama también muestra un avance similar, aunque con un patrón ligeramente diferente. Si bien el número de errores es menor en comparación con alfa, persiste un leve desequilibrio en la clasificación, con una tendencia a clasificar incorrectamente más casos de una de las clases. A pesar de esto, gama logra un balance razonable, posicionándose como una opción intermedia que combina aspectos positivos de las otras bandas.

Figura 25

Matrices de Confusión de KNN para Datos con Sobremuestro para Todos los Canales



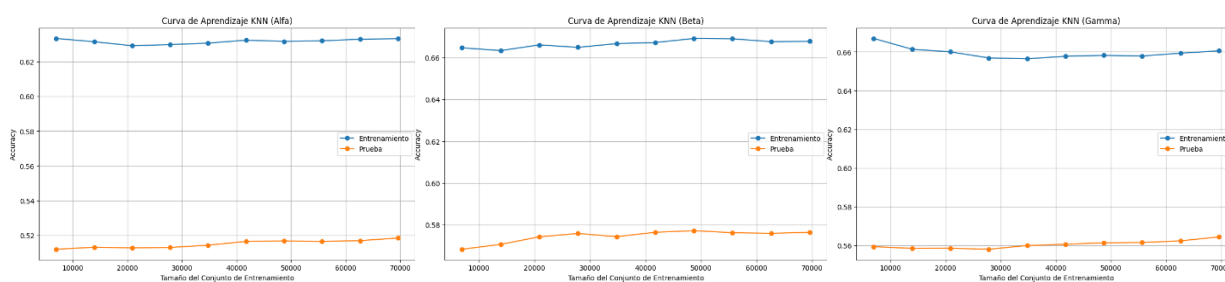
Nota. Matrices de confusión del modelo KNN para todos los electrodos con técnica de Oversampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Clasificación para Electrodo Específicos. Las curvas de aprendizaje del modelo KNN con datos balanceados mediante oversampling para todos los canales muestran un patrón consistente en las tres bandas, la precisión en el conjunto de entrenamiento es significativamente mayor que en el conjunto de prueba, y ambas curvas permanecen relativamente estables a medida

que aumenta el tamaño del conjunto de datos con una brecha entre ellas. En alfa, la precisión en el conjunto de prueba es extremadamente baja, cercana al azar, lo que indica una inadecuada generalización. En beta, aunque la precisión en el conjunto de prueba es algo mayor, sigue siendo insuficiente en comparación con la precisión en el entrenamiento. Finalmente, en gama, la brecha entre las curvas de entrenamiento y prueba persiste, aunque con ligeras variaciones.

Figura 26

Curvas de Aprendizaje de KNN para Datos con Oversampling para Canales Específicos



Nota. Curvas de aprendizaje del modelo KNN para electrodos específicos con técnica de Oversampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Las matrices de confusión para el modelo KNN con datos balanceados mediante Oversampling, presentadas en la figura 27, revelan un desempeño similar en las tres bandas de frecuencia, con una proporción significativa de errores en ambas clases. A pesar del uso de Oversampling para equilibrar las clases, el modelo sigue enfrentando desafíos en la clasificación precisa, reflejando limitaciones inherentes a su capacidad de generalización.

En la banda alfa, el modelo registra 6,557 falsos negativos y 3,958 falsos positivos. Estos valores destacan dificultades importantes en la predicción tanto de la clase 0 (estudiantes confundidos) como de la clase 1 (estudiantes no confundidos). La alta tasa de falsos negativos indica que el modelo tiene problemas para identificar correctamente a los estudiantes confundidos, mientras que los falsos positivos reflejan errores al clasificar incorrectamente a estudiantes no

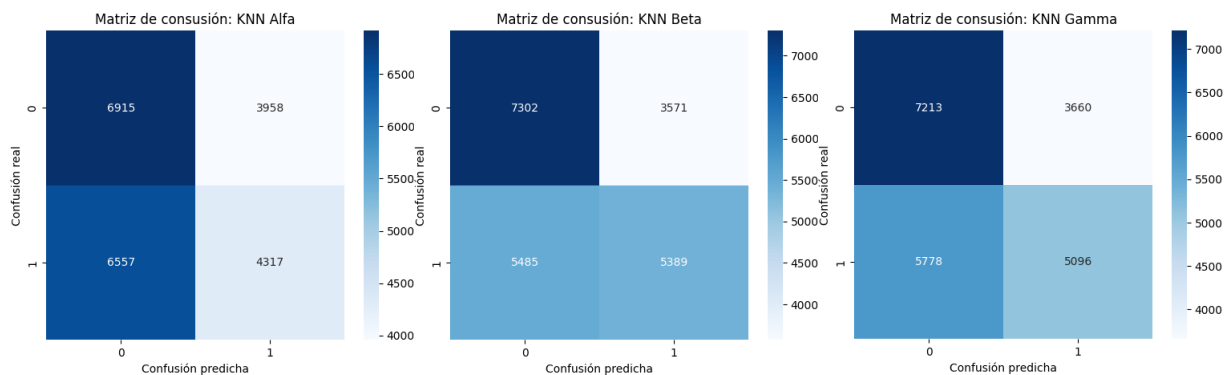
confundidos como confundidos. Este desempeño limitado sugiere que las características de la banda alfa no son suficientemente discriminativas, lo que compromete la precisión del modelo. En la banda beta, se observa una mejora ligera en el desempeño, con 5,485 falsos negativos y 3,571 falsos positivos. Esto indica un mayor balance entre sensibilidad y especificidad en comparación con alfa. La reducción en los falsos negativos refleja una mejora en la capacidad del modelo para identificar correctamente a los estudiantes confundidos, mientras que la disminución de falsos positivos sugiere un progreso en la correcta clasificación de los estudiantes no confundidos. Estas mejoras sugieren que las características extraídas de la banda beta son más representativas y útiles para la tarea de clasificación.

En la banda gama, el modelo muestra 5,778 falsos negativos y 3,660 falsos positivos. Aunque estos valores son intermedios entre alfa y beta, reflejan un comportamiento consistente con las otras bandas. La proporción de errores sigue siendo significativa, indicando que, aunque hay cierta mejora en la identificación de ambas clases, el modelo aún no logra una precisión óptima. Sin embargo, la banda gama destaca por mantener un equilibrio razonable entre la sensibilidad y la especificidad, lo que la posiciona como una opción viable en comparación con alfa.

Estos resultados resaltan que, aunque el balanceo mediante Oversampling mejora la detección de estudiantes confundidos al reducir parcialmente el sesgo hacia una de las clases, el modelo sigue enfrentando limitaciones. El desempeño general sugiere que KNN, en su configuración actual, no está capturando completamente los patrones subyacentes en los datos balanceados, lo que podría deberse a la complejidad inherente de los datos o a la falta de ajustes en los hiperparámetros del modelo

Figura 27

Matrices de Confusión de KNN para Datos con Sobremuestro para Canales Específicos



Nota. Matrices de confusión del modelo KNN para electrodos específicos con técnica de Oversampling aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Regresión Logística

En la sección 2.3 se presentan y analizan los resultados obtenidos al implementar la función LogisticRegression de la librería scikit-learn para entrenar los modelos de regresión logística. Los modelos fueron configurados con los parámetros predeterminados, asegurando una implementación estándar y consistente. Estos modelos se entrenaron y evaluaron en tres configuraciones de datos: sin balanceo, con Undersampling (Undersampling) y con Oversampling (Oversampling), tanto para los electrodos significativos como para todos los canales. El objetivo de esta sección es identificar cuál de estos enfoques logra un mejor desempeño en términos de precisión y capacidad de generalización del modelo. A través de este análisis comparativo, se busca evaluar cómo el balanceo de los datos afecta el rendimiento del modelo y determinar si existe una estrategia óptima para mejorar la clasificación en las distintas configuraciones evaluadas.

Datos sin Balanceo

Las curvas de aprendizaje de la regresión logística para todos los canales, con datos sin balanceo (figura 28), evidencian un comportamiento consistente en todas las bandas de frecuencia, caracterizado por una brecha significativa entre las precisiones del conjunto de entrenamiento y prueba. Este patrón revela problemas fundamentales en la capacidad del modelo para generalizar y adaptarse a la complejidad de los datos.

En todos los canales, la precisión en el conjunto de entrenamiento comienza alta, lo que indica que el modelo inicialmente se ajusta bien a los datos disponibles. Sin embargo, esta precisión decrece de forma irregular a medida que aumenta el tamaño del conjunto de datos. Este fenómeno sugiere que el modelo enfrenta dificultades para manejar un volumen mayor de datos, lo que podría deberse a una falta de complejidad en su estructura o a su incapacidad para capturar relaciones subyacentes en los datos de manera efectiva.

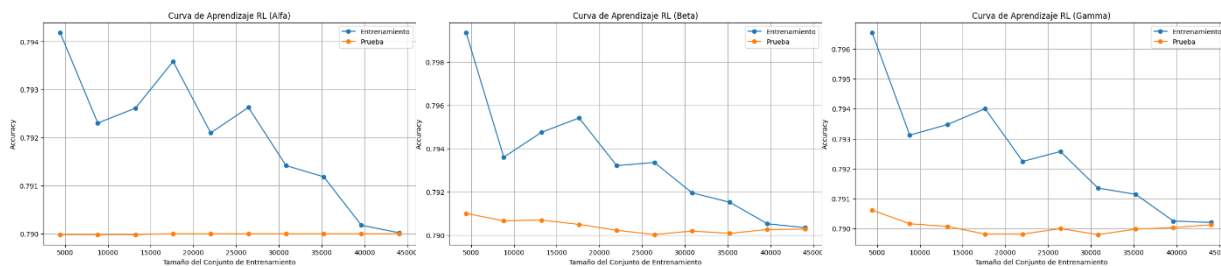
En el conjunto de prueba, la precisión permanece baja y prácticamente constante, con valores cercanos al azar (aproximadamente 50 % en una tarea binaria). Este resultado evidencia que el modelo no está logrando aprender patrones significativos que permitan una generalización adecuada. La brecha persistente entre las precisiones de entrenamiento y prueba refuerza la hipótesis de que el modelo está sobreajustado a las clases mayoritarias y no puede manejar el desequilibrio en la distribución de clases.

El desbalance de clases juega un papel crucial en estos resultados. La regresión logística, al ser un modelo lineal, puede ser particularmente vulnerable a este problema, ya que tiende a priorizar la clase mayoritaria, dejando a la clase minoritaria con un desempeño predictivo deficiente. Esto no solo afecta la sensibilidad del modelo (capacidad para identificar la clase minoritaria), sino también su precisión global y su capacidad para adaptarse a nuevos datos.

Clasificación para Todos los Electrodo

Figura 28

Curvas de Aprendizaje de Regresión Logística para Datos sin Balancear para Todos los Canales



Nota. Curvas de aprendizaje del modelo regresión logística para todos los electrodos sin técnicas de balanceo aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Las matrices de confusión de la regresión logística para los datos sin balanceo reflejan un desempeño limitado, dominado por un sesgo hacia la clase mayoritaria (clase 1), lo que pone en evidencia la incapacidad del modelo para manejar el desbalance de clases. Este comportamiento se observa de manera consistente en los diferentes canales.

En la banda alfa, el modelo clasifica correctamente 10,874 instancias de la clase 1, lo que indica un alto nivel de precisión en la predicción de la clase mayoritaria. Sin embargo, no logra identificar ninguna instancia de la clase 0, lo que evidencia una completa falta de sensibilidad hacia la clase minoritaria. Este resultado refleja un modelo altamente sesgado, que prioriza la precisión en la clase mayoritaria a costa de ignorar completamente la minoritaria.

El patrón observado en alfa se repite en la banda beta, aunque con una ligera mejora. El modelo clasifica correctamente 10,791 instancias de la clase 1, mientras que identifica apenas 84 instancias de la clase 0 de forma correcta. Aunque la detección de la clase 0 es marginalmente mejor que en alfa, sigue siendo insuficiente para proporcionar un equilibrio entre las clases, lo que

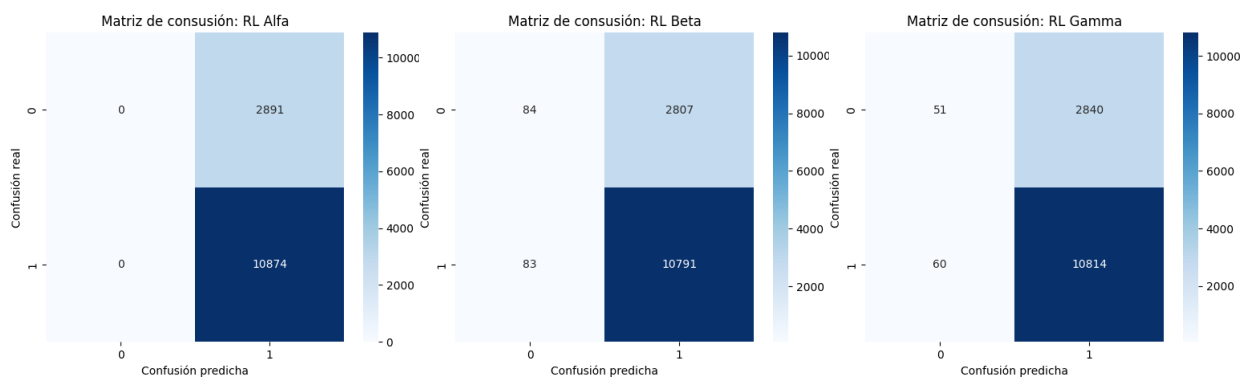
limita la utilidad del modelo en escenarios prácticos.

En la banda gama, el modelo muestra un comportamiento similar al de los otros canales. Clasifica correctamente 10,814 instancias de la clase 1, pero solo 51 instancias de la clase 0 son identificadas correctamente. Esto confirma la tendencia del modelo a enfocarse casi exclusivamente en la clase mayoritaria, dejando a la clase minoritaria prácticamente desatendida. Estas matrices evidencian cómo el desbalance de clases impacta negativamente el desempeño de la regresión logística, llevándola a priorizar la predicción de la clase mayoritaria a expensas de la minoritaria. Este comportamiento afecta particularmente la sensibilidad del modelo hacia la clase 0, que en este caso corresponde a los estados de confusión de los estudiantes.

La incapacidad para identificar estos casos representa una limitación crítica, especialmente en aplicaciones donde la detección de la clase minoritaria tiene una importancia estratégica.

Figura 29

Matrices de Confusión de Regresión Logística para Datos sin Balancear para Todos los Canales

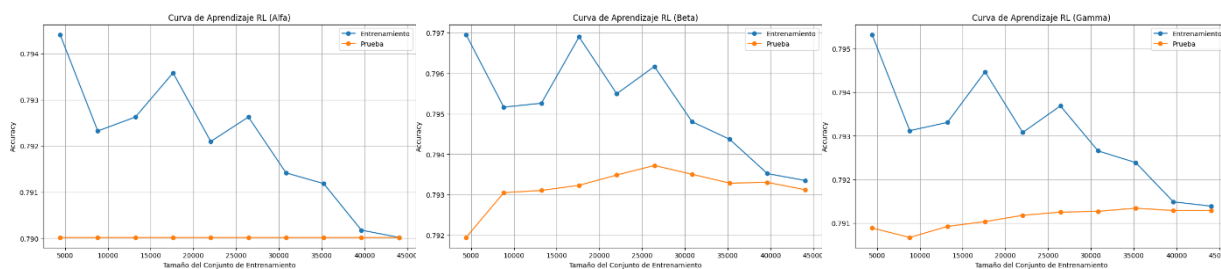


Nota. Matrices de confusión del modelo regresión logística para todos los electrodos sin técnicas de balanceo aplicada en las tres bandas de frecuencia (alfa, beta y gama).

Clasificación para Electrodo Específicos. La figura 30 muestra que al igual que en todos los canales, para canales específicos existe una brecha notable entre las precisiones de los conjuntos de entrenamiento y prueba. En la banda alfa, la precisión del entrenamiento presenta fluctuaciones significativas, comenzando alta, pero disminuyendo con el aumento del tamaño de los datos, mientras que la precisión de prueba se mantiene baja, cercana al azar. Para beta, aunque la precisión de entrenamiento es más estable, sigue mostrando una disminución gradual con el tamaño del conjunto de datos, y la precisión en prueba mejora ligeramente, pero sigue siendo limitada. En gama, se observa un comportamiento similar, con una baja precisión en prueba y una fluctuación inicial en la precisión de entrenamiento que converge a valores más bajos. Estas curvas reflejan que el modelo no logra generalizar adecuadamente, lo que confirma la influencia significativa del desbalance de clases.

Figura 30

Curvas de Aprendizaje de Regresión Logística para Datos sin Balancear para Canales Específicos



Nota. Curvas de aprendizaje del modelo regresión logística para electrodos específicos sin técnicas de balanceo aplicadas en las tres bandas de frecuencia (alfa, beta y gama).

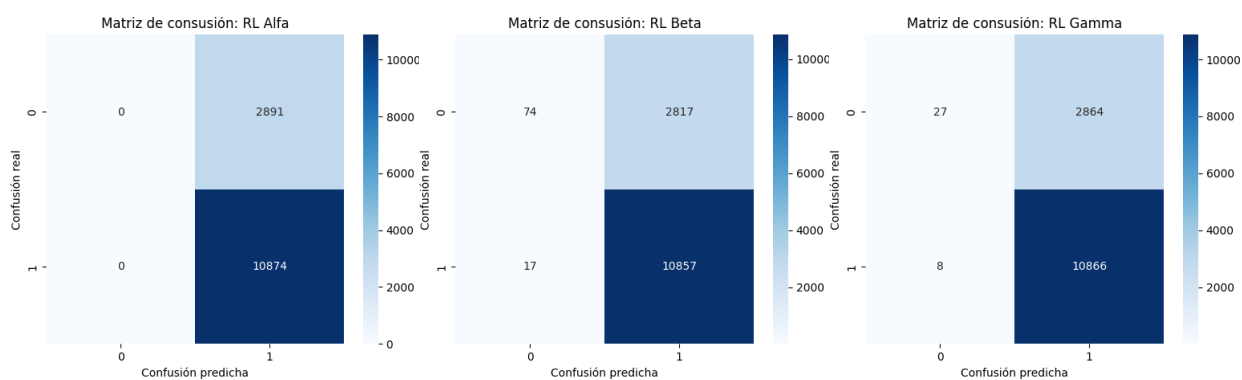
Al igual que en el caso anterior, las matrices de la figura 31 evidencian un marcado sesgo hacia la clase mayoritaria (clase 1). En la banda alfa, el modelo no logra predecir ninguna instancia de la clase 0, clasificando todas las instancias como clase 1, con 10,874 predicciones correctas

para esta clase y 2,891 falsos positivos. En beta, se observa un ligero mejor desempeño con 74 instancias correctamente clasificadas de la clase 0 y 17 falsos negativos, aunque el sesgo hacia la clase 1 persiste con 10,857 predicciones correctas para dicha clase. Finalmente, en gama, el modelo logra clasificar correctamente 27 instancias de la clase 0 y solo 8 falsos negativos, mientras que mantiene 10,866 predicciones correctas de la clase 1.

Figura 31

Matrices de Confusión de Regresión Logística para Datos sin Balancear para Canales

Específicos



Nota. Matrices de confusión del modelo regresión logística para electrodos específicos sin técnicas de balanceo aplicadas en las tres bandas de frecuencia (alfa, beta y gama).

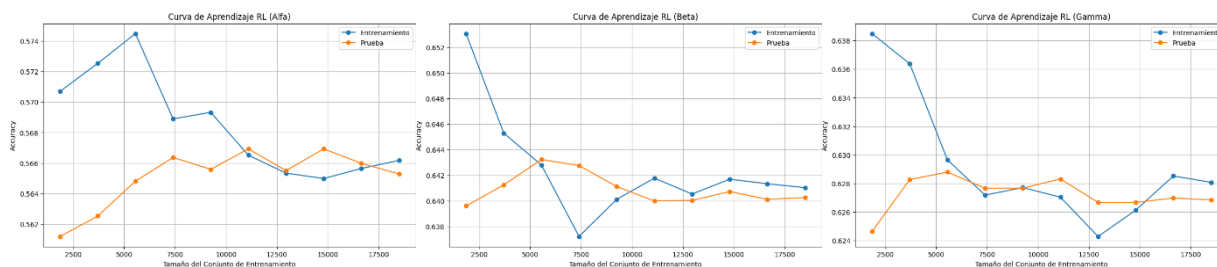
Datos con Undersampling

Clasificación para Todos los Electrodo. Las curvas de aprendizaje para la regresión logística aplicando balanceo mediante Undersampling que se presentan en la figura 34 muestran una tendencia similar a los datos sin balancear en la relación entre los conjuntos de entrenamiento y prueba. En la banda alfa, se observa inicialmente una alta precisión en el entrenamiento que disminuye conforme aumenta el tamaño del conjunto de datos, mientras que la precisión en prueba mejora gradualmente hasta converger con la del entrenamiento, sugiriendo

una mejora en la generalización. En beta, la precisión de entrenamiento sigue un patrón similar, con una caída inicial marcada que estabiliza el modelo, mientras que la precisión en prueba alcanza un rendimiento cercano al del entrenamiento con menor fluctuación. En gama, se repite un comportamiento comparable, donde las curvas de entrenamiento y prueba convergen, aunque con menor precisión general respecto a los otros canales. En conjunto, el balanceo undersampled parece haber reducido el impacto del desbalance de clases, mejorando la capacidad del modelo para generalizar y ofreciendo un rendimiento más consistente entre ambos conjuntos.

Figura 32

Curvas de Aprendizaje de Regresión Logística para Datos con Undersampling para Todos los Canales.



Nota. Curvas de aprendizaje del modelo regresión logística para todos los electrodos con técnica de Undersampling en las tres bandas de frecuencia (alfa, beta y gama).

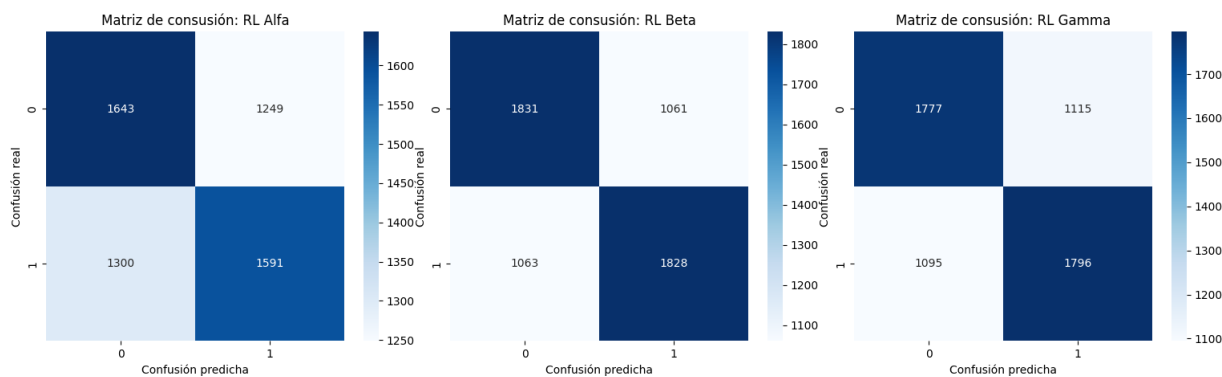
Las matrices de confusión para la regresión logística aplicando balanceo mediante undersampling muestran una mejora notable en la capacidad del modelo para identificar ambas clases en comparación con los datos sin balanceo. En la banda alfa, el modelo clasifica correctamente 1,643 instancias de la clase 0 y 1,591 de la clase 1, pero aún enfrenta desafíos importantes con 1,249 falsos positivos y 1,300 falsos negativos, lo que refleja una precisión moderada y una sensibilidad limitada hacia la clase minoritaria. En la banda beta, se evidencia

un desempeño superior, con 1,831 instancias correctamente clasificadas de la clase 0 y 1,828 de la clase 1. Además, los errores disminuyen, alcanzando 1,061 falsos positivos y 1,063 falsos negativos, lo que sugiere un mejor balance entre sensibilidad y especificidad.

En la banda gama, el modelo mantiene un rendimiento competitivo, clasificando correctamente 1,777 instancias de la clase 0 y 1,796 de la clase 1. Los errores son comparativamente menores, con 1,115 falsos positivos y 1,095 falsos negativos, lo que posiciona a esta banda como una opción intermedia en términos de equilibrio y precisión. Estas matrices reflejan que el balanceo mediante Undersampling es efectivo para mitigar el sesgo hacia la clase mayoritaria, permitiendo al modelo prestar más atención a la clase minoritaria (confusión).

Figura 33

Matrices de Confusión de Regresión Logística para Datos con Undersampling para Todos los Canales.



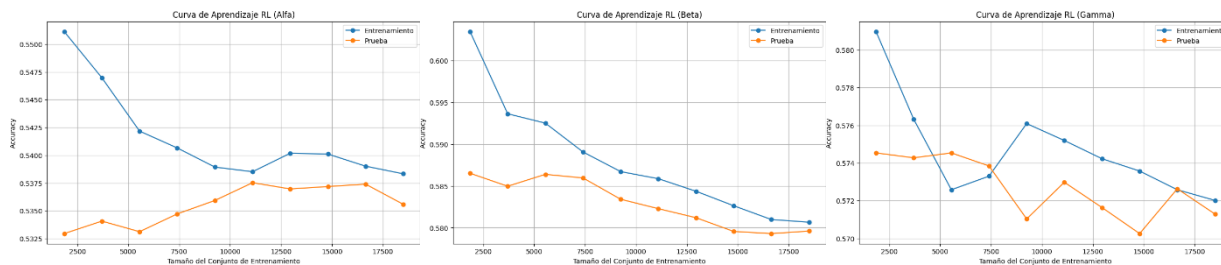
Nota. Matrices de confusión del modelo regresión logística para todos los electrodos con técnica de Undersampling en las tres bandas de frecuencia (alfa, beta y gama).

Clasificación para Electrodo Específicos. Las curvas de aprendizaje para la regresión logística con balanceo mediante Undersampling en los canales específicos se muestran en la figura 34, reflejan una mejora en la alineación entre los conjuntos de entrenamiento y prueba, aunque

persisten brechas significativas en algunos casos. En alfa, la precisión de entrenamiento disminuye rápidamente a medida que aumenta el tamaño del conjunto de datos, mientras que la precisión en prueba mejora ligeramente, acercándose a la del entrenamiento hacia el final, lo que indica una mejor generalización. En beta, se observa un comportamiento similar, con una caída inicial pronunciada en la precisión del entrenamiento y un incremento sostenido en la precisión del conjunto de prueba, logrando una convergencia notable. En gama, las curvas muestran fluctuaciones más pronunciadas, pero la tendencia general apunta a una menor diferencia entre entrenamiento y prueba, aunque con menor precisión general en comparación con los otros canales.

Figura 34

Curvas de Aprendizaje de Regresión Logística para Datos con Undersampling para Canales Específicos



Nota. Curvas de aprendizaje del modelo regresión logística para electrodos específicos con técnica de Undersampling en las tres bandas de frecuencia (alfa, beta y gama).

Las matrices de confusión de la figura 35 reflejan una mejora en el equilibrio de las predicciones entre las clases 0 y 1 tras aplicar balanceo mediante Undersampling. Aunque el modelo sigue presentando errores significativos, se observa un desempeño más balanceado en comparación con el modelo aplicado a datos sin balancear.

En la banda alfa, el modelo clasifica correctamente 1,530 instancias de la clase 0 y 1,568

de la clase 1. Sin embargo, persisten 1,323 falsos negativos y 1,362 falsos positivos, lo que indica una moderada precisión y sensibilidad. Estos errores reflejan que, si bien el modelo ha mejorado su capacidad para identificar ambas clases, aún tiene dificultades para discriminar con precisión, especialmente en escenarios donde las características de las clases son similares.

En la banda beta, se logra un mejor equilibrio entre las predicciones correctas y los errores. El modelo clasifica correctamente 1,654 instancias de la clase 0 y 1,620 de la clase 1, lo que representa una mejora respecto a alfa. Los errores disminuyen, con 1,271 falsos negativos y 1,238 falsos positivos, sugiriendo un aumento en la sensibilidad y especificidad del modelo. Este comportamiento indica que las características de la banda beta son más representativas, permitiendo una clasificación más precisa.

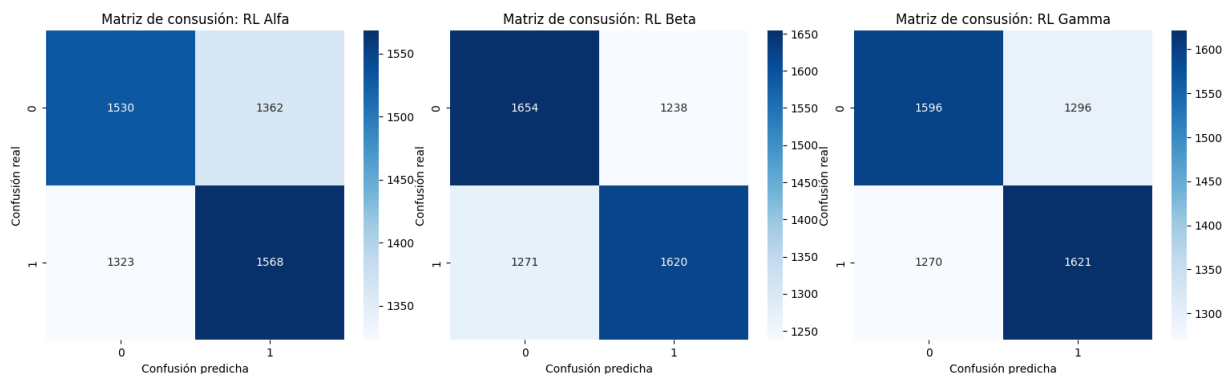
En la banda gama, el modelo mantiene un desempeño similar al de beta, con 1,596 instancias correctamente clasificadas de la clase 0 y 1,621 de la clase 1. Los errores son comparables, con 1,270 falsos negativos y 1,296 falsos positivos. Esto evidencia que el modelo logra un equilibrio razonable en esta banda, aunque aún persisten errores residuales que limitan su efectividad.

Los resultados muestran un comportamiento consistente en los tres canales, con mejoras evidentes en el equilibrio de las predicciones entre las clases 0 y 1 en comparación con los datos sin balancear. Sin embargo, los errores significativos en ambas clases sugieren que el modelo aún enfrenta desafíos para capturar patrones más complejos en los datos. Esto puede deberse a limitaciones inherentes al modelo de regresión logística, que tiende a ser menos efectivo en contextos donde las relaciones entre las variables no son estrictamente lineales

Figura 35

Matrices de Confusión de Regresión Logística para Datos con Undersampling para Canales

Específicos



Nota. Matrices de confusión del modelo regresión logística para electrodos específicos con técnica de Undersampling en las tres bandas de frecuencia (alfa, beta y gama).

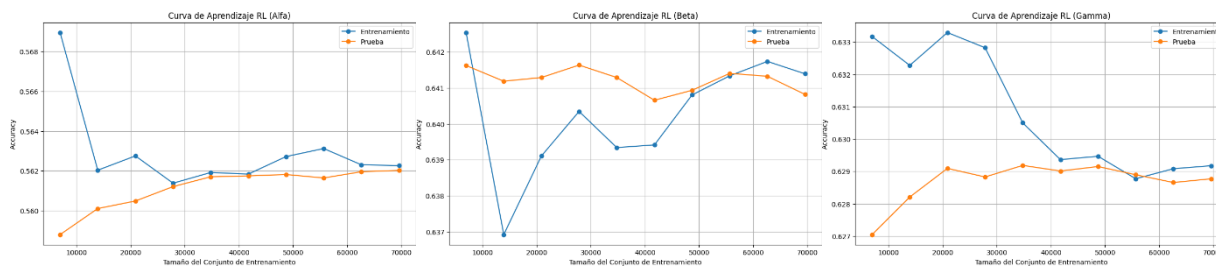
Datos con Oversampling

Clasificación para Todos los Electrodo. Al implementar el balanceo mediante Oversampling se obtienen las curvas de aprendizaje que se presentan en la figura 36. Se observa un desempeño variable con una tendencia a reducir la brecha entre los conjuntos de entrenamiento y prueba. En la banda alfa, la precisión en el conjunto de entrenamiento disminuye y posteriormente se mantiene relativamente estable, mientras que la precisión del conjunto de prueba presenta una ligera mejora progresiva, aunque con valores aún bajos, lo que indica una limitada capacidad de generalización. Para beta, se observa una tendencia de convergencia entre las curvas de entrenamiento y prueba, con una mejora significativa en el conjunto de prueba, sin embargo, la curva de entrenamiento reduce la precisión drásticamente mientras que los datos de prueba permanecen constantes, esto sugiere que el modelo no está aprendiendo de los datos de entrenamiento. En gama, las curvas muestran fluctuaciones pronunciadas en el entrenamiento, pero la precisión del conjunto de prueba incrementa gradualmente, acercándose a la del entrenamiento. En general, el balanceo mediante Oversampling parece mejorar la capacidad de

generalización del modelo, especialmente en las bandas beta y gama.

Figura 36

Curvas de Aprendizaje de Regresión Logística para Datos con Oversampling para Todos los Canales

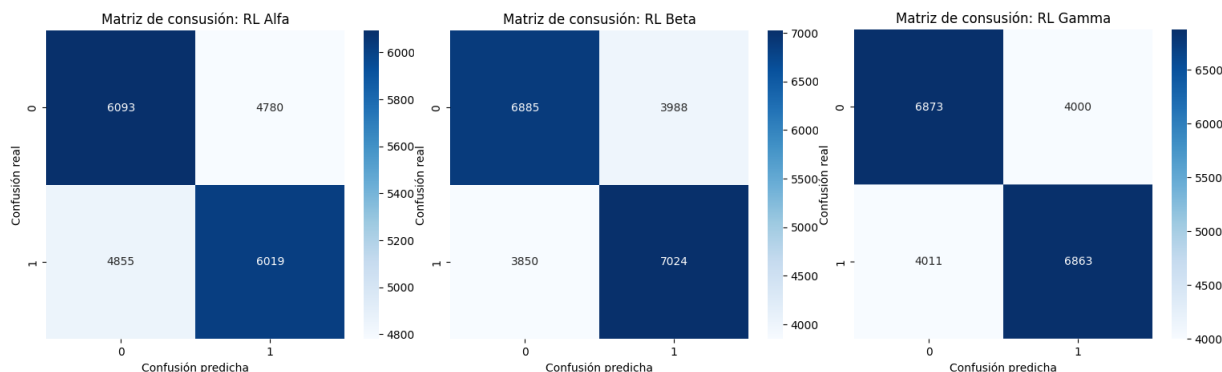


Nota. Curvas de aprendizaje del modelo regresión logística para todos los electrodos con técnica de Oversampling en las tres bandas de frecuencia (alfa, beta y gama).

Las matrices de confusión respaldan los resultados obtenidos en las curvas de aprendizaje, donde se evidencia una mejora en el equilibrio de las predicciones entre las clases 0 y 1, aunque persisten errores significativos. En la banda alfa, el modelo clasifica correctamente 6093 instancias de la clase 0 y 6019 de la clase 1, mientras que presenta 4780 falsos positivos y 4855 falsos negativos, lo que indica una distribución relativamente equilibrada de los errores. Para beta, el desempeño mejora ligeramente, con 6885 instancias correctamente clasificadas de la clase 0 y 7024 de la clase 1, mientras que los errores son menores en comparación, con 3988 falsos positivos y 3850 falsos negativos. En la banda gama, se logran 6873 predicciones correctas de la clase 0 y 6863 de la clase 1, mientras que los errores se distribuyen en 4000 falsos positivos y 4011 falsos negativos, mostrando un balance casi simétrico. Estas matrices reflejan que el Oversampling ayuda al modelo a reducir el sesgo hacia la clase mayoritaria, mejorando su capacidad para identificar ambas clases de manera más uniforme en los tres canales.

Figura 37

Matrices de Confusión de Regresión Logística para Datos con Oversampling para Todos los Canales

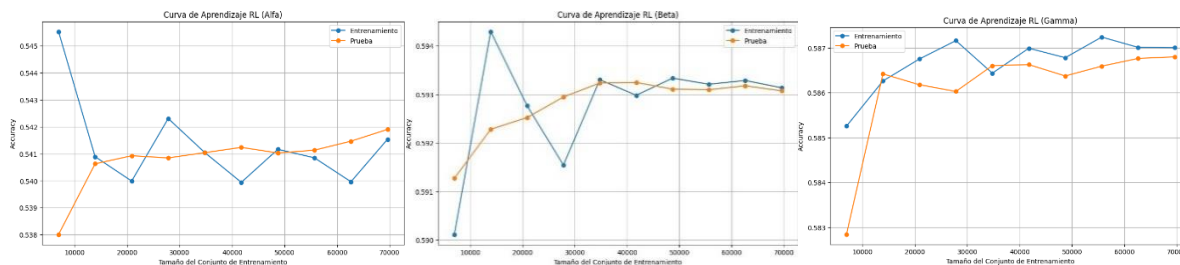


Nota. Matrices de confusión del modelo regresión logística para todos los electrodos con técnica de Oversampling en las tres bandas de frecuencia (alfa, beta y gama).

Clasificación para Electrodo Específicos. Al analizar las curvas de aprendizaje mostradas en la figura 38 para los electrodos más relevantes, se observa que, en la banda alfa, la curva de entrenamiento presenta una estabilidad creciente, mientras que la precisión en el conjunto de prueba mejora progresivamente, alcanzando una convergencia hacia el final con valores moderados. Esto sugiere que el modelo logra encontrar un punto de equilibrio adecuado. En las bandas beta y gama, las curvas destacan por una convergencia temprana, donde la precisión en el conjunto de prueba se aproxima rápidamente a la del entrenamiento, lo que indica un mejor equilibrio entre ambos y una mayor capacidad de generalización. En general, el análisis refleja una reducción significativa del sobreajuste, una mejora en la precisión de los conjuntos de prueba y un desempeño más equilibrado entre los diferentes canales específicos.

Figura 38

Curvas de Aprendizaje de Regresión Logística para Datos con Undersampling para Canales Específicos

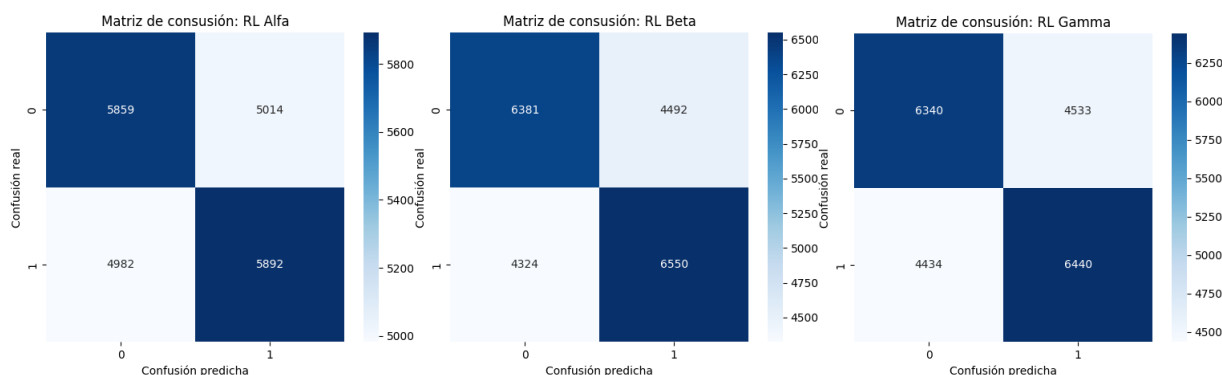


Nota. Curvas de aprendizaje del modelo regresión logística para electrodos específicos con técnica de Oversampling en las tres bandas de frecuencia (alfa, beta y gama).

Las matrices de confusión de la figura 39 reflejan un desempeño equilibrado entre las predicciones de las clases 0 y 1. En la banda alfa, el modelo clasifica correctamente 5859 instancias de la clase 0 y 5892 de la clase 1, mientras que se registran 5014 falsos positivos y 4982 falsos negativos, mostrando una ligera equidad en los errores. Para beta, el desempeño es similar, con 6381 instancias correctamente clasificadas de la clase 0 y 6550 de la clase 1, acompañadas de 4492 falsos positivos y 4324 falsos negativos, lo que sugiere una mejora respecto a alfa. En gama, se logran 6340 predicciones correctas de la clase 0 y 6440 de la clase 1, mientras que los errores se distribuyen en 4533 falsos positivos y 4434 falsos negativos, reflejando una buena consistencia entre ambas clases. Estas matrices muestran que el balanceo mediante Oversampling permite al modelo abordar de manera más efectiva el problema del desbalance de clases, logrando una clasificación más uniforme entre ambas categorías en los tres canales analizados.

Figura 39

*Matrices de Confusión de Regresión Logística para Datos con Oversampling para Todos
Canales Específicos*



Nota. Matrices de confusión del modelo regresión logística para electrodos específicos con técnica de Oversampling en las tres bandas de frecuencia (alfa, beta y gama).

Comparación de Modelos

La sección detalla el cálculo de métricas clave de evaluación, como sensibilidad, especificidad y precisión, para cada uno de los modelos desarrollados utilizando los diferentes métodos de balanceo (Undersampling, Oversampling y sin balanceo). Este análisis se realiza tanto para todos los electrodos como para los electrodos específicos, con el propósito de identificar qué modelo logra un mejor desempeño en términos de capacidad de clasificación y generalización. Estas métricas permiten evaluar no solo la proporción de verdaderos positivos correctamente identificados (sensibilidad), sino también la capacidad del modelo para evitar falsos positivos (especificidad) y su exactitud global (precisión). A través de esta evaluación integral, se busca establecer cuál de las configuraciones logra un equilibrio óptimo entre estas métricas, ofreciendo una clasificación confiable y robusta en los distintos escenarios analizados.

En la tabla 8 se presentan los resultados de los modelos de clasificación aplicados a los

datos sin balancear utilizando todos los electrodos. Los modelos analizados muestran una alta sensibilidad, destacándose la regresión logística y los árboles de decisión, con valores superiores al 97 %. Sin embargo, presentan una especificidad extremadamente baja, particularmente en la banda alfa, donde la regresión logística alcanza un mínimo de 0 y los árboles de decisión registran 0.0006. Aunque la precisión se mantiene en valores moderados (0.79 en promedio para todos los modelos y bandas), el marcado desequilibrio entre sensibilidad y especificidad evidencia un sesgo significativo hacia la predicción de la clase mayoritaria (estado de comprensión). Por otro lado, la tabla 9 resume los resultados de los modelos aplicados a datos de electrodos específicos. En este caso, se observa una ligera mejora en la especificidad, aunque esta sigue siendo baja en general. En las bandas beta y gama, los árboles de decisión destacan como los modelos más consistentes, logrando una precisión cercana a 0.798. Esta mejora, aunque limitada, sugiere que la utilización de electrodos específicos podría favorecer un mejor equilibrio entre las métricas, aunque persiste el sesgo hacia la clase mayoritaria. Además, esta observación resalta la necesidad de técnicas de balanceo para abordar el desbalance en los datos y mejorar el desempeño general de los modelos.

Tabla 8

Comparación de Modelos sin Balanceo para Todos los Canales

Modelo	Banda de potencia	Sensibilidad	Especificidad	Precisión
Árboles de Decisión	alfa	0.9999	0.0006	0.7900
	beta	0.9789	0.1432	0.8034
	gama	0.9714	0.1342	0.7956
KNN	alfa	0.9605	0.0515	0.7696
	beta	0.9486	0.2552	0.8030

	gama	0.9453	0.2193	0.7928
Regresión Logística	alfa	1	0	0.7899
	beta	0.9921	0.0387	0.7919
	gama	0.9947	0.0224	0.7905

Tabla 9

Comparación de Modelos sin Balanceo para Canales Específicos

Modelo	Banda de potencia	Sensibilidad	Especificidad	Precisión
Árboles de Decisión	alfa	0.9999	0.0006	0.7900
	beta	0.9880	0.0857	0.7985
	gama	0.9906	0.0691	0.7970
KNN	alfa	0.9555	0.0549	0.7664
	beta	0.9456	0.1736	0.7835
	gama	0.9427	0.1556	0.7774
Regresión Logística	alfa	1	0	0.7899
	beta	0.9985	0.0217	0.7933
	gama	0.9993	0.0124	0.7920

Para los modelos que emplean el método Undersampling, los resultados obtenidos se presentan en las tablas 10 y 11. En la tabla 10, que corresponde a todos los electrodos, se observa que los modelos muestran un mejor equilibrio entre sensibilidad y especificidad en comparación con los datos sin balancear. Los árboles de decisión y KNN se destacan en las bandas beta y gama, alcanzando sensibilidades cercanas al 69 % y especificidades en torno al 62 %. Aunque la precisión se reduce ligeramente en comparación con los datos sin balancear,

logrando valores superiores al 0.63 en beta con KNN y árboles de decisión, esta reducción se justifica por la mejora en la sensibilidad, lo que indica una mayor capacidad del modelo para clasificar correctamente ambas clases.

Por otro lado, la tabla 11, que presenta los resultados para los electrodos específicos, muestra que, aunque la sensibilidad es ligeramente menor en comparación con el análisis de todos los electrodos, la especificidad mejora en algunos casos. Por ejemplo, en beta, KNN logra una especificidad cercana al 61 %. En este contexto, el modelo de regresión logística se distingue por su desempeño más equilibrado, aunque menos sobresaliente en términos absolutos, con métricas de precisión que rondan el 0.55 para la mayoría de las bandas.

En general, el uso del método Undersampling permite un mejor balance entre las métricas clave, particularmente en bandas como beta y gama, destacando el impacto positivo de esta técnica en la mejora de la sensibilidad y la especificidad, a pesar de una ligera reducción en la precisión global. Esto sugiere que el Undersampling es una estrategia efectiva para abordar el desbalance de clases.

Tabla 10

Comparación de Modelos con Undersampling para Todos los Canales

Modelo	Banda de potencia	Sensibilidad	Especificidad	Precisión
Árboles de Decisión	alfa	0.6624	0.5688	0.5324
	beta	0.6402	0.6220	0.6311
	gama	0.6897	0.5193	0.6045
KNN	alfa	0.4801	0.5895	0.5348
	beta	0.6900	0.5812	0.6356
	gama	0.6624	0.5688	0.6155

Regresión Logística	alfa	0.5503	0.5681	0.5592
	beta	0.6323	0.6331	0.6327
	gama	0.6212	0.6144	0.6178

Tabla 11

Comparación de Modelos con Undersampling para Canales Específicos

Modelo	Banda de potencia	Sensibilidad	Especificidad	Precisión
Árboles de Decisión	alfa	0.6796	0.3800	0.5298
	beta	0.6938	0.4543	0.5740
	gama	0.7509	0.3869	0.5689
KNN	alfa	0.4064	0.6089	0.5076
	beta	0.5472	0.6123	0.5798
	gama	0.5191	0.6054	0.5623
Regresión Logística	alfa	0.5423	0.5290	0.5357
	beta	0.5603	0.5719	0.5661
	gama	0.5607	0.5518	0.5562

Para los modelos que emplean el método Oversampling, los resultados obtenidos se presentan en las tablas 12 y 13. En la tabla 12 se presentan los resultados utilizando todos los electrodos, se evidencia un mejor balance entre sensibilidad y especificidad en comparación con los datos sin balancear. Los modelos de KNN y regresión logística se destacan particularmente en las bandas beta y gama, logrando sensibilidades cercanas al 65 % y especificidades superiores al 66 % en algunos casos. La precisión también experimenta una mejora moderada, alcanzando valores de hasta 0.66 en beta para KNN, lo que refleja un buen desempeño general en la clasificación de

ambas clases. Estos resultados sugieren que el Oversampling contribuye a un balance más equitativo entre las métricas clave, especialmente en bandas donde el desbalance original impactaba más significativamente.

Por su parte, la tabla 13, que presenta los resultados para los electrodos específicos, muestra que los modelos mantienen un equilibrio razonable entre sensibilidad y especificidad, aunque con valores ligeramente más bajos en comparación con todos los electrodos. En beta, KNN y regresión logística logran un desempeño competitivo, con sensibilidades y especificidades superiores al 58 %, mientras que la precisión se mantiene entre 0.57 y 0.59. En la banda alfa, los árboles de decisión muestran un desempeño más discreto en términos de precisión (0.54), pero logran sensibilidades relativamente altas.

En general, el método Oversampling muestra ser una estrategia efectiva para mejorar el desempeño de los modelos de clasificación, especialmente en términos de especificidad, donde los resultados son notablemente superiores a los obtenidos sin balanceo o con Undersampling. Este enfoque es particularmente beneficioso para modelos como KNN y regresión logística, que alcanzan un desempeño equilibrado y consistente, especialmente en bandas como beta y gama, lo que resalta la utilidad del Oversampling para abordar el desbalance de clases y mejorar la capacidad de generalización del modelo.

Tabla 12

Comparación de Modelos con Oversampling para Todos los Canales

Modelo	Banda de potencia	Sensibilidad	Especificidad	Precisión
Árboles de Decisión	alfa	0.6976	0.4076	0.5526
	beta	0.6362	0.6294	0.6328
	gama	0.5800	0.6539	0.6170

KNN	alfa	0.5318	0.5784	0.5551
	beta	0.6529	0.6692	0.6611
	gama	0.6243	0.6632	0.6438
Regresión Logística	alfa	0.5637	0.5600	0.5618
	beta	0.6537	0.6352	0.6445
	gama	0.6231	0.6345	0.6288

Tabla 13

Comparación de Modelos con Oversampling para Canales Específicos

Modelo	Banda de potencia	Sensibilidad	Especificidad	Precisión
Árboles de Decisión	alfa	0.6875	0.3878	0.5376
	beta	0.8165	0.3581	0.5873
	gama	0.5483	0.6174	0.5828
KNN	alfa	0.3976	0.6326	0.5151
	beta	0.4885	0.6692	0.5789
	gama	0.4636	0.6603	0.5620
Regresión Logística	alfa	0.5416	0.5361	0.5389
	beta	0.6025	0.5898	0.5961
	gama	0.5848	0.5899	0.5873

Los modelos aplicados a los datos sin balancear mostraron un sesgo hacia a clase mayoritaria (comprensión), caracterizado por sensibilidades altas (superiores al 97%), pero especificidades extremadamente bajas, en algunos casos cercanas a cero. Este desequilibrio, aunque logra precisiones moderadas (0.79), compromete la capacidad de identificar

correctamente la clase minoritaria (estado de confusión), especialmente en bandas como alfa. La técnica de Undersampling permitió un mejor equilibrio entre sensibilidad y especificidad. En las bandas beta y gama, los árboles de decisión y KNN alcanzaron sensibilidades cercanas al 69% y especificidades superiores al 60%, logrando un desempeño más balanceado. Aunque la precisión disminuyó, el incremento en la sensibilidad refleja una mejora en la capacidad del modelo para clasificar ambas clases. El método de Oversampling se destacó como la técnica más efectiva, logrando un balance entre las tres métricas evaluadas (sensibilidad, especificidad y precisión). En bandas como beta y gama, KNN y regresión logística lograron especificidades superiores al 66% y precisiones cercanas al 0.66, consolidándose como las configuraciones más robustas. Esta técnica demostró ser particularmente efectiva tanto con todos los electrodos como con los electrodos específicos.

Conclusiones

La investigación presentada demuestra que es posible detectar estados de confusión en estudiantes durante actividades de aprendizaje a distancia mediante el uso de señales de electroencefalografía (EEG) y modelos de aprendizaje automático. Este enfoque aborda los desafíos de la falta de comunicación no verbal en entornos virtuales, permitiendo intervenciones personalizadas y en tiempo real que mejoran el rendimiento académico y fortalecen los procesos cognitivos.

Los análisis realizados identificaron que los electrodos en regiones temporales, parietales y occipitales (T7, T8, P7, P8, O1 y O2) presentan una mayor asociación con los estados de confusión, especialmente en las bandas beta y gama, vinculadas con la atención activa y procesos cognitivos avanzados. En contraste, los electrodos frontales (como F8 y AF4) mostraron menor relevancia en este contexto.

Entre los modelos evaluados, KNN con Oversampling destacó como el mejor, logrando una mayor especificidad tanto con todos los electrodos como con los seleccionados. Aunque presentó una ligera disminución en la precisión debido a la reducción en sensibilidad al implementarse en los electrodos específicos, mostró un desempeño notable en la banda beta.

El balanceo de datos mediante Oversampling, combinado con señales EEG de las bandas beta y gama y electrodos en las áreas temporal, parietal y occipital, proporciona una metodología robusta para identificar estados cognitivos de confusión. Este enfoque facilita la implementación de estrategias educativas personalizadas que optimizan la experiencia de aprendizaje en línea y mejoran la comprensión en tiempo real.

Referencias

- Abo-Zahhad, M., Ahmed, S. M., & Abbas, S. N. (2015). A new EEG acquisition protocol for biometric identification using eye blinking signals. *International Journal of Intelligent Systems and Applications*, 48-54.
- Ahire, N., Awale, R. N., & Wagh, A. (2023). Electroencephalogram (EEG) based prediction of attention deficit hyperactivity disorder (ADHD) using machine learning. *Applied Neuropsychology: Adult*, 1-12.
- Al-Khatir Al-Arimi, A. (2014). Distance learning. *Procedia - Social and Behavioral Sciences*, 82-88.
- Arana-Llanes, J., González-Serna, J., Pineda-Tapia, R., & Rendón-Miranda, J. (2016). Detección de Estilos de Aprendizaje e Interpretación de Ondas EEG para la Adaptación del Estado Cognitivo en Estudiantes E-learning. *Revista de Docencia e Investigación Educativa*, 2, 8-15.
- Betancourt, G., Bruges, A. L., & Sfer, A. L. (2017). Evaluación del modelo de aprendizaje en la modalidad de educación abierta y a distancia en la Universidad de Cartagena.
- Casson, A. J., Abdulaal, M., Dulabh, M., Kohli, S. a., & Trimble, E. (2018). Electroencephalogram. In *Seamless Healthcare Monitoring* (pp. 45-81). Springer International Publishing.
- Chen, C.-M., Wang, J.-Y., & Yu, C.-M. (2017). Assessing the attention levels of students by using a novel attention aware system based on brainwave signals. *British Journal of Educational Technology*, 348-369.
- Dakoure, C., Benlamine, M. S., & Frasson, C. (2021). Confusion detection using cognitive ability tests. *The International FLAIRS Conference Proceedings*.

- Fraga, M. (2024). *Lóbulo parietal: Integración sensorial y percepción espacial*. Neuroclass.
- Fuentes-Martinez, V. (2024). Plataforma EEG para la monitorización grupal de la atención en entornos de enseñanza. *Universidad de Granada*, 123-135.
- García Amaya, I. A., Aceves Fernández, M. A., Tovar Arriaga, S., & Ramos Arreguín, J. M. (2020). Regresión Logística y Máxima Verosimilitud para clasificación de personas diagnosticadas con Trastorno de Déficit de Atención e Hiperactividad. *La Mecatrónica en México*, 79-90.
- Hu, B., Li, X., Sun, S., & Ratcliffe, M. (2016). Attention Recognition in EEG-Based Affective Learning Research Using CFS+KNN Algorithm}. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*.
- Hu, P.-C., & Kuo, P.-C. (2017). Adaptive learning system for E-learning based on EEG brain signals. *IEEE 6th Global Conference on Consumer Electronics (GCCE)*.
- Impulsa, N. (2018). *Funciones de los lóbulos cerebrales*.
- Kyriakides, G., & Margaritis, K. G. (2019). *Hands-On Ensemble Learning with Python: Build highly optimized ensemble machine learning models using scikit-learn and Keras*. Packt Publishing.
- Lancheros-Cuesta, D. J., Arias, J., Forero, Y., & Duran, A. (2018). Evaluation of e-learning activities with NeuroSky MindWave EEG. *2018 13th Iberian Conference on Information Systems and Technologies (CISTI)*, 1-6.
- León, A. (2007). Qué es la educación. *Educere*, 11.
- Liang, Y., Liu, X., Qiu, L., & Zhang, S. (2018). An EEG study of a confusing state induced by information insufficiency during mathematical problem-solving and reasoning. *Computational Intelligence and Neuroscience*, 1-13.

- Luo, Y., Zhou, Z., Yang, H., & He, B. (2023). Study on the human-brain confusion levels and corresponding EEG levels. *Applied and Computational Engineering*, 113-120.
- Mineducación. (2022). *Nota Orientadora Modalidades*. Colombia:
https://www.mineducacion.gov.co/1780/articles-408425_recurso_6.pdf.
- Ni, Z., Yuksel, A. C., Ni, X., Mandel, M. I., & Xie, L. (2017). *Confused or not confused?: Disentangling brain activity from EEG data using bidirectional LSTM recurrent neural networks*. Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics.
- Omar, M. (2022). *EEG data / Distance learning*. Data set: Kaggle.
- Sethi, C., Dabas, H., Dua, C., Dalawat, M., & Sethia, D. (2018). *EEG-based attention feedback to improve focus in E-learning*. Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence.
- Sun, H., Wu, C.-W., Zhang, D., Makay, Ö., Zhao, Y., Carcofaro, P., . . . Pappalardo, V. (2019). New paradigms for neural monitoring in thyroid surgery. *Surgical Technology International*, 79-86.
- Tambe, N., & Khachane, A. (2016). Mood based E-learning using EEG. *International Conference on Computing Communication Control and Automation (ICCUBEA)*, 1-4.
- Trigka, M., Dritsas, E., & Mylonas, P. (2023). Mental confusion prediction in E-learning contexts with EEG and machine learning. In *Lecture Notes in Networks and Systems* (pp. 195-200). Springer Nature Switzerland.
- UNAM. (2019). *MANUAL SISTEMA 10/20 INTERNACIONAL*. México.
- Vachha, B. A., Massoud, T. F., & Huang, S. Y. (2022). Anatomy of the Cerebral Cortex, Lobes, and Cerebellum.