

## **Análisis comparativo de técnicas de aprendizaje automático y aprendizaje profundo para la detección de TDAH**

Ismael Santarrosa-López<sup>1</sup>, Giner Alor-Hernández<sup>1</sup>,  
Maritza Bustos-López<sup>1</sup>, Laura Nely Sánchez-Morales<sup>2</sup>,  
José Luis Sánchez-Cervantes<sup>2</sup>

<sup>1</sup> Tecnológico Nacional de México,  
Instituto Tecnológico de Orizaba,  
México

<sup>2</sup> Consejo Nacional de Ciencia y Tecnología,  
Instituto Tecnológico de Orizaba,  
México

{m16011220, giner.ah, laura.sm, jose.sc}@orizaba.tecnm.mx,  
maritbustos@gmail.com

**Resumen.** El Trastorno por Déficit de Atención e Hiperactividad (TDAH) es un trastorno neurológico que afecta del 5-10% de los niños y 2-5% de los adultos en el mundo, causando maltrato infantil, rechazo escolar y aislamiento social. Por ello es importante detectar y tratar el TDAH de manera efectiva. En este trabajo, se presenta un análisis comparativo para determinar qué conjunto de datos y técnicas de Aprendizaje Automático y Aprendizaje Profundo ofrecen mejores resultados con base a diferentes métricas conocidas para la detección del TDAH. Los resultados indicaron que los algoritmos Máquina de Vectores de Soporte, Regresión Lineal, K-Vecinos más Cercano, Árbol de Decisiones y CatBoost son los de mejor clasificación con resultados de entre 66-97% de correcta clasificación de TDAH, superiores a los resultados obtenidos por los algoritmos de Aprendizaje Profundo, por lo que las técnicas de Aprendizaje Automático indican ser más efectivas en la detección de TDAH.

**Palabras clave:** Aprendizaje automático, aprendizaje profundo, biomarcadores, conjunto de datos, TDAH.

### **Comparative Analysis of Machine Learning and Deep Learning Techniques for ADHD Detection**

**Abstract.** Attention Deficit Hyperactivity Disorder (ADHD) is a neurological disorder affecting 5-10% of children and 2-5% of adults worldwide, causing child abuse, school rejection, and social isolation. It is therefore important to detect and treat ADHD effectively. In this paper, a comparative analysis is presented to determine which Machine Learning and Deep Learning datasets and techniques provide better results based on different known metrics for ADHD detection. The

results indicated that the Support Vector Machine, Linear Regression, K-Nearest Neighbors, Decision Tree, and CatBoost algorithms are the best classifiers with results between 66-97% correct ADHD classification, superior to the results obtained by the Deep Learning algorithms so that Machine Learning techniques indicate to be more effective in ADHD detection.

**Keywords:** ADHD, biomarkers, dataset, deep learning, machine learning.

## 1. Introducción

El Trastorno por Déficit de Atención e Hiperactividad (TDAH) es un trastorno neurológico que afecta a cerca del 5% a 11% de los niños y 2% a 5% de los adultos en todo el mundo, según la Organización Mundial de la Salud (OMS) [1] y se estima que en México hay aproximadamente 1.5 millones de los 33 millones de niños y adolescentes en todo el país que tienen la probabilidad de padecer este trastorno según el CDC [2].

El TDAH tiene un impacto significativo en la vida de las personas, incluyendo dificultades en la escuela, en el trabajo y en las relaciones interpersonales [3]. Por lo tanto, es importante detectar y tratar el TDAH de manera efectiva. Este trastorno se caracteriza por dificultades en la atención, la concentración y el control de impulsos; estos síntomas llevan a problemas en los entornos ya mencionados.

Además, el TDAH también se relacionó con un mayor riesgo de abuso de sustancias y trastornos de salud mental, como la ansiedad y la depresión. Un estudio realizado por el Instituto Nacional sobre el Abuso de Drogas encontró que las personas con TDAH son más propensas a abusar de sustancias y tener problemas relacionados con las drogas [1].

La importancia del uso de técnicas de Inteligencia Artificial en la identificación de enfermedades ha sido reconocida por la comunidad científica y médica. Estas técnicas permiten la automatización y mejora de procesos de diagnóstico y terapéuticos, lo que resulta en una atención más eficiente y efectiva de la salud. El aprendizaje automático permite a las máquinas aprender a partir de datos y hacer predicciones sobre nuevos datos sin ser explícitamente programadas [4].

Por su parte, el aprendizaje profundo es un tipo de aprendizaje automático que imita la estructura y función de la corteza cerebral, lo que le permite a la IA aprender a partir de imágenes, audio y otras formas complejas de datos [4]. El uso de técnicas de aprendizaje automático y aprendizaje profundo en la identificación de enfermedades demostró ser efectivo en numerosos estudios y aplicaciones. Por ejemplo, el uso de técnicas de aprendizaje profundo para la detección de enfermedades cardiovasculares [5] o la detección temprana del Parkinson [6].

En la literatura existen trabajos que han reportado revisiones y análisis de conjuntos de datos para la detección del TDAH, como los realizados por Vikas Khullar et al. [7] y Hui Wen Loh et al. [8]. Sin embargo, estos trabajos tienen un enfoque basado en el análisis de un solo conjunto de datos o en un grupo pequeño de técnicas de inteligencia artificial, ya sea sólo de Aprendizaje Automático o sólo de Aprendizaje Profundo.

Como diferencia, en este trabajo se utilizan múltiples técnicas de inteligencia artificial para el análisis y evaluación de 4 conjuntos de datos diferentes para identificar

qué algoritmos ofrecen los mejores resultados, así como también los biomarcadores más relacionados para la identificación del TDAH.

La estructura de este artículo es la siguiente: en la sección 2 se presenta el estado del arte con un conjunto de investigaciones y trabajos relacionados con la identificación del TDAH, en la sección 3 se presentan los conjuntos de datos utilizados para este análisis y sus características, en la sección 4 se presenta los resultados del análisis y evaluación realizado a los conjuntos de datos con técnicas de inteligencia artificial previamente seleccionadas; en la sección 5 se presenta la discusión de los resultados y finalmente, en la sección 6, las conclusiones y el trabajo a futuro.

## **2. Trabajos relacionados**

A continuación, se presentan algunos de los trabajos relacionados con respecto al diagnóstico del TDAH mediante técnicas de inteligencia artificial.

### **2.1. Enfoques basados en técnicas de aprendizaje profundo para la identificación del TDAH**

Arthi y Tamilarasi [9], realizaron un sistema de red neuronal híbrido que consta de mapas autoorganizados de Kohonen seguidos de una función de base radial que utilizó valores de pertenencia difusos como entrada. El modelo se entrenó en dos fases sobre datos de TDAH.

Se observó que el modelo híbrido de redes SOM y RBF tiene una mejor clasificación en comparación con las redes neuronales de retropropagación. Kuang y He [10] propusieron el uso de una Red de Creencia Profunda (DBN), esta se compone por una pila de máquinas de Boltzmann (RBM).

DBN se aplicó para predecir los subtipos de TDAH, pero se necesitó un preprocesamiento para los datos específicos de fMRI. El método propuesto demostró ser eficaz para discriminar el TDAH del control y los subtipos. Con la idea de acelerar el proceso de clasificación clínico entre pacientes con TDAH y AOS, Chu et al.[11] realizaron tres modelos de aprendizaje automático para encontrar la mejor manera de ayudar al médico a realizar el diagnóstico.

Los resultados mostraron que los algoritmos de redes neuronales se adaptaron bien a la tarea de clasificación en cuestión, porque tenían una tasa de clasificación errónea menor que el modelo CART y CHAID.

Bo Miao & Yulin Zhang [12] utilizaron el proceso de Amplitud Fraccional Alterada de Fluctuación de Baja Frecuencia que sirvió para obtener la actividad cerebral en reposo, obteniendo un subconjunto de características para después complementar con el algoritmo Relief y obtener las características candidatas y así lograr la clasificación con la SVM, resultando en una clasificación viable de pacientes con TDAH en función a los datos de las neuroimágenes.

Shao L. et al. [13] utilizaron varias técnicas de aprendizaje automático, incluidas las redes neuronales profundas, para clasificar el TDAH. Se propuso como alternativa el bosque profundo, siendo un conjunto de conjuntos de árboles de decisión.

**Tabla 1.** Comparación de los conjuntos de datos para el diagnóstico de TDAH.

Conjunto de datos	Núm. de atributos	Núm. de registros	Descripción de los datos
HYPERAKTIV	+25	51 sujetos con TDAH y 51 sin TDAH	Datos de movimiento, ritmo cardíaco y concentración
ADHD200	23	776	Datos atómicos de Resonancia Magnética Multifuncional
Working Memory and Reward in Children with and without ADHD	8	79 niños	Imágenes por Resonancia Magnética Funcional 3D
Working Memory and Reward in Adults	8	24 adultos	Imágenes por Resonancia Magnética Funcional 3D
Eeg Data for ADHD	12	61 niños con TDAH y 60 sin TDAH	Datos atómicos de señales cerebrales (EEG)

Los resultados experimentales en los conjuntos de datos mostraron que este método logró un rendimiento superior que los métodos informados en la literatura sobre los conjuntos de datos de prueba de retención.

Los autores Mao, et al. [14] utilizaron la computación granular mediante el análisis de exploraciones de fMRI a través de perspectivas espaciales y temporales. Se utilizó CNN para conocer las características espaciales de cada cuadro de todo el escaneo.

Se utilizaron núcleos de convolución 3D y después se extendieron a la dimensión temporal, aprendiendo patrones espaciales y temporales simultáneamente con la operación de convolución 4D (este encapsula varias imágenes a la vez y aprender el movimiento del cerebro). Los resultados arrojaron una precisión del 71,3 % y un AUC del 0,80.

En el estudio realizado por Tosun [15], se reveló el canal más efectivo y el estado de registro más efectivo para el diagnóstico de TDAH. Los datos de EEG se aplicaron a la memoria a corto plazo (LSTM), la máquina de vectores de soporte y los clasificadores de redes neuronales artificiales, pero la mayor precisión se obtuvo con LSTM.

Esta métrica se calculó como 88,88% en el canal "Fp1, F7" y 92,15% en el estado de reposo con los ojos cerrados demostrando ser eficaz en el diagnóstico del TDAH. En cuanto a Khullar, et al. [16], se utilizó el algoritmo de red neuronal convolucional bidimensional (CNN) y la red neuronal convolucional bidimensional híbrida: memoria a largo plazo a corto plazo (2D CNN-LSTM) para la clasificación del TDAH a partir de controles de desarrollo típico.

El método propuesto logró una mejora significativa en el análisis y detección de parámetros. Los resultados construyeron un modelo adecuado e inteligente para diagnosticar comparativamente el TDAH a partir de controles sanos.

Los métodos tuvieron una buena precisión con el modelo 2D híbrido donde mostró más del 98% de precisión. En el estudio, se analizaron los efectos de los estímulos fóticos a diferentes frecuencias y en diferentes canales en el diagnóstico de TDAH.

## **2.2. Enfoques basados en técnicas de aprendizaje automático para la identificación del TDAH**

Anuradha, et al. [17] usó el algoritmo SVM; se implementó en una herramienta llamada Clementine, este software se utiliza para realizar minería de datos. Después de ejecutar una base de datos de acuerdo con el SVM, proporcionó su propia interpretación de los resultados. El algoritmo SVM demostró un porcentaje de 88,674 % de éxito.

Por otro lado, Bautista et al. [18] utilizó una extensión del método Deformación Dinámica del Tiempo (DTW) para medir la similitud entre dos secuencias temporales; dicha extensión se codificó con clasificadores de clase GMM y APE; junto a dispositivo Kinect© para obtener patrones de comportamiento. El modelo se aplicó a un conjunto de datos multimodal y obtuvo mejoras con respecto a las técnicas puras de DTW.

En Duda et al. [19] utilizaron seis algoritmos diferentes de aprendizaje automático en datos de 2,925 sujetos, implementando métodos de selección de características para adaptar cada algoritmo y reducir el conjunto original. Los algoritmos SVC, LDA, Categorical Lasso y Logistic Regression arrojaron resultados de entre el 0.962 y 0.965 de exactitud, siendo los mejores modelos para su clasificación.

Por su parte, Uluyagmur-Ozturk et al. [20] se enfocaron en la clasificación de los participantes con TDAH y los que contaban con el Trastorno del Espectro Autista (TEA) en función al reconocimiento de emociones; los datos se usaron en los clasificadores de Árboles de Decisión. Los resultados indicaron que el TDAH y el TEA se clasificaron con un 90 % de precisión mediante el uso del algoritmo AdaBoost.

En el trabajo de Itani et al. [21] desarrollaron modelos de soporte de diagnóstico interpretables con árboles de decisión, C4.5 y algunas reglas, capaces de implementar a una base de datos médica. Al aplicar este marco en el conjunto ADHD-200 junto con imágenes de fMRI, logró predecir sujetos con TDAH. Mohammadhasani et al. [22] desarrolló un módulo de Instrucción Asistida por Computadora (CAI) que incluye agentes pedagógicos virtuales, utilizados en el aprendizaje en línea, que guían a los usuarios en entornos multimedia y, brindar apoyo cognitivo y emocional a los estudiantes con TDAH.

El estudio demostró que los agentes generan beneficios educativos para los estudiantes con TDAH. Khanna & Das [23] desarrollaron un método para analizar la variación pupilo métrica utilizando el aprendizaje automático, con la hipótesis de que reflejaría con precisión los pacientes con TDAH. El resultado de este método superó las tasas de diagnóstico clínico en un 80%. Christiansen et al. [24] realizaron un estudio con el objetivo de validar si el algoritmo CAARS logra discriminar pacientes con TDAH. Al aplicar el algoritmo, logró diferenciar sujetos con TDAH, obesidad y ludopatía con una precisión global del 80%.

Por último, los autores Maniruzzaman M. et al. [25] presentaron una investigación sobre los factores de riesgo en niños con TDAH. El estudio ilustró que el clasificador Random Forest proporciona una excelente clasificación y predicción de niños con TDAH dando más del 85% de correcta clasificación.

En nuestro caso, para el desarrollo de este trabajo, se utilizaron técnicas de aprendizaje automático y técnicas de aprendizaje profundo, estas técnicas se aplicaron a varios conjuntos de datos, cuya naturaleza de los datos van desde datos numéricos, datos provenientes de Electroencefalogramas (EEG) hasta imágenes de resonancia magnética (MRI); estos conjuntos se describen en la siguiente sección.

**Tabla 2.** Comparación de resultados obtenidos en el conjunto de datos HYPERAKTIV.

Algoritmos AA	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Linear Regression	0.9047	0.9090	0.9166	0.9000	0.9166
CatBoost	0.9047	0.9047	0.9047	0.9047	0.9047
SVM	0.8571	0.8750	0.8750	0.8571	0.8750
Gradient Boosting	0.8571	0.8545	0.8611	0.8421	0.8611
AdaBoost	0.8571	0.8545	0.8611	0.8421	0.8611
Random Forest	0.8095	0.8136	0.8194	0.8000	0.8194
Decision Tree	0.7619	0.7777	0.7777	0.7619	0.7777
LightGBM	0.7619	0.7777	0.7777	0.7619	0.7777
XGBoost	0.7619	0.7777	0.7777	0.7619	0.7777
KNN	0.6666	0.6636	0.6666	0.6315	0.6666

### 3. Conjuntos de datos para la identificación del TDAH

En esta sección se describen los distintos conjuntos de datos utilizados para la identificación del TDAH. Se utilizaron 5 conjuntos de datos, cada uno con distintos enfoques y características.

- **HYPERAKTIV:** HYPERAKTIV [26] es un conjunto de datos público con datos relacionados con la frecuencia cardiaca y de datos de movimiento de 51 pacientes adultos sin TDAH y 52 pacientes clínicos con TDAH. HYPERAKTIV incluye datos salud, atributos como la edad y el sexo, datos de salida de una prueba neuropsicológica informatizada e información sobre el estado mental del paciente. El conjunto de datos HYPERAKTIV se utilizó en [27] para el diagnóstico del TDAH en adultos utilizando algoritmos de aprendizaje automático.
- **ADHD200:** Por su parte, ADHD200 [28] ofrece un conjunto de datos preprocesados procedentes del Concurso Mundial ADHD-200. ADHD200 consiste en 776 conjuntos de datos anatómicos y de RMf en estado de reposo, incluyendo 285 de niños y adolescentes con TDAH. ADHD200 incluye datos sobre edad, sexo, estado de diagnóstico, estado de medicación, medidas dimensionales de los síntomas del TDAH, puntuación de algunos de los cuestionarios más comunes para la detección del TDAH y cociente intelectual (CI). El conjunto de datos ADHD200 se utilizó en [29] para el desarrollo de métodos de clasificación del TDAH basados en CNN 3D y redes neuronales de aprendizaje profundo multicanal en [30]. Los resultados obtenidos en ambos estudios indicaron una exactitud del 69,15% al 95%, respectivamente.
- **Working Memory and Reward in Children with and without ADHD:** este conjunto de datos (disponible en OpenNeuro.org [31]) fue el resultado de datos registrados por resonancia magnética funcional (fMRI) y puntajes longitudinales de medidas estandarizadas para la capacidad cognitiva, síntomas de TDAH y habilidades de lectura.

**Tabla 3.** Comparación de los resultados obtenidos en el conjunto de datos ADHD-200.

Algoritmo AA	Accuracy	Precision	Recall	F1-Score	ROC-AUC
KNN	0.9750	0.9705	0.9791	0.9787	0.9791
CatBoost	0.9750	0.9736	0.9772	0.9767	0.9772
Decision Tree	0.9750	0.9800	0.9687	0.9795	0.9687
Gradient Boosting	0.9750	0.9800	0.9687	0.9795	0.9687
LightGBM	0.9750	0.9800	0.9687	0.9795	0.9687
XGBoost	0.9750	0.9800	0.9687	0.9795	0.9687
Random Forest	0.9750	0.9800	0.9687	0.9795	0.9687
AdaBoost	0.9750	0.9800	0.9687	0.9795	0.9687
Linear Regression	0.9500	0.9479	0.9479	0.9583	0.9479
SVM	0.9250	0.9194	0.9270	0.9361	0.9270

El objetivo de las tareas era explorar la memoria de trabajo y el procesamiento de la retroalimentación en niños con desarrollo típico y con diagnóstico de TDAH. Se recogieron datos de 79 niños con edades comprendidas entre los 8, 6 y 12 años; 35 niños tenían un diagnóstico formal de TDAH [32].

- **Working memory and reward in adults:** En un estudio posterior, se creó un conjunto de datos titulado "Working Memory and Reward in Adults" [33] (disponible en OpenNeuro.org) con las mismas características que su predecesor y que incluye datos de 24 individuos adultos que realizaron las mismas tareas. El conjunto de datos anteriormente mencionado tiene antecedentes en el estudio presentado por [34] y [35]. Booth et al. [34] analizaron datos de fMRI obtenidos de cuatro tareas de memoria de trabajo visoespacial (VSWM) para detectar casos de TDAH. Los resultados indicaron una exactitud del 92,5% en la clasificación del TDAH.

Hammer et al. [35] probaron el efecto interactivo de la retroalimentación y la recompensa en la memoria de trabajo visoespacial en niños con TDAH. Las pruebas recogieron datos MRI de 17 niños con TDAH y 17 niños de control mientras realizaban un seguimiento espacial de letras en una pantalla. Los resultados indicaron que el rendimiento de los niños con TDAH es similar a los niños de sin TDAH sólo cuando se les dio retroalimentación mediante una recompensa.

- **EEG data for ADHD:** Por su parte, EEG DATA FOR ADHD [36] es un conjunto de datos que contiene información de 61 niños con TDAH y 60 pacientes sanos. El registro EEG se realizó durante tareas de atención visual, se pedía a los niños que contaran el número de personajes de cada imagen; cada imagen se mostraba inmediatamente después de la respuesta del niño.

El registro EEG se basó en registros EEG estándar de 19 canales (Fz, Cz, Pz, C3, T3, C4, T4, Fp1, Fp2, F3, F4, F7, F8, P3, P4, T5, T6, O1, O2) a una frecuencia de muestreo de 128 Hz, y se colocaron electrodos A1 y A2 en los lóbulos de las orejas.

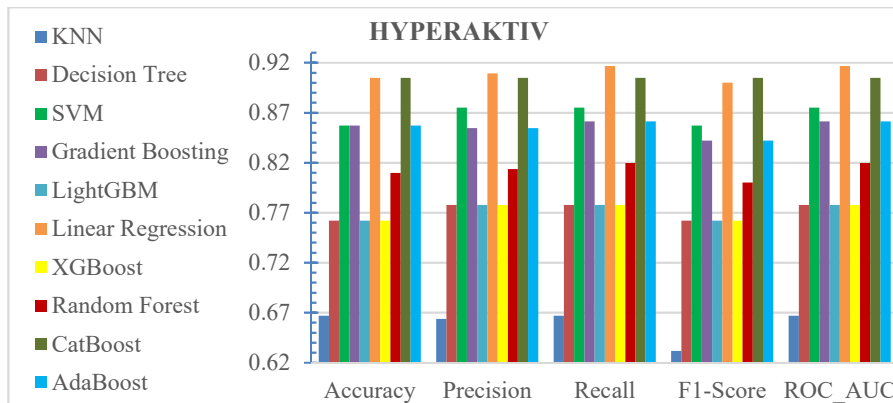


Fig. 1. Gráfica de los resultados de la tabla 2 para el conjunto de datos HYPERAKTIV.

El conjunto de datos EEG DATA FOR ADHD se utilizó en otros estudios para la detección del TDAH a partir de las características del EEG. En este sentido, Mohammadi et al. [37] clasificaron las características no lineales del EEG en niños con y sin TDAH como equivalentes a la atención. Los resultados confirmaron el defecto en el segmento del cerebro anterior de los niños con TDAH.

Barua et al. [38] utilizaron señales de EEG para plantear un nuevo modelo de clasificación manual para diferenciar a los individuos con TDAH. El modelo propuesto utilizó la Transformada Q Wavelet Sintonizable (TQWT) para generar sub-bandas Wavelet y un nuevo patrón de motivo ternario (TMP). Los resultados obtenidos mediante validaciones cruzadas arrojaron porcentajes de 95,57% y 77,93% de precisión en la clasificación.

La tabla 1 presenta una breve comparación de los conjuntos de datos utilizados, en esta tabla se muestran características como: cantidad de registros, número de atributos y breve descripción de los datos. Es importante mencionar que durante el análisis de los Datasets para su implementación, dos de ellos pertenecen al mismo grupo de datos, “Working Memory and Reward in Children with y without ADHD” y “Working Memory and Reward in Adults”, por lo que este último se omitió, quedando así sólo 4 conjuntos de datos para el análisis.

#### 4. Análisis comparativo de técnicas de aprendizaje automático y aprendizaje profundo

En esta sección se detalla el análisis realizado de los conjuntos de datos revisados anteriormente con distintas técnicas de inteligencia artificial, las cuales aplican técnicas de aprendizaje automático y aprendizaje profundo.

Dentro de los algoritmos de aprendizaje automático se utilizaron: KNN (K-Nearest-Neighbor), Decision Tree, SVM (Support Vector Machine), LightGBM [39], Linear Regression [40], Gradient Boosting, XGBoost, CatBoost y AdaBoost [41].



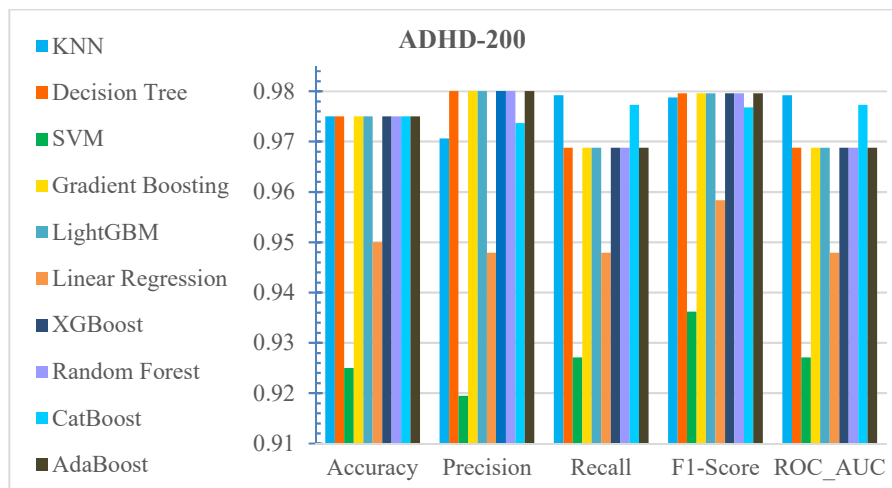


Fig. 2. Gráfica de resultados de la tabla 3 para el conjunto de datos ADHD-200.

Para la parte de aprendizaje profundo, se utilizaron los siguientes algoritmos: CNN (Convolutional Neural Network) y Multi-modality 3D CNN (es una variante de CNN, pero aplicada a imágenes en 3D) [42].

Estas técnicas de Aprendizaje Profundo se seleccionaron dado que la información de los conjuntos de datos a aplicar fue extraída con Matlab y que son compatibles con la herramienta Image Processing Toolbox de MATLAB.

Para los algoritmos de aprendizaje automático, se realizó una configuración experimental donde en primera instancia, se filtraron las características relevantes de cada conjunto de datos, evitando incluir aquellas características cuya presencia no aporta valor al resultado y omitiendo valores nulos.

Los conjuntos de datos se dividieron en un 80% de entrenamiento y un 20% de pruebas. Se agregaron más de un valor a cada parámetro utilizado, parámetros que varían según el algoritmo a utilizado, con el fin de utilizar la propiedad `best_estimator_` al momento de realizar el entrenamiento para obtener los parámetros que mejor se ajustan y así obtener una clasificación óptima.

En cada algoritmo, se obtuvieron las siguientes métricas: exactitud (accuracy), precisión, exhaustividad (recall), F1-Score y el área bajo la curva (ROC-AUC). Se utilizó la matriz de confusión y otras métricas por ser ampliamente aceptadas en la literatura. Los algoritmos pertenecientes al aprendizaje automático fueron aplicados a los conjuntos de HYPERAKTIV y ADHD-200 en el lenguaje de programación Python junto con el apoyo de la biblioteca de funciones Scikit-Learn.

En la tabla 2 y en la tabla 3, se muestran los resultados obtenidos respectivamente. La Figura 1 y la Figura 2 muestra una gráfica comparativa de los resultados. Los conjuntos de datos evaluados con algoritmos de aprendizaje profundo se realizaron con Matlab. Se obtuvieron las cinco métricas ya mencionadas. En la Tabla 4 se muestra los resultados obtenidos para el resto de conjuntos de datos.

En la Figura 3a y la Figura 3b que se muestran a continuación, exponen de manera gráfica los resultados mostrados de la tabla 4 respectivamente.

**Tabla 4.** Resultados obtenidos con los algoritmos de aprendizaje profundo para los conjuntos de datos EEG data for ADHD (EEG ADHD) y Working Memory and Reward in Children with and without ADHD (WMRC ADHD).

Conjunto	Algoritmo AP	Accuracy	Precision	Recall	F1-Score	ROC-AUC
EEG ADHD	CNN	0.5833	0.5833	0.5533	0.5833	0.5533
WMRC ADHD	Multi-modality 3D CNN	0.6833	0.6833	0.6833	0.6538	0.6833

Los resultados obtenidos para los distintos conjuntos de datos analizados varían de significativamente según la procedencia de los datos, por lo que es un aspecto importante a considerar para la fiabilidad de los resultados.

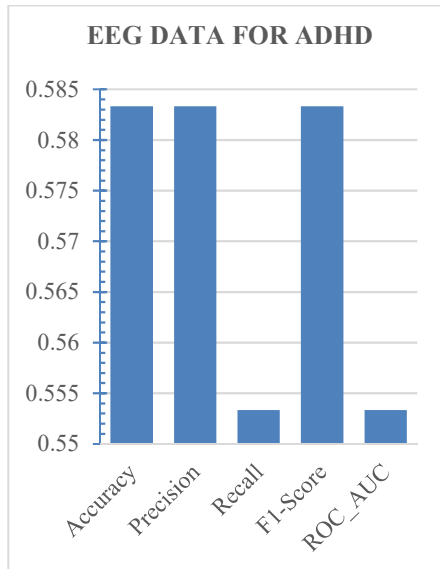
## 5. Discusión

Los resultados obtenidos durante el análisis realizado arrojaron un rango de precisión del 57% al 91%, el rendimiento de cada modelo varía significativamente según el conjunto de datos utilizado. Este resultado podría deberse a diferentes factores, como la calidad o naturaleza de los datos, la distribución de las clases, la cantidad de datos de entrenamiento y la complejidad del modelo utilizado. Es posible que el modelo tenga un alto nivel de clasificación, pero aun así tenga problemas con la capacidad de generalización a nuevos datos. Por lo tanto, es importante tener en cuenta el resultado arrojado en otras métricas, como recall, F1-score o la curva ROC.

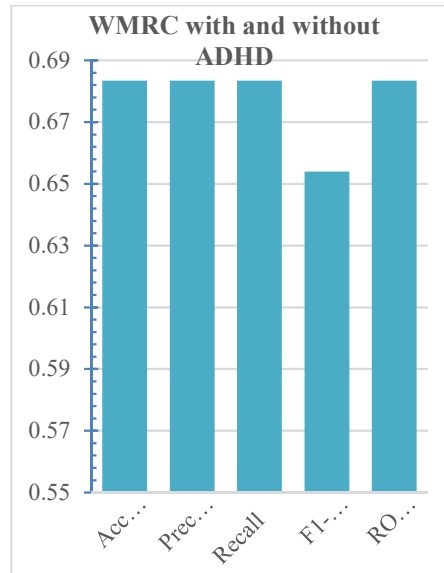
Ahora, profundizando más en los resultados obtenidos, se tiene que los algoritmos Regresión Lineal, CatBoost y SVM fueron los que arrojaron más del 85% de exactitud en el conjunto de datos HYPERAKTIV; mientras tanto que el conjunto de datos ADHD200 tuvo valores arriba del 97% de exactitud con los algoritmos de KNN, CatBoost y AdaBoost. Para el resto de los conjuntos de datos (EEG data for ADHD y Working Memory and Reward in Children with and without ADHD) fueron los menos precisos, arrojando resultados de entre el 58% y 68% respectivamente.

Es importante mencionar que debe considerarse la aplicación de un estándar para los conjuntos de datos relacionados al TDAH para garantizar la precisión y confiabilidad de los datos; además de hacer la colección de datos pública para futuras investigaciones y aún más si los datos fuesen multimodal, es decir, si la recolección de información de distintos estudios aplicados, como por ejemplo, imágenes fMRI, datos EEG y actividades que midan la concentración e hiperactividad de un mismo paciente se combinasen para favorecer aún más el desarrollo de herramientas con inteligencia artificial y de la exactitud resultante de las mismas para mejorar la detección, el diagnóstico y el tratamiento del TDAH.

Otro tema a discutir aplica a la selección de otros algoritmos a los aplicados aquí, considerando sus fortalezas y debilidades. La inclusión de otras técnicas de validación para evaluar su precisión y su comparación con los modelos tradicionales. Por último, considerando los resultados de este análisis, se determina que los biomarcadores más óptimos para la detección del TDAH son la combinación de: 1) datos de ondas cerebrales (obtenidos por EEG y que mide la concentración del individuo), 2) el registro de la actividad física y 3) ritmo cardíaco.



**Fig. 3a.** Gráfica de los resultados vistos en la tabla 4 (EEG DATA FOR ADHD).



**Fig. 3b.** Gráfica de los resultados vistos en la tabla 4 (Working Memory and Reward in Children with and without ADHD)

No obstante, la variable relacionada con fMRI no se encuentra descartada, ya que el conjunto de datos utilizado contiene muy pocos registros para dar un diagnóstico preciso; pero en un futuro, con la obtención de un mayor número de registros, se podría obtener un mayor grado de precisión.

## 6. Conclusiones

La falta de precisión en los diagnósticos para la detección del TDAH evita que las personas con este trastorno no reciban el tratamiento adecuado a tiempo, generando problemas en su entorno social, escolar o laboral. En este artículo se analizaron diferentes conjuntos de datos con distintas técnicas de aprendizaje automático y aprendizaje profundo para identificar qué algoritmos ofrecen los mejores resultados, así como también los biomarcadores más relacionados para la identificación del TDAH.

Durante el desarrollo de este análisis se detectó que la exactitud en la detección puede variar significativamente según el conjunto de datos utilizado y las técnicas de inteligencia artificial utilizadas. Los algoritmos con mayor puntaje obtenido para esta métrica fueron SVM, Regresión Lineal, KNN, Decision Tree y CatBoost, aplicados sobre los conjuntos de datos HYPERAKTIV y ADHD200, donde los resultados arrojaron un rango entre un 85% y 97% de exactitud.

El desafío de la detección del TDAH con técnicas de aprendizaje automático y aprendizaje profundo requiere la colaboración de especialistas médicos y técnicos en aprendizaje automático para desarrollar algoritmos precisos y eficientes.

En resumen, este análisis arrojó resultados prometedores, demostrando que las técnicas de aprendizaje automático y aprendizaje profundo son útiles en la detección del TDAH, pero que necesitan de un conjunto de datos con suficiente información y características específicas para realizar una correcta clasificación.

Como trabajo a futuro debe considerarse el desarrollo de un test estadístico con el fin de garantizar validez y fiabilidad de los resultados, así como para tomar decisiones informadas basadas en datos objetivos.

También es importante considerar el uso de más conjuntos de datos similares o diferentes para complementar la información, también se considera el posible uso de otras técnicas de inteligencia artificial a las utilizadas en este trabajo, considerar también evaluar otro tipo de métricas como el tiempo de respuesta y realizar la evaluación en otros lenguajes de programación como R.

**Agradecimientos.** Este trabajo de investigación fue patrocinado por el Consejo Nacional de Ciencia y Tecnología de México (CONACYT) y la Secretaría de Educación Pública (SEP) de México a través del programa PRODEP. Adicionalmente, se agradece al Tecnológico Nacional de México (TecNM) por apoyar este proyecto.

## Referencias

1. OMS: Salud mental del adolescente (2023) [www.who.int/es/news-room/fact-sheets/detail/adolescent-mental-health](http://www.who.int/es/news-room/fact-sheets/detail/adolescent-mental-health)
2. Centers for disease control and prevention: Trastorno por déficit de atención e hiperactividad (TDAH)
3. Rivera, F. B.: La elevada prevalencia del TDAH: posibles causas y repercusiones socioeducativas. *Psicología educativa*, vol. 22, no. 2, pp. 81–85 (2016) doi: 10.1016/j.pse.2015.12.002
4. Jakhar, D., Kaur, I.: Artificial intelligence, machine learning and deep learning: definitions and differences. *Clinical and experimental dermatology*, vol. 45, no.1, pp. 131–132 (2020) doi: 10.1111/ced.14029
5. Gimenez, M., Peláez, D. B., Fisac, J. E. O., McVeigh, E. R., Carbayo, M. L.: Desarrollo de una herramienta para la detección automática del plano valvular mitral mediante algoritmos de deep learning. In: Libro de Actas del XXXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica, pp. 33–36 (2018)
6. Wang, W., Lee, J., Harrou, F., Sun, Y.: Early detection of parkinson's disease using deep learning and machine learning. *IEEE Access*, vol. 8, pp. 147635–147646 (2020) doi: 10.1109/ACCESS.2020.3016062
7. Khullar, V., Salgotra, K., Singh, H. P., Sharma, D. P.: Deep learning-based binary classification of ADHD using resting state MR images. *Augmented Human Research*, vol. 6, no. 1, p. 5 (2021) doi: 10.1007/s41133-020-00042-y
8. Loh, H. W., Ooi, C. P., Barua, P. D., Palmer, E. E., Molinari, F., Acharya, U. R.: Automated detection of ADHD: Current trends and future perspective. *Computers in Biology and Medicine*, vol. 146, pp. 1–18 (2022) doi: 10.1016/j.combiomed.2022.105525
9. Arthi, K., Tamilarasi, A.: A hybrid model in prediction of ADHD using artificial neural networks. *International Journal of Information Technology and knowledge management*, vol. 2, No. 1, pp. 209–215 (2009)
10. Kuang, D., He, L.: Classification on ADHD with deep learning. In: International Conference on Cloud Computing and Big Data, pp. 27–32 (2014) doi: 10.1109/CCBD.2014.42

11. Chu, K., C., Huang, H. J., Huang, Y. S.: Machine learning approach for distinction of ADHD and OSA. In: IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 1044–1049 (2016) doi: 10.1109/ASONAM.2016.7752370
12. Miao, B., Zhang, Y.: A feature selection method for classification of ADHD. In: 4th International Conference on Information, Cybernetics and Computational Social Systems, pp. 21–25 (2017) doi: 10.1109/ICCSS.2017.8091376
13. Shao, L., Zhang, D., Du, H., Fu, D.: Deep forest in ADHD data classification. In: IEEE Access, vol. 7, pp. 137913–137919 (2019) doi: 10.1109/ACCESS.2019.2941515
14. Mao, Z., Su, Y., Xu, G., Wang, X., Huang, Y., Yue, W., Sun, L., Xiong, N.: Spatio-temporal deep learning method for ADHD fMRI classification. *Information Sciences*, vol. 499, pp. 1–11 (2019) doi: 10.1016/j.ins.2019.05.043
15. Tosun, M.: Effects of spectral features of EEG signals recorded with different channels and recording statuses on ADHD classification with deep learning. *Physical and Engineering Sciences in Medicine*, vol. 44, no. 3, pp. 693–702 (2021) doi: 10.1007/s13246-021-01018-x
16. Anuradha, J., Tisha, Ramachandran, V., Arulalan, K. V., Tripathy, B. K.: Diagnosis of ADHD using SVM algorithm. In: Proceedings of the Third Annual ACM Bangalore Conference, pp. 1–4 (2010) doi: 10.1145/1754288.1754317
17. Bautista, M. A., Hernández-Vela, A., Escalera, S., Igual, L., Pujol, O., Moya, J., Violant, V., Anguera, M. T.: A gesture recognition system for detecting behavioral patterns of ADHD. In: *IEEE Transactions on Cybernetics*, vol. 46, no.1, pp. 136–147 (2014) doi: 10.1109/TCYB.2015.2396635
18. Duda, M., Ma, R., Haber, N., Wall, D. P.: Use of machine learning for behavioral distinction of autism and ADHD. *Translational Psychiatry*, vol. 6, no. 2, pp. e732–e732 (2016) doi: 10.1038/tp.2015.221
19. Uluyagmur-Ozturk, M., Arman, A. R., Yilmaz, S. S., Findik, O. T. P., Genc, H. A., Carkaxhiu-Bulut, G., Yazgan, M. Y., Teker, U., Cataltepe, Z.: ADHD and ASD classification based on emotion recognition data. In: 15th IEEE International Conference on Machine Learning and Applications, pp. 810–813 (2017) doi: 10.1109/icmla.2016.0145
20. Itani, S., Lecron, F., Fortemps, P.: A multi-level classification framework for multi-site medical data: Application to the ADHD-200 collection. *Expert Systems with Applications*, vol. 91, pp. 36–45 (2018) doi: 10.1016/j.eswa.2017.08.044
21. Mohammadhasani, N., Fardanesh, H., Hatami, J., Mozayani, N., Fabio, R. A.: The pedagogical agent enhances mathematics learning in ADHD students. *Education and Information Technologies*, vol. 23, pp. 2299–2308 (2018) doi: 10.1007/s10639-018-9710-x
22. Khanna, S., Das, W.: A novel application for the efficient and accessible diagnosis of ADHD using machine learning. In: 2020 IEEE/ITU International Conference on Artificial Intelligence for Good, pp. 51–54 (2020) doi: 10.1109/AI4G50087.2020.9311012
23. Christiansen, H., Chavanon, M. L., Hirsch, O., Schmidt, M. H., Meyer, C., Müller, A., Rumpf, H. J., Grigorev, I., Hoffmann, A.: Use of machine learning to classify adult ADHD and other conditions based on the Conners' Adult ADHD Rating Scales. *Scientific reports*, vol. 10, no. 1, pp. 1–10 (2020) doi.org/10.1038/s41598-020-75868-y
24. Maniruzzaman, M., Shin, J., Hasan, M. A. M.: Predicting children with ADHD using behavioral activity: A machine learning analysis. *Applied Sciences*, vol. 12, no. 5, p. 2737 (2022) doi: 10.3390/app12052737
25. Hicks, S. A., Stautland, A., Fasmer, O. B., Førland, W., Hammer, H. L., Halvorsen, P., Mjeldheim, K., Oedegaard, K. J., Osnes, B., Gjaever-Syrstad, V. E., Riegler, M. A., Jakobsen, P.: HYPERAKTIV: An activity dataset from patients with attention-deficit/hyperactivity disorder. In: Proceedings of the 12th ACM Multimedia Systems Conference, pp. 314–319 (2021) doi: 10.1145/3458305.3478454
26. Kaur, A., Kahlon, K. S.: Accurate identification of ADHD among adults using real-time activity data. *Brain Sciences*, vol. 12, no. 7, p. 831 (2022) doi: 10.3390/brainsci12070831
27. Nichols, N.: ADHD200 (2023) data.world/nicholsn/adhd-200

28. Zou, L., Zheng, J., Miao, C., Mckeown, M. J., Wang, Z. J.: 3D CNN based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI. *IEEE Access*, vol. 5, pp. 23626–23636 (2017) doi: 10.1109/ACCESS.2017.2762703
29. Chen, M., Li, H., Wang, J., Dillman, J. R., Parikh, N. A., He, L.: A multichannel deep neural network model analyzing multiscale functional brain connectome data for attention deficit hyperactivity disorder detection. *Radiology: Artificial Intelligence*, vol. 2, no. 1, pp. 1–9 (2019) doi: 10.1148/ryai.2019190012
30. Booth, J. R., Cooke, G. E., Gayda, J., Hammer, R., Lytle, M. N. Stein, M. A., Tennekoon, M. M.: Working memory and reward in children with and without attention deficit hyperactivity disorder (ADHD). *OpenNeuro* (2023) doi: 10.18112/openneuro.ds002424.v1.1.1
31. Lytle, M. N., Hammer, R., Booth, J. R.: A neuroimaging dataset on working memory and reward processing in children with and without ADHD. *Data in Brief*, vol. 31, p. 105801 (2020) doi: 10.1016/j.dib.2020.105801
32. Booth, J. R., Cooke, G. E., Gayda, J., Hammer, R., Lytle, M. N., Stein, M. A., Tennekoon, M.: Working memory and reward in adults. *OpenNeuro* (2023) doi: 10.18112/openneuro.ds002687.v1.1.1
33. Hammer, R., Cooke, G. E., Stein, M. A., Booth, J. R.: Functional neuroimaging of visuospatial working memory tasks enables accurate detection of attention deficit and hyperactivity disorder. *NeuroImage: Clinical*, vol. 9, pp. 244–252 (2015) doi: 10.1016/j.nicl.2015.08.015
34. Hammer, R., Tennekoon, M., Cooke, G. E., Gayda, J., Stein, M. A., Booth, J. R.: Feedback associated with expectation for larger-reward improves visuospatial working memory performances in children with ADHD. *Developmental Cognitive Neuroscience*, vol. 14, pp. 38–49 (2015) doi: 10.1016/j.dcn.2015.06.002
35. Nasrabadi, A. M., Allahverdy, A., Samavati, M., Reza-Mohammadi, M.: EEG data for ADHD/control children. *IEEE Dataport* (2020) doi: 10.21227/rzfh-zn36
36. Mohammadi, M. R., Khaleghi, A., Nasrabadi, A. M., Rafieivand, S., Begol, M., Zarafshan, H.: EEG classification of ADHD and normal children using non-linear features and neural network. *Biomedical Engineering Letters*, vol. 6, pp. 66–73 (2016) doi: 10.1007/s13534-016-0218-2
37. Barua, P. D., Dogan, S., Baygin, M., Tuncer, T., Palmer, E. E., Ciaccio, E. J., Acharya, U. R.: TMP19: A novel ternary motif pattern-based ADHD detection model using EEG signals. *Diagnostics*, vol. 12, no. 10, p. 2544 (2022) doi: 10.3390/diagnostics12102544
38. Mahesh, B.: Machine learning algorithms-a review. *International Journal of Science and Research*, vol. 9, pp. 381–386 (2020) doi: 10.21275/ART20203995
39. Su, X., Yan, X., Tsai, C. L.: Linear regression. *Wiley interdisciplinary reviews: Computational Statistics*, vol. 4, no. 3, pp. 275–294 (2012) doi: 10.1002/wics.1198
40. Bentéjac, C., Csörgő, A., Martínez-Muñoz, G.: A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, vol. 54, pp. 1937–1967 (2021) doi: 10.1007/s10462-020-09896-5
41. Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., Farhan, L.: Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data*, vol. 8, pp. 1–74 (2021) doi: 10.1186/s40537-021-00444-8